



UNIVERSITAT DE
BARCELONA

Computer-Aided Drug Design applied to marine drug discovery

Disseny de fàrmacs assistit per ordinador aplicat
a la cerca de possibles fàrmacs marins

Laura Llorach Parés

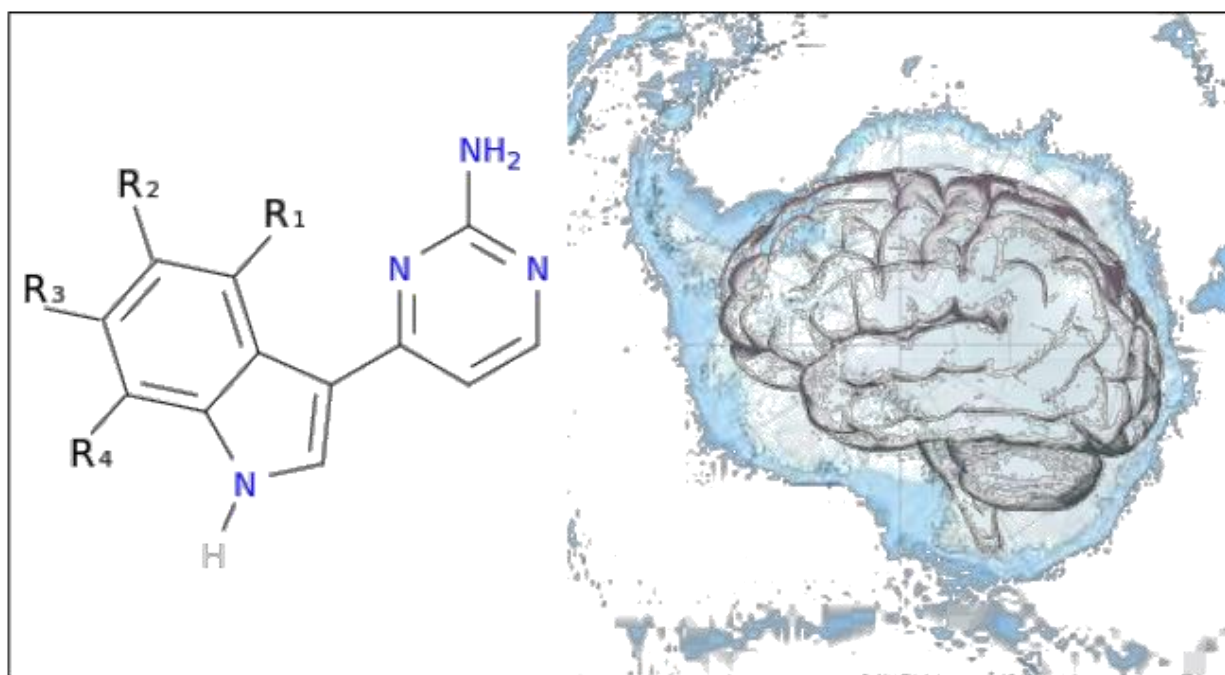


Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – SenseObraDerivada 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – SinObraDerivada 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-NoDerivs 4.0. Spain License.**

Computer-Aided Drug Design applied to marine drug discovery



Doctoral Thesis
Laura Llorach Pares
Barcelona, 2019



UNIVERSITAT DE
BARCELONA

Programa de Doctorat en Biologia Marina

Departament de Biologia evolutiva, ecologia i ciències ambientals

Computer-Aided Drug Design applied to marine drug discovery

Disseny de fàrmacs assistit per ordinador aplicat a la cerca de
possibles fàrmacs marins

Memòria presentada per

Laura Llorach Pares

per optar al grau de

Doctora per la Universitat de Barcelona

Laura Llorach Pares

Barcelona, 2019

Vist i plau de la directora de la tesi

Dra. Conxita Avila Escartín

Catedràtica del departament de Biologia
evolutiva, ecologia i ciències ambientals

Facultat de Biologia Universitat de Barcelona

Vist i plau del co-director de la tesi

Dr. Melchor Sánchez Martínez

Director científic Mind The Byte.SL

“Who am I? Who am I?
I am Jean Valjean!”

Les Misérables, Victor Hugo



Agraïments / Acknowledgements

Portava temps pensant com escriure aquest capítol. Per una banda, pensava en aquelles persones a les que volia agrair que haguessin format part d'aquest procés que és fer una tesi. I per l'altra, em plantejava no fer agraïments i dedicar-li al meu fill, en Roger.

Com podeu veure, m'he acabat decantant per fer un agraïment general, ja que han sigut unes quantes les persones a les que necessito dedicar unes paraules.

Voldria començar pel principi de la història i agrair a la Mire (Mireia Rodríguez), amiga i companya d'aventures i viatges, que sense tu no hauria sigut prou valenta per llençar-me a fer la carrera a Portugal. I mira, vist en perspectiva, va ser un gran esforç, però ha valgut la pena. A tú, Pepe (Josep Ferrer), per dedicar tantes hores a fer-me entendre estadística i a la vegada, ensenyar-me a pensar. Sense les teves classes no hauria tret la carrera ni seria qui sóc ara.

A l'Alfons Nonell, de MtB, per donar-me l'oportunitat de fer les pràctiques a la seva empresa. Aquí va ser on va començar tot. Gràcies a les pràctiques del màster vaig començar a conèixer l'apassionant món de la química computacional i el disseny de fàrmacs. També una menció a tots els companys de MtB, tant amb els que hem compartit poc temps com aquells amb els que n'hem compartit més, que durant aquests anys hem viscut canvis, alguns bons i altres no tant, experiències, coneixements i alguna que altra cerveseta.

De MtB s'ha de fer una menció especial pel Mel, Melchor, encara recordo el primer dia que vaig venir a l'oficina per fer una entrevista. I els primers dies... No entenc com vares poder tenir tanta paciència i ensenyar-me tant i tan bé. De tu ho he après gairebé tot. Mentor i company. Et puc assegurar que recordo perfectament apuntar-me a la llibreta "ls -ltr" per que no entenia res del que feia. Això de Linux els primers dies va ser dur, molt dur. Després, amb el temps, he après a fer els meus scripts i bon servei que t'han fet a vegades, eh?! Durant aquests anys hem treballat amb projectes molt interessants on els dos hem après molt i, en el fons, crec que formavem un bon equip. Tots dos recordem i recordarem, un projecte que ens va costar suor i sang, però entre els dos el vàrem aconseguir tirar endavant amb perseverança i esforç. Com a co-director de tesis també t'he d'agrair l'esforç que has hagut de fer aquests últims dies de correccions per que tot estigués a punt. Per tot, gràcies.

A la Conxita, una dona amb les idees clares i apassionada per l'Antàrtida i les molècules marines. Gràcies al teu esperit investigador i al projecte BLUEBIO, vaig poder començar a conèixer les molècules marines, primer durant les pràctiques del màster amb les quals vàrem fer el treball de final de màster i després durant tot el doctorat, on aquestes molècules marines s'han acabat convertit en "les meves

molècules". Crec que ens ho vàrem passar força bé en la nostra estada a Creta; congrés, platjeta i algun imprevist que altre amb el tema aeroport. Ens va servir per coneixe'ns una miqueta més, ja que al fer un doctoral industrial, no ens hem pogut conèixer com tocava. De totes maneres, sempre has estat allà i ajudat en tot el que he necessitat. Conxita, un plaer haver-te tingut de directora de tesi!

Ara voldria agrair a la família el seu suport. Malgrat que crec que no han entès mai què estic fent, han intentat respectar-ho i en aquests últims temps, on m'he dedicat a escriure, han ajudat en tot el que han pogut. Als avis i àvies, gràcies per preguntar-me com tenia la feina i per intentar, una vegada i un altre, entendre a que em dedico i que faig. Podeu seguir preguntant, us ho seguiré intentant explicar el millor que puc.

Als meus pares, que tot i que crec que quan els hi vaig dir, primer, que faria la carrera a Portugal mentre treballava, després, que amb 30 anys deixava una bona feina per començar un màster i després que faria un doctorat mentre treballava, no varen ser els pares més feliços del món. Crec que per a ells, això de prioritzar la formació acadèmica a una feina estable els hi va costar de pair, però amb el temps, han vist que tan malament tampoc m'ha anat, no?! Us estimo, pares! Ja sabeu que sóc tossuda i que de vegades pensem diferent, però al final, cada un hem de fer el nostre camí, i el meu camí no està tan lluny del vostre.

Als meus sogres i sobretot a la meva sogra, que ha hagut de fer d'àvia durant moltes tardes per que jo em pogués dedicar a la tesis. Suegri! Gràcies per tot. Per fer d'àvia, que t'encanta, per cuidar-nos a tots, ajudar-nos a casa, fer-nos tants menjars i tant bons... Ets una gran àvia, una gran mare i una sogra increïble!

A l'Oriol, el meu mariduky! Carinyu, que això ja s'acaba i tot tornarà a la normalitat. Sé que has hagut d'aguantar el meu caràcter i mal humor, que puc arribar a ser molt pesada, ja ho sé. Que moltes vegades voldries fer les teves "cosas jueguiles" i no pots. Però has de saber que ets un gran home, ple de bondat i que treus el millor de mi. Tenir-te al costat es una sort i espero no perdre't. Gràcies per fer de marit i de pare tant bé. T'estimo!

Fill, a tú també t'estimo. No t'assabentes de gran cosa per què tot just ara faràs un any, i tú no ho saps, però has fet gairebé la meitat d'un doctorat! Una part a la panxa de la mama i l'altre aquí al meu costat. Ets el millor regal del món. Per tú, Roger, tota aquesta tesis i tot l'esforç, decisions i història que hi ha al darrera, per què res és senzill, per què tot val la pena.

Finalment, a tota aquella gent que no ha tingut una relació directa amb la tesi però que han estat sempre allà. Nenes, llop de neu, Patri. Gràcies pels ànims!

Contents

General Introduction and Objectives	1
General introduction	3
Biodiversity in marine ecosystems.....	3
Antarctic benthic ecosystems	4
Mediterranean benthic ecosystems.....	9
Marine natural products.....	11
Computer-Aided Drug Design and Discovery	12
Alzheimer's Disease.....	25
Objectives of this thesis	29
Supervisor's Report	31
Section I. Target elucidation from a computational approach	33
Chapter 1. <i>In silico</i> studies to find new therapeutic indications for marine molecules.	37
Abstract	38
Resum.....	39
Introduction	40
Results and discussion.....	43
Materials and Methods	58
Conclusions.....	61
Appendix.....	63
References	67
Section II. Elucidation of different pharmacophoric features of marine compounds and a precise <i>in silico</i> binding study	71
Chapter 2. Computer-aided drug design applied to marine drug discovery: Meridianins as Alzheimer's disease therapeutic agents	75
Abstract	76
Resum.....	77
Introduction	78
Results	81
Discussion	90

Materials and Methods	95
Conclusions.....	104
Appendix.....	106
References	113
Chapter 3. Kororamides, convolutamines, and indole derivatives as possible tau and dual specificity kinases inhibitors for Alzheimer’s Disease	121
Abstract	122
Resum.....	123
Introduction	124
Results and discussion.....	126
Materials and Methods	149
Conclusions.....	154
Appendix.....	157
References	162
Section III. A combined computational and experimental study.....	169
Chapter 4. Meridianins and lignarenones as potential GSK3β inhibitors and inducers of structural synaptic plasticity	173
Abstract	174
Resum.....	175
Introduction	176
Results and discussion.....	179
Materials and Methods	192
Conclusions.....	196
Appendix.....	198
References	200
General Discussion.....	207
General discussion.....	209
CADD potential in drug discovery	210
Protein kinase inhibitors and MNPs pharmacophoric properties	213
Pharmacokinetic properties to found hits from marine scaffolds.....	216

Concluding remarks and future perspectives	217
Final Conclusions	219
Final conclusions	221
Catalan Summary	223
Objectius d'aquesta tesi.....	225
Resultats	227
Discussió general	230
Conclusions finals	241
General References	243
General references	245
Appendix I	269
Appendix II	337

General Introduction and Objectives

General introduction

Biodiversity in marine ecosystems

Ecology is often described as the biology of the ecosystems. Margalef, in his book “Ecología”, defined ecology as “the study of systems to a level at which individuals or whole organisms can be considered elements of interaction among them or with a loosely organized environmental matrix” (Margalef, 1974). An ecosystem comprehends a group of living organisms, or communities, in conjunction with their abiotic components, all of them interacting as a system (Willis, 1997). The abiotic factors are chemical or physical components of the environment, such as water, light, or temperature. Ecosystems are controlled by internal factors, such as degradation and decomposition, or external factors, such as climate or topology. Marine ecosystems are the largest on Earth, covering more than 70% of the surface of the planet (UNESCO, 2017). They are formed by oceans, seas and nearshore systems, such as salt marshes and mudflats. **Marine biodiversity** is the result of life evolution for billions of years and it is of great interest and value in many senses, but for a long time it has been underestimated (Snelgrove, 2016). In fact, between one and two-thirds of marine species are considered to be not yet described (Appeltans et al., 2012). Among these huge amounts of marine species, as in terrestrial ecosystems, a wide net of interactions exists, generating all sorts of defences and protective systems to survive (Faulkner & Ghiselin, 1983; Van Donk et al., 2011). Marine organisms communicate through intra- and/or interspecific interactions, which are often regulated by natural products (NPs) (infochemicals) and comprise what it is known as **chemical ecology** (Dayton et al., 1994; Hay & Fenical, 1996). **Chemodiversity** is intrinsically related to biodiversity and is the result of constant organism-to-organism interactive process, and thus, high biodiversity and ecological interactions are linked to high chemical diversity (Barre, 2010; Núñez-Pons & Avila, 2015). The unique chemical diversity of NPs, and marine natural products (MNPs) in particular, has been for long a major source of drug candidates (Blunt et al., 2018a). Many marine invertebrates, due to the strong predation pressure, possess unusual bioactive compounds that are essential for them to survive, playing important roles in predatory and competitive interactions (Leal et al., 2012; Avila, 2016; Blunt et al., 2018a). Therefore, with the aim to understand the **pharmacological potential** of natural compounds from marine benthic invertebrates we develop here a computer-aided drug design study over a group of secondary metabolites from selected marine organisms.

Antarctic benthic ecosystems

Terra Australis, nowadays known as **Antarctica**, from the antithesis of Arctic and which means “opposite to the arctic” or “opposite to the north” (Hince, 2000), was the name given to this southern region of 14 million km² full of ice and icebergs. The waters of the **Southern Ocean** (SO) contain an incredible biodiversity that is still being described nowadays (De Broyer & Danis, 2011; De Broyer et al., 2016). The **Antarctic Circumpolar Current** (ACC) is a current that flows from west to east around Antarctica, thus isolating the frozen continent. The origins of Antarctica go back to about 25 million years ago, during the early Cenozoic, when the Antarctic region broke away from South America forming the Drake Passage, and provoking a gradual cooling, believed to have had a huge influence on the development of both ACC (Scher & Martin, 2006) and the **Polar Front** (PF; Figure 1) (Clarke et al., 2005). The PF has also historically been called the Antarctic Convergence, being an area separating warmer tropical air masses from colder polar air in the mid-latitudes (Gordon, 1971).

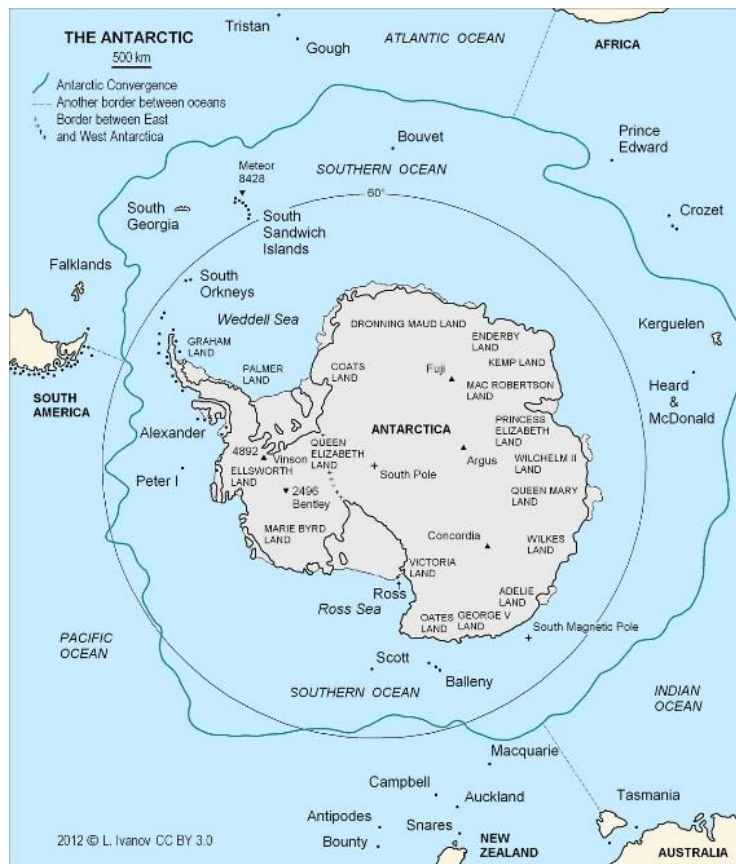


Figure 1. General map of Antarctica surrounded by the Southern Ocean (SO). Polar Front (PF), also called Antarctic Convergence, delineated in turquoise. Source: https://www.bugbog.com/maps/antarctic_circle_map/.

The PF constitutes a natural barrier that limits the exchange of cold and warm water, turning this area into a distinctive biogeographical region. This particularity let the Antarctic region in isolation, affecting the evolution of its fauna (Clarke & Crame, 1989). This is also associated with the high degree of endemisms found in Antarctic marine ecosystems (Clarke & Johnston, 2003). With the aim of classifying the SO biodiversity, the Register of Antarctic Marine Species (RAMS) published an accurated list of more than 8100 species, where an 88% are **benthic species** (De Broyer et al., 2011). The benthic community is exposed to considerable **predatory pressure** exerted by both macro- and micropredators (McClintock & Baker; Oshel & Steele; Dayton et al., 1994; Figuerola et al., 2013; Moles et al., 2015). Several studies demonstrated that most Antarctic benthic invertebrates present **natural products** in their crude organic extracts that act as feeding repellents to avoid predation, thus using different **defensive** chemical mechanisms (McClintock & Janssen, 1990; Amsler et al., 2001; Iken et al., 2002; Avila et al., 2008; Koplovitz et al., 2009; Slattery, 2009; Moles et al., 2015). Overall, the organisms living in Antarctic benthic ecosystems have developed very effective chemical defensive strategies, based on **secondary metabolites (natural products)**, which are crucial for species survival. Moreover, these unique chemical compounds can potentially be further exploited for the development of new drugs, considering its potential pharmacological properties (Núñez-Pons et al., 2015; Avila, 2016).

Natural products from Antarctic organisms have been reviewed several times recently (Avila et al., 2008; Moles et al., 2015; Blunt et al., 2018a). From polar organisms such as sponges, cnidarians, molluscs, bryozoans, and tunicates, a wide variety of biological compounds with diverse activities has been isolated, such as antitumorals, anti-bacterials, and anti-inflammatories (Avila, 2016; Tian et al., 2017). In this thesis I have worked with several compounds from diverse Antarctic benthic species, with special attention to the compounds listed below.

Sponges are organisms full of pores and channels, allowing water circulation through them. Their distribution is worldwide in all oceans, including tropical and polar regions with an approximate number of 5.000-10.000 known species (Bergquist, 2001). Sponges (Porifera) are very effective competitors for space, producing toxins and preventing other sessile organisms, such as ascidians, from growing on top or nearby, being one of the most diverse sources of bioactive natural products known (Proksch, 1994; Wang, 2006; Mehbub et al., 2014; Figueroa et al., 2015; Blunt et al., 2018a). Within the phylum Porifera the most diverse class is Demosponges (Figure 2), including 76,2% of all described species (WoRMS). The genera *Latrunculia* du Bocage, 1869, *Dendrilla* Lendenfeld, 1883, and *Aplysilla* Schulze, 1878 belong to Demosponges.



Figure 2. The Antarctic Demosponge *Latrunculia apicalis* (Ridley & Dendy, 1886). Adapted from WoRMS (WoRMS).

In order to study the pharmacological properties of marine compounds, three secondary metabolites isolated from diverse Antarctic sponges are included in this thesis (**Chapter 1**). These are **discorhabdin B**, an alkaloid from the species *Latrunculia apicalis* Ridley & Dendy, 1886 (Yang et al., 1995), **dendrinolide**, a diterpenoid from *Dendrilla membranosa* (Pallas, 1766) (Fontana et al., 1997), and **polyrhaphin A**, another diterpenoid isolated from *Aplysilla polyraphis* Laubenfels, 1930 (Figure 3).

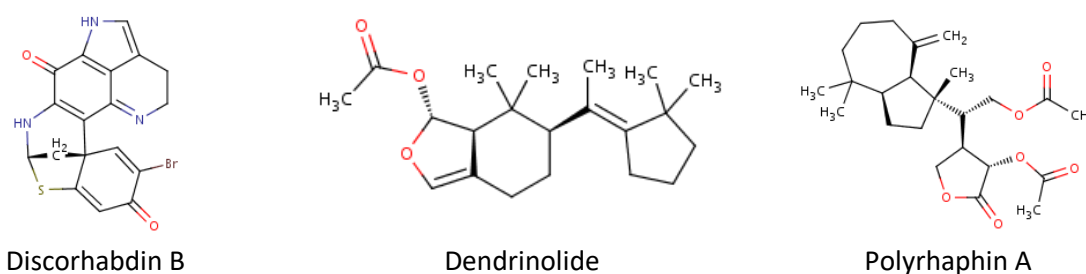


Figure 3. Chemical structure of compounds from the sponges *Latrunculia*, *Dendrilla* and *Aplysilla*: discorhabdin B, dendrinolide, and polyrhaphin A.

Mollusca, with around 85.000 described species, is the second largest phylum of invertebrates and the largest marine phylum, representing an enormous diversity of species; they can also live in freshwater and terrestrial habitats (WoRMS; Rosenberg, 2014). The most diverse class is Gastropoda, with around 70.000 species, some of them with commercial interest as human food sources. Marine gastropods have been of great interest also for their astonishing natural products and amazing chemical ecology. In fact, some drug leads from gastropods are currently in clinical trials, despite less than 1% of the molluscan secondary metabolites have been investigated so far (Avila, 2006; Benkendorff, 2010).

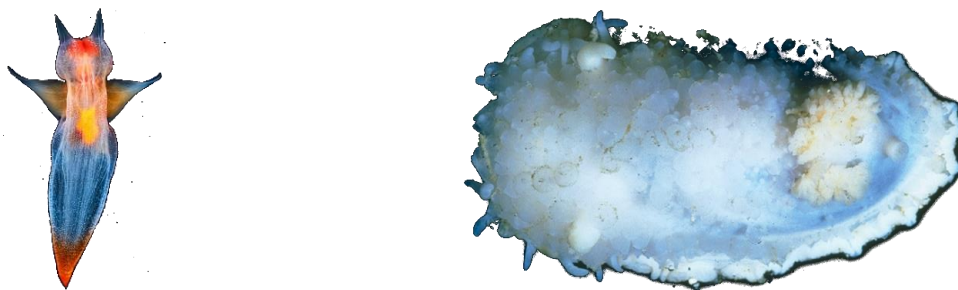


Figure 4. The gorgeous Pteropod *Clione limacina* (Phipps, 1774) is shown on the left. Adapted from www.hiveminer.com. And the Heterobranch *Bathydoris hodgsoni* Eliot, 1907 is shown on the right. Photo by C. Avila.

With the aim of elucidating the pharmacological potential of molluscan secondary metabolites, in this thesis I study two different species of gastropods, belonging to the genera *Bathydoris* Bergh, 1884 and *Clione* Pallas, 1774 (Figure 4). On **Chapter 1**, interesting results are reported about the pharmacological potential of **Hodgsonal**, a drimane sesquiterpene isolated from the mantle extract of the Antarctic heterobranch mollusc *Bathydoris hodgsoni* from the Weddell sea (Iken et al., 1998), and also of **Pteroenone**, a defensive metabolite belonging to the polyketide family, isolated from the pteropod *Clione limacina*, a shell-less pelagic mollusc collected in McMurdo Sound (Yoshida et al., 1995)(Figure 5).



Figure 5. Chemical structure of the secondary metabolites of the genera *Bathydoris* and *Clione*: Hodgsonal and Pteroenone.

Sea cucumbers are echinoderms from the class Holothuroidea (Figure 6), with rugged skin and long bodies. They can be found on the sea floor worldwide, with a number of described species of about 1.700. Sea cucumbers are known because of their defensive systems and thus, for the toxins they may contain. Diverse chemical active compounds have been identified from these animals (Khotimchenko, 2018) and in this thesis, the pharmacological potential of a metabolite of the genus *Staurocucumis* is analysed (**Chapter 1**). More precisely, **Liouvilloside**, a triterpeneglycoside isolated from the Antarctic sea

cucumber *Staurocucumis liouvillei* (Vaney, 1914) Ekman, 1927 collected near the sub-Antarctic Island of Bouvet (South Atlantic Ocean) was studied here (Antonov et al., 2008).

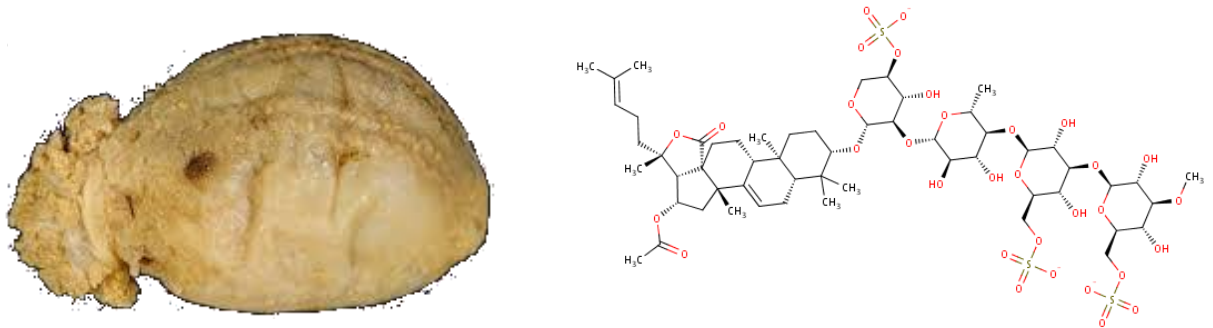


Figure 6. The Holothuroidea *Staurocucumis liouvillei* (Vaney, 1914) Ekman, 1927 and the chemical structure of the triterpene glycoside, Liouvilloside (Antonov et al., 2008). Adapted from WoRMS (WoRMS).

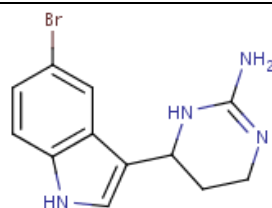
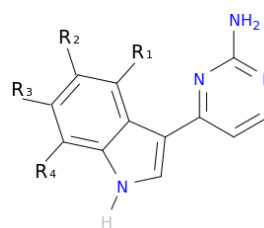
Ascidians are marine animals living in all oceans, they are usually sessile ciliary-mucus filter feeders, and comprise more than 2800 described species (Lambert, 2005). Ascidians or sea squirts belong to the subphylum Tunicata of sac-like marine invertebrate filter feeders and are usually cylindrical animals (Holland, 2016). Most ascidians' metabolites have been isolated from whole-body extractions but their complex organized body-plan and circulatory systems in comparison with other sessile invertebrates, may allow them to encapsulate bioactive compounds to avoid toxicity (López-Legentil et al., 2006; Núñez-Pons et al., 2012a). Within the class Ascidiacea, one of the most prolific genus, with forty species described from the SO, is *Aplidium* Savigny, 1816 (Figure 7). From these colonial genus several very interesting bioactive compounds have been obtained, including meridianins, aplicyanins, and rosinones (Franco et al., 1998; Reyes et al., 2008; Appleton et al., 2009; Šíša et al., 2009; Carbone et al., 2012; Núñez-Pons et al., 2012b).



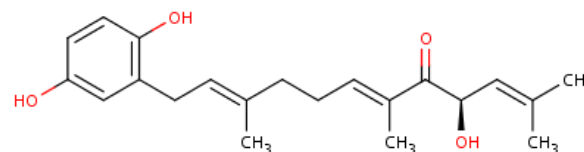
Figure 7. *Aplidium meridianum* (Sluiter, 1906). Adapted from Maggioni and collaborators (Maggioni et al., 2018).

Along this thesis, I have studied nine different bioactive natural compounds of the genus *Aplidium*, collected at the Eastern Weddell Sea, Antarctica (Núñez-Pons et al., 2012a), trying to elucidate some of their pharmacological properties (**Chapter 1**). We selected the seven indole alkaloids **meridianins A-G**, one brominated indole, **aplicyanin**, and one meroterpenoid, **rossinone B** (Figure 8).

Meridianin A	R1 = OH, R2 = H, R3 = H, R4 = H
Meridianin B	R1 = OH, R2 = H, R3 = Br, R4 = H
Meridianin C	R1 = H, R2 = Br, R3 = H, R4 = H
Meridianin D	R1 = H, R2 = H, R3 = Br, R4 = H
Meridianin E	R1 = OH, R2 = H, R3 = H, R4 = Br
Meridianin F	R1 = H, R2 = Br, R3 = Br, R4 = H
Meridianin G	R1 = H, R2 = H, R3 = H, R4 = H



Aplicyanin



Rossinone B

Figure 8. Chemical structure of the natural products of the genus *Aplidium*: meridianins A-G, aplicyanin, and rossinone B.

Mediterranean benthic ecosystems

The Ancient Greeks once call it “The Great Sea” and around 6th centuries after the Romans, the term *Mare Mediterraneum*, which means “in the middle of the land”, was first used to name the **Mediterranean Sea**. It covers an approximate area of 2.5 million km², and it is connected to the Atlantic

Ocean via a narrow strait of 14 km, called the Strait of Gibraltar. Cool waters from the Atlantic Ocean enter through the Strait, and their low salinity turns the waters circulation westward along the North African coasts till the Levantine Sea, where it starts to circulate eastwards along the Greek and South Italian coasts (Millot & Taupier-Letage, 2005). Before exiting the Mediterranean Sea through the depths of the Strait of Gibraltar, the seawater circulates along Italian, French and Spanish coasts (Millot, 1989). It has been calculated that this circulation process in the Mediterranean Sea takes around 100 years, which makes this sea very exposed to climate change (Millot, 1989). Overall, these low currents favourably affect the biodiversity of the Mediterranean waters (MWs) building a stable and rich ecosystem, estimated to contain between 4% and 18% of the world's marine species (Bianchi & Morri, 2000) (Figure 9).

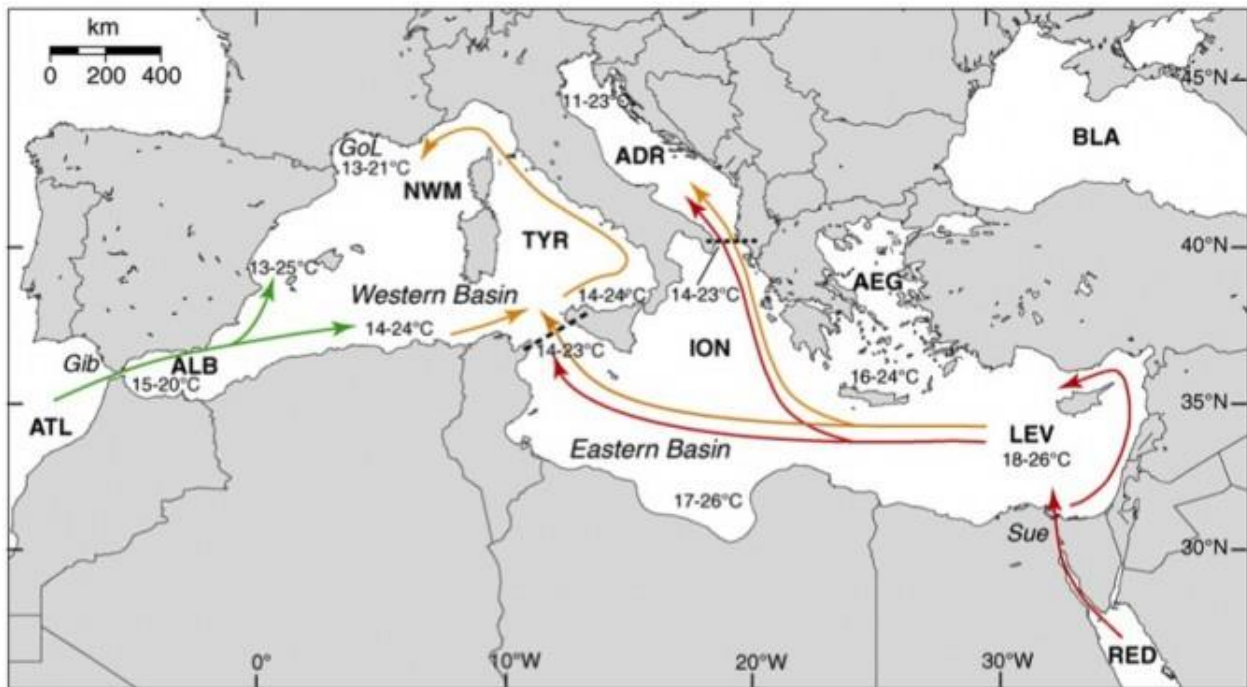


Figure 9. General map of the Mediterranean Sea and its water masses circulation. Source: <https://www.greenprophet.com/2010/08/mediterranean-garbage-patch/mediterranean-temp-mao-2/>.

From a chemical point of view, the bioactivity related to the natural products found in the organisms living in the Mediterranean Sea is, by far, less studied than in the Atlantic or Pacific regions (Uriz et al., 1991; Leal et al., 2012; Blunt et al., 2018a). As mentioned above, it is now known that marine invertebrates are involved in a wide variety of interactions, most of them chemically mediated (Paul et al., 2006; Egan et al., 2008; Puglisi et al., 2014). Due to this, as said, marine invertebrates are a potential

source of **bioactive natural products** which can act as protective defence against predators, and that may be further investigated for therapeutic aims.

In this thesis I have also analysed two secondary metabolites from the genus *Scaphander* Monfort, 1810, due to the pharmacological potential showed by gastropods (Carté, 1996; Pati et al., 2015) (Figure 10). The Mediterranean Cephalaspidean *Scaphander lignarius* (Lineé 1758) is a marine heterobranch gastropod mollusc inhabiting European coasts, from Iceland and Norway to the Mediterranean Sea (Cutignano et al., 2012). In the particular case of this thesis, the mollusc was collected in Blanes (Mediterranean coast of Catalonia) (Cutignano et al., 2008). *S. lignarius* typical metabolites are the lignarenones, a family of phenyl containing polyketides, particularly, **lignarenone A** and **lignarenone B** (Figure 10).

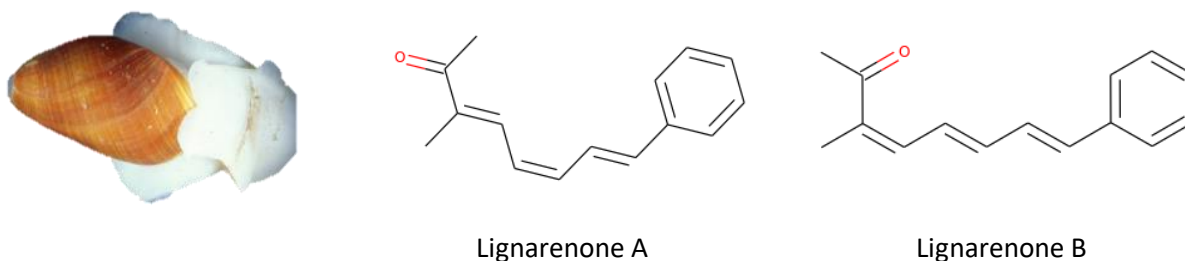


Figure 10. The Mediterranean heterobranch *Scaphander lignarius* (Lineé 1758) and the chemical structure of its two secondary metabolites: lignarenone A and lignarenone B. Adapted from the Bluebio team.

Marine natural products

Historically, **natural products** (NPs) have been widely studied from diverse disciplines ranging from ecology to pharmacy, where they are of capital importance. NPs exhibit a wide range of relevant biochemical features, such as specific scaffolds and pharmacophoric patterns, which are an invaluable source, for instance, for natural-product-inspired drug design and chemical synthesis, as well as for other disciplines such as nutrition, playing an important role in chemical sciences, with a frequent application in health too (J. Li & Vederas, 2009; Harvey et al., 2015; Rodrigues et al., 2016). Organic compounds from **terrestrial natural products** (TNPs), originating from terrestrial plants, microorganisms, vertebrates and invertebrates, have been extensively used in the past and present for the treatment of many diseases, as well as used as templates for synthetic design (Chin et al., 2006; Dias et al., 2012). Probably, two of the most famous examples are “aspirin”, acetylsalicylic acid which is an anti-inflammatory agent isolated from the willow tree *Salix alba* L. 1753, and morphine, isolated from *Papaver somniferum*, L., 1753 (Dias et al., 2012).

Due to their evolution and biodiversity, marine ecosystems are yielding more NPs than their terrestrial counterparts (Tringali, 2012). Taking into account that marine species constitute nearly half of the total planet biodiversity, this opens up the possibility to discover **potential therapeutic agents** from marine NPs (Thakur et al., 2005; Baker, 2015). In the past 40 years, but specially in the last two decades, the role of **marine natural products** (MNPs) in drug discovery has emerged as a hot research line (Newman et al., 2000; Newman & Cragg, 2016; Molinski et al., 2009a; Baker, 2015). Due to the long evolutionary processes and the specific conditions found on the seas and oceans, like the Mediterranean but especially in Antarctica, together with the known capacity of NPs to bind proteins (Breinbauer et al., 2002), MNPs represent a potentially huge source of therapeutic compounds with a high potency and selectivity (Paterson & Anderson, 2005; Baker, 2015). The identification of NPs that are capable of **modulate protein functions** in pathogenesis-related pathways, is one of the most promising lines followed in drug discovery (Koehn & Carter, 2005; Folmer et al., 2008; Cragg et al., 2009). Taking into account all these features, we may say that MNPs are optimized biologically active metabolites which can be used as a template to **design drug-like compounds** (Paul et al., 2006; Puglisi et al., 2014; Núñez-Pons et al., 2015; Prachayasittikul et al., 2015; Avila, 2016). Two recent examples of the uses of MNPs to develop drugs for human health are ziconotide and trabectedin. Ziconotide (commercialized as “Prialt”), approved by the United States Food and Drug Administration (FDA) in 2004, is used now for the treatment of chronic pain (Atanassoff, 2000; Reig & Abejón, 2009). More recently, in 2007, PharmaMar launched to the market trabectedin (commercialized as “Yondelis”), the first marine cancer medication approved by the European Medicines Agency (EMA) (European Medicines Agency, 2007). In addition to these two, there are other antiviral, cytostatic or antihyperlipidemic approved drugs and some others, around 20, are currently in clinical trials (Lindequist, 2016), such as the recently described SYL1801 of Sylentis (PharmaMar), whose application in eye drops could be a new therapeutic option for the treatment of retinal diseases characterized by neovascular processes, as announced in May 2019.

Computer-Aided Drug Design and Discovery

The action of identifying new molecules with certain therapeutic activity is known as **drug discovery**. The discovery and development of new drugs, for instance **small molecules** or peptides that inhibit the function of a protein (biomolecular target) related to a particular pathological pathway, is a complicated procedure that requires a lot of human and economic resources. In 2016, the cost of a new

molecular entity (NME), from the basic research until approved by the FDA as a final product, was estimated to be 1-1.8\$ billion and took around 12-15 years (Hughes et al., 2011; DiMasi et al., 2016). Due to these high investments in terms of human and economic resources, the low efficiency of some drugs, and the elevated failure rate in the drug discovery process, new methods and approaches have been developed to try to solve or to reduce these problems (Ou-Yang et al., 2012). From the last thirteen years, **computer-aided drug design** (CADD) has been settled down as one of the main effective methods to tackle the aforementioned difficulties (Sliwoski et al., 2014). The use of computational approaches allows the rapid exploration of the chemical space with the aim of finding novel lead compounds and predict, for instance, if a given molecule can bind to a target, and if so, how strong will be the binding (Katsila et al., 2016). CADD methods are powerful tools capable to complement experimental approaches, such as high-throughput screening (HTS) techniques, reducing the number of molecules to test, thus limiting the number of experiments to carry out, and as a consequence, optimizing the use of research time and budget (Leelananda & Lindert, 2016). The use of CADD techniques is time to time getting more attention, and today, CADD has become an effective and indispensable tool in therapeutic development. This has happened because of the huge development that this field has suffered respect to the initial times, including new and better tools and also stronger pipelines that have been proved to work in several studies, due to the increasing knowledge of biological structures, and also to the increasing computer power, among other reasons. CADD techniques can cover several steps of the **drug discovery pipeline** (Figure 11). CADD can be used to predict effectiveness, possible side effects, to improve the bioavailability, or to perform compound optimization. CADD methods can be classified into two general category types. (I) **Structure-based drug design** (SBDD), is based on the knowledge of the 3-dimensional (3D) disease-related target protein structure through methods such as x-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy (Jhoti & Leach, 2007), although, if the structure of the target of interest is not available, CADD can help through homology modelling. (II) **Ligand-based drug design** (LBDD) relies mostly on the knowledge of small molecules that bind to the target of interest. Depending on the case, either LBDD or SBDD can be the principal approach, but very often SBDD and LBDD are used together as they complement each other. In SBDD, among the most used methods, docking calculations and molecular dynamics (MD) simulations are used (Durrant & McCammon, 2011; Meng et al., 2011). SBDD methods begin with the identification of possible binding sites, active sites, or allosteric cavities on the target surface with specific features as hydrogen donors and acceptors, hydrophobic characteristics, and size determination (Anderson, 2003). Thus, SBDD can be applied on the understanding of how the

orientation or the pose of a given molecule could interact with a biological target, ultimately elucidating the main pharmacophoric properties which exert a therapeutic effect. In LBDD, molecular similarity approaches, quantitative structure-activity relationship (QSAR) techniques, or pharmacophore modelling, are some of the most used techniques. LBDD methods, for instance, can be applied to perform virtual screening (VS) in order to find analogue compounds to a molecule of interest, or to find molecules that fulfill certain pharmacological, biological and/or toxicological properties, or to improve compound features through hit to lead (H2L) optimization cycles to develop drug-like compounds (Acharya et al., 2011; Lionta et al., 2014; Yu & Mackerell, 2017).

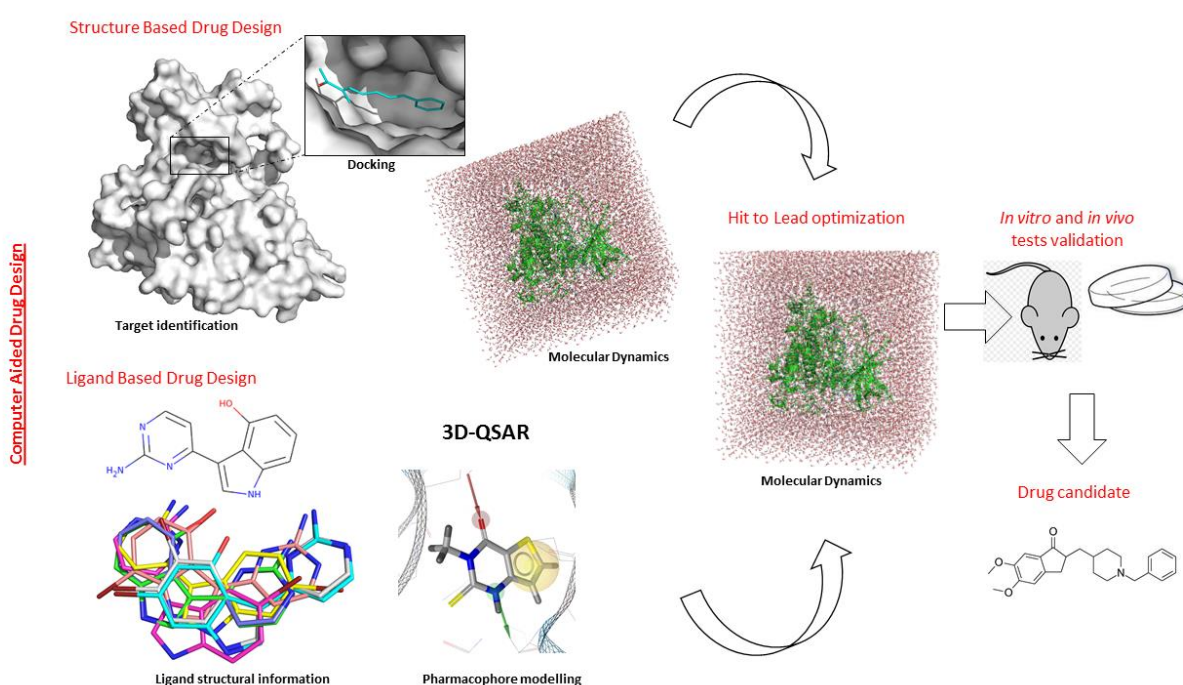


Figure 11. Schematic representation of a computer-aided drug design (CADD) pipeline. Original from the author.

It is always interesting to go a little bit deeper into all these computational techniques to understand the magnitude of the evolution that this field has suffered during the last years. To start with, and before describing any technique, it has to be noted that the protein structure determination is fundamental in computational biology, computational chemistry, chemo/bioinformatics, and/or computational biophysics fields. In CADD, specially for SBDD as commented above, having a good determined structure is, usually, proportionally related to the results obtained, as the structure

information is key for understanding the interactions between small molecules and the protein, which is an essential point on the drug discovery process (Reich & Webber, 1993).

Structure determination

One of the most successful and historically used approach for proteins and biological macromolecules structure determination is the **X-ray crystallography**, where a trustworthy source of protein needs to be available and purified until achieving a soluble material. Then, the protein must be crystallized, and the crystal obtained has to diffract to sufficient resolution to be processed (Smyth & Martin, 2000; Slabinski et al., 2007). **NMR spectroscopy** had his boom in the field of structural biology on the 2000s, starting to play a major role in the determination of structures and dynamics of proteins, and other biological macromolecules, since it allow determining protein conformations/ensembles at a resolution better than 2 Å (Cavanagh et al., 1995; Cavalli et al., 2007). This determination consists in several steps, including specialized techniques as quantum mechanical (QM) properties determination (Elyashberg, 2015). The major advantage of NMR spectroscopy over X-ray crystallography is that the determination can be done in solution, which allows for study of protein dynamics, and the difficulty of fixing the protein in a crystal disappears. This is a substantial improvement in comparison with X-ray crystallography. **Cryogenic electron microscopy (cryo-EM)** is planned to be the future of biological macromolecules structure determination, as it allows its study in native conditions at near atomic resolution while capturing multiple dynamic states (Murata & Wolf, 2018). In 2017, the Nobel Prize in Chemistry was awarded to Dubochet, Frank and Henderson “for developing cryo-electron microscopy for high-resolution structure determination of biomolecules in solution” (Dubochet, 2012; Henderson & McMullan, 2013; Chen & Frank, 2016; Frank, 2016). Unlike X-ray crystallography and NMR spectroscopy, cryo-EM requires a much smaller amount of sample and it allows to determine wide molecular mass range of proteins, from kilo-Daltons (protein complexes) to mega-Daltons (virus particles) (Murata et al., 2018). Nowadays this technique is becoming more used with time, producing better resolutions, better resolved structures, but it is not the most used or common technique yet. According to the statistics of the Protein Data Bank (PDB), a repository of information about the 3D structures of proteins, nucleic acids and complex assemblies, 90% of the protein structures were resolved by X-ray crystallography, 12.000 (of more than 120.000) by NMR, while the number of structures resolved by Cryo-EM is 3947, currently not comparable with the other two techniques. Nevertheless, in recent years, Cryo-EM is suffering a very high growth in the number of protein structure diposits (Liu et al., 2014).

Having a resolved structure of the target of interest is the first step for using SBDD techniques. However, it has to be taken into account that having a structure resolved, by any technique, does not mean to have a perfect picture of the protein structure and being ready to start using computation over it as a starting point for simulations. Usually, the structures present in the PDB do not fill the computational needs. This may be due to high atomic resolutions that are translated in a bad description of protein regions associated with non-natural, or lack, of secondary structure, or missing atoms from the aminoacid sequence. Also, often, the resolved protein structure does not include the region of interest. These “errors” can compromise the computational simulations, so they should be solved or at least reduced as much as possible. In that sense, a popular computational method used to alleviate this problem, when predicting the 3D coordinates of structures, is **homology modelling** (HM) also known as template-based protein modelling. It is mainly used to obtain structures whose coordinates are not available, or that are lacking some regions of interest. For other of the mentioned errors, like the presence of missing atoms, there are software tools like PDBFixer (PDBFixer, 2019), that help to fix protein structures. The principle behind HM is that evolutionary-related proteins often share similar structures, and this is because it is well known that the protein structure remains more conserved than the sequence during evolution (Lesk & Chothia, 1980; Illergård et al., 2009; Kaczanowski & Zielenkiewicz, 2010). Exploiting this fact, homology modelling relies on the identification of one or more known protein structures similar to the structure of a query sequence (or sequence of interest), making an alignment of those structures and mapping the shared regions/residues between both the query and the similar template. Using the retrieved information, a model is constructed and finally is evaluated using different criteria such as Ramachandran angles, sequence similarity or sequence coverage (Fiser, 2010).

Structure and ligand based applications

Once the 3D structure of a protein is known, finding its orthosteric pocket (active site) or additional binding pockets (allosteric cavities or just binding regions on its surface) is the next important step in SBDD. But before delving into this topic it is interesting to first explain LBDD, because it is important to understand that the computational drug discovery process is not linear and methods and techniques coming from structural or ligand sides, are highly complementary and their efficacy increases when they are used together. As said above, given the case, maybe a SBDD or LBDD approach can be better suited than a combination of both, but usually, mixing methods from both approaches increase the probability of success, as can be seen on **Chapters 1-4**. Here, we describe the main methods used in

this thesis, coming from LBDD and SBDD. Further information is available, for instance, in Sliwoski et al (2014) and Yu and Mackerell (2017).

Virtual screening (VS) is a computational technique used in drug discovery to search libraries of small molecules in order to identify those structures which are most likely to bind to a drug target, typically a protein receptor or enzyme. VS can be performed using structure or ligand-based techniques (Sliwoski et al., 2014; Gimeno et al., 2019). One of the main techniques used in **virtual (ligand) screenings** is **molecular similarity** (Willett, 2006; Eckert & Bajorath, 2007; Cereto-Massagué et al., 2015). It is used to score and ranking molecules according to their likelihood to another molecule(s), since it is a knowledge-driven approach which requires structural information of the bioactive ligand(s) of interest. Another important variant is based on **pharmacophore mapping**. The International Union of Pure and Applied Chemistry (IUPAC) defines a **pharmacophore** to be “an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response”. It also constitutes a central unit or a key scaffold of chemical compounds that should be preserved to design effective drugs (Wermuth et al., 1998). In drug discovery, pharmacophore features are widely used for VS, de novo design and/or lead optimization experiments (Yang, 2010).

The molecular-similarity VS method relies on the similarity-property principle, which states that similar molecules should exhibit similar properties (Klopmand, 1992). This technique is usually employed over large libraries and/or databases of compounds which contain diverse information associated to each molecule, such as binding targets or distribution profiles. Because of that, these methods have been highly used to elucidate the plausible targets, off-targets, or other pharmacological properties of the studied compounds. This can be done by correlating the structural similarity with the possibility of sharing a similar biological profile. This correlation is the idea behind the so-called Structure Activity Relationship (SAR) principle, first introduced in 1865 (Crum-Brown & Fraser, 1868; Blake, 1884), that derived into the so-called quantitative SAR (QSAR) methods. **QSAR methods** started to be used in the pharmaceutical context as an attempt to correlate chemical structure (2D and/or 3D) with activity using statistical approaches (Perkins et al., 2003). This was done with the aim of solving the problems they encountered in the late 1990s, where some studies started to point out that poor pharmacokinetics (PK) and toxicity predictions were an important cause of costly late-state failures in the drug development process (Van de Waterbeemd & Gifford, 2003). Actually, these methods have now become a common technique in the field. QSAR methods are now widely applied in drug discovery, especially on the study of absorption, distribution, metabolism, excretion, and toxicity (**ADMET**) properties (Gola et al., 2006).

VS techniques based on ligands are powerful and widely used approaches, but there also exists a structure-base counterpart, as commented before, which is the **docking-based virtual screening**. This method allows the scanning of thousands of proteins to identify potential targets for a single molecule or a library of compounds by using molecular docking calculations (see below) (Xu et al., 2018). Docking-based screenings constitute an important computational tool for identifying new targets of existing drugs and, especially, are highly valuable for predicting the bioactivity of a small molecule where the protein target is still unknown (Lapillo et al., 2019). The usage of these techniques is clearly explained and put in context in the following chapters (**Chapters 1-4**), where it is demonstrated how from a given chemical compound (or a set of them) these techniques can be applied, for instance, to elucidate possible targets (Toledo-Sherman & Chen, 2002; Shoichet, 2004), to determine their biological profile or to find similar compounds (Varney et al., 1992; Shoichet, 2004).

The work done in this thesis, as can be observed in the following chapters, has been mainly approached from a structural perspective, as it is mainly based on SBDD techniques (although LBDD methods have also been important to achieve the thesis objectives). Probably, the two main methods encompassed on SBDD are docking calculations and MD simulations. Because of that, and also as they are mostly used here, we proceed to describe them in more detail. These methods can be applied in different steps along the drug discovery pipeline, as seen in the following chapters.

Docking calculation process concerns the study and prediction of ligand conformation and orientation (pose) within a target binding site (Kitchen et al., 2004). This calculation was first mentioned in the early 1980s (Kuntz et al., 1982), and still today, is one of the most popular CADD tools used in drug discovery (De Vivo & Cavalli, 2017). The docking protocol can be described as a multi-step process full of complexity (Brooijmans & Kuntz, 2003), but basically involves two steps: the prediction of the binding pose and the evaluation of its strength. The procedure begins with the application of docking algorithms that facilitate the prediction of the best pose (including also ligand-target interactions) of a given small molecule in the orthosteric site or other allosteric binding region of the protein; thereafter the binding affinity of the protein-ligand complex is estimated (Meng et al., 2011).

Back to the history, the first explanation of binding was provided by Emil Fischer in 1894, describing the specific action of an enzyme with single substrate using the lock and key analogy (Fischer, 1894) (Figure 12).

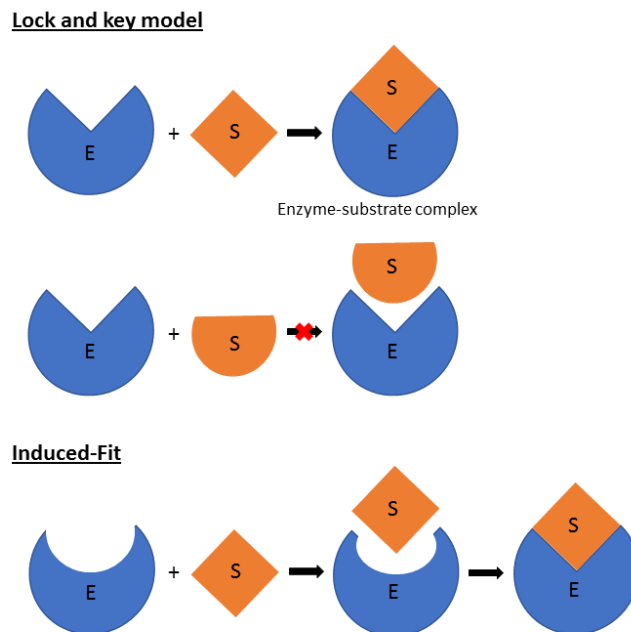


Figure 12. Fischer’s original “lock and key” model proposed in 1894 and Koshland “induce-fit” theory (Fischer, 1894; Koshland, 1963). E: Enzyme. S: Substrate. Original from the author.

Later on, as this lock and key hypothesis did not take into account the flexible nature of the protein, another theory was proposed, the so-called **induce-fit theory**, which refuses the idea that the substrate only fits into the active site, and proposes a continuous change in the conformation of the enzyme in response to the substrate binding (Koshland, 1963) (Figure 12). In agreement with this theory, both ligands and protein receptors should be considered as flexible entities during docking. However, probably the most used docking approach, the so-called classical or rigid docking, does not take this into account. This variant only allows the ligand movement fixing the target conformations. This represents a clear drawback mainly due to computer limitations resources, but also for the desire of preserving a certain protein conformation. Anyway, in general, the lack of protein movement is considered a limitation. To overcome this issue, in **flexible docking** the ligand and the receptor are allowed to move. There are different variants of flexible docking based on the way the intrinsic protein dynamics is incorporated into the equation. For instance, there are approaches where the receptor remains rigid with the exception of the side chains of selected residues which are allowed to move or even the receptor is fully flexible (Meng et al., 2011). Another approach to incorporate protein flexibility could be the use of **ensemble docking**, which consists of the generation of different conformations of the target experimentally (coming from NMR models or X-ray crystal structures), or computationally, generally, obtained by MD simulations (Amaro et al., 2018). Over these ensembles, classical docking

experiments are performed; however, as different protein conformations are considered, the flexibility is indirectly captured (Korb et al., 2012; De Vivo et al., 2017). Finally, a very useful method (explained in the following paragraphs) is the post-processing of classical docking calculations by MD simulations (De Vivo et al., 2016).

From the binding region perspective, there are two possible scenarios. 1) when the binding pocket is previously known, and **classical (rigid) and/or flexible docking** calculations are performed over it elucidating the preferential binding pose (Taylor et al., 2002) (Figure 13), and 2) when the binding cavity is not known and the protein surface has to be explored with the aim of founding plausible cavities (catalytic cavities, allosteric cavities or just binding regions where the ligand can be retained for a certain period of time) where the molecule can bind and exert some activity (Hetényi & van der Spoel, 2006) (Figure 13). After the elucidation of all possible pockets, classical or flexible docking techniques can be applied over them to determine which are the most favourable cavities and molecules poses. The whole process is usually known as **blind docking**. A clear example of the application and utility of blind docking calculations can be seen on **Chapter 1**. On **Chapter 4**, following the first approach of the blind docking methods, an exploration of the cavities was performed with the aim of elucidate plausible cavities. Besides, on **Chapters 1-4** the important and crucial contribution of docking techniques on the first steps of drug discovery process is shown.

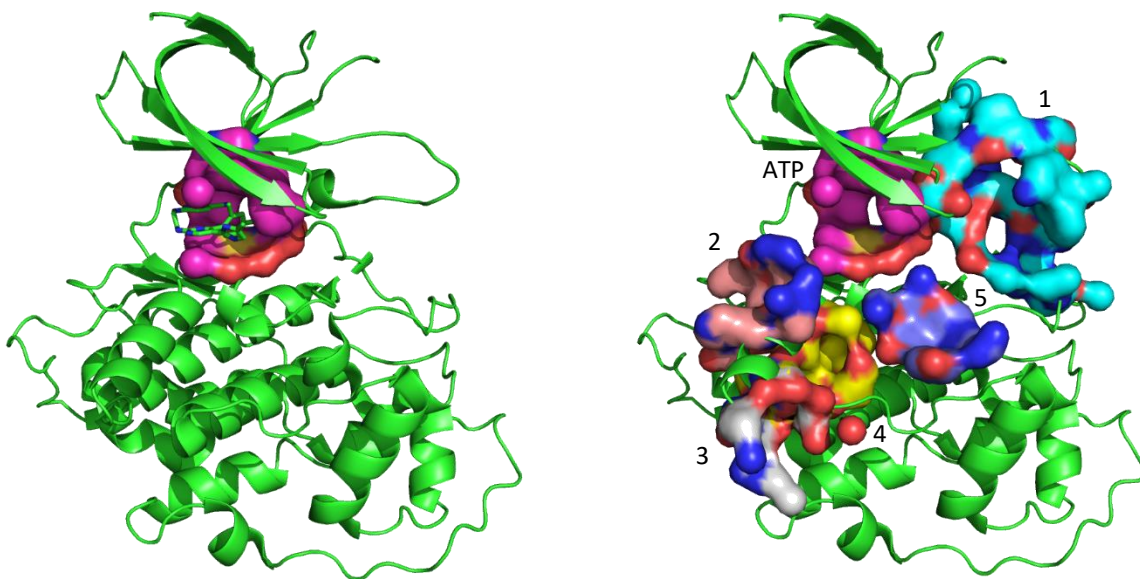


Figure 13. Graphical representation of crystallographic human structure of glycogen synthase kinase beta (GSK3 β) (Protein Data Bank ID (PDB) 6B8J) (Wagman et al., 2017). On the left GSK3 β structure with its crystallographic ligand 65C placed in the adenosine triphosphate (ATP) active site and where classical docking can be performed over. On the right, the ATP cavity and also other five proposed allosteric cavities (1-5) found after performing a search of the protein surface (blind docking). Adapted from Llorach-Pares, et al. (2019).

In all variants of docking calculations, and once the cavity or cavities of binding are known, predict the optimal placement for a molecule, given certain degrees of freedom, is by itself challenging, as a high accuracy is needed to identify the best possible conformational pose of the ligand that fits better the receptor structure. This step needs to be fast enough to allow the analysis of hundreds or thousands of compounds in the same run, and it is conditioned for the high number of degrees of freedom, which significantly increases the computation time and also the number of false-positive results (Andrusier et al., 2008). The following and complementary step is the prediction of the biological activity, in terms of binding energy by the use of scoring functions, and the subsequent evaluation of the interactions between the small molecule and the target (Meng et al., 2011). **Chapter 2** is a nice example where the aforementioned step process is put in context and helps to understand the applicability of binding interactions studies. To add more complexity to these methods, as said above, poor crystallographic resolution of targets, implicit flexibility (Koshland, 1963), induced fit events (Tobi & Bahar, 2005), and the water involvement on the target-molecule binding, make these type of calculations a scientifically complex process (Kitchen et al., 2004).

All organisms are regulated by a correct protein function. A malfunction of this regulation can result in some disease. Usually, the protein function is regulated by the binding of a substrate to the orthosteric cavity (active site). However, there are cases where other additional/alternative pockets can have this role. **Allosteric regulation**, an emerging concept in drug discovery in the last years (Abdel-Magid, 2015), can control protein function by the binding of small molecules, or other entities like peptides or even other proteins, to the target. It is used to be a single protein or protein complex, in a cavity at some distance (until tens of Å) from the orthosteric site (Laskowski et al., 2009; Amaro, 2017). However, a molecule binding to a cavity different from the principal one does not mean it is an allosteric cavity, because this depends on the effect that the compound can exert over the protein. This effect, in general, can be positive (activating), provoking an increase of the target protein activity, or negative (inhibiting), causing a decrease of the protein activity (although the scenario can be, in some cases and for certain proteins, a little bit more complex) (Tian et al., 2012; Morra & Colombo, 2018; Greener & Sternberg, 2018). Also, the molecules that bind on it, do not need to be chemically similar to the natural ligands as there is no competition between them (Laskowski et al., 2009). A nice example on how the allosteric modulation can inhibit a protein function can be found in **Chapter 4**, where two marine molecules are proposed to inhibit the activity of a kinase by binding to an allosteric pocket (called substrate cavity).

Focusing on protein activity, all proteins are intrinsically **dynamic/flexible** entities (Kim et al., 1993), and thus, its biological function/activity relies on their flexibility. To be more precise, the internal motions of proteins result in conformational changes, which are at a time, essential for their functions (Henzler-Wildman & Kern, 2007). The study of protein dynamic movements is necessary to understand the structure-function relationship (Quan et al., 2014), that, in fact, could be reformulated as structure-dynamics-function. Conformational changes on protein structures can be caused by protein-protein binding, ligand binding or post-translational modifications (Teilum et al., 2009), which can directly affect their function. Measuring, analysing, and understanding proteins dynamics and the associated conformational changes, is a must. In this regard, **MD simulations** are a versatile and powerful computational method widely used to obtain information on the time evolution of protein motions (Karplus, 2002; Adcock & McCammon, 2006). More precisely, MD simulations allow the study of the physical movement of atoms and molecules, ranging from simple systems of few atoms or just one small chemical compound to more complex scenarios like proteins, or chemical compounds bound to proteins. The atoms and molecules are allowed to interact for a fixed period of time, through the integration on Newton's laws of motion, constructing trajectories that allowed to describe the temporal evolution of the particles of a given system, and thus, to observe its dynamic evolution. There are several variants of this technique, some of them addressed to accelerate the dynamic process and span the time-scale. In order to do that, an option is to apply an external force, like targeted (TMD) (Schlitter et al., 1994), steered (SMD) (Suan & Khanh, 2013) or accelerated MD (AMD) (Hamelberg et al., 2004) methods.

The first MD simulation of a protein was carried out in 1974 by Andrew McCammon and Martin Karplus, and consisted in a 9.2 picoseconds (ps) trajectory of small globular protein, bovine pancreatic trypsin inhibitor (BPTI), in vacuum (McCammon et al., 1977). More than ten years were needed to report the simulation of the same protein but solvated in water (Levitt & Sharon, 1988). From that moment, computational power has been growing quite fast, thus allowing the performance of more complex simulations over time, that are also more "useful", as they can help to solve more complex problems related to diverse areas like biology, chemistry, or physics. Because of that, nowadays, the use of these methods is very popular in different fields, as it happens in drug discovery. Regarding the use of MD methods in drug discovery, its main advantage is that it allows to consider the structural dynamics/flexibility of the proteins, alone and/or in complex, for instance, with ligands, other proteins, or DNA. Unlike other static techniques, like rigid docking, this kind of simulations takes into account the

entropic effects and enables a more accurate estimation of the thermodynamics and kinetic association of target-ligand complexes (De Vivo et al., 2016).

MD, in addition to the characterization of the structural landscape of a protein, or a protein complex, and/or extracting conformational ensembles, is widely used, specially in drug discovery, to understand the ligand-target binding and unbinding mechanisms (De Vivo et al., 2016). In relation to that, one popular use of this technique is to post-processing docking calculations. As mentioned above, classical docking (despite being a reasonable good technique to predict the optimal placement of a ligand within a binding pocket, as it has a proven track record of success) has several limitations, as classical docking does not consider protein flexibility and the scoring functions used to have accuracy limitations. These limitations are usually translated into a bad description of the binding mode and the associated binding energy, and thus a wrong ranking of the analysed compounds (Kitchen et al., 2004). Flexible docking methods can improve the results of the rigid counterpart, but these variants still have a strong dependency on the scoring function. In that sense, MD simulations can optimize the predicted docking poses and also validate the stability of the docked complex (De Vivo et al., 2016; Aravindhana et al., 2017). If the docking pose is not “good” enough, it could be possible to see how the ligand leaves the binding site during the simulation (usually in hundreds of ps). This two-step protocol (docking+MD) constitutes a good approach to solve docking drawbacks, thus allowing us the prediction of, theoretically, more reliable protein-ligand binding modes (Alonso et al., 2006). The workflow combining docking calculations (that can be used to screen large compound libraries filtering out a significant part) and MD simulations (that despite being more computationally expensive, can be used efficiently, over the best docking poses), has been extensively used in the literature (Alonso et al., 2006; Aravindhana et al., 2017). Applying short post-processing MDs over hundreds of compounds is, nowadays, feasible in a short period of time (around a week in a desktop GPU), which reinforces this approach, since it is fast enough to be used regularly in any SBDD workflow. As a consequence of that, it is being increasingly used.

There are different variations of MD simulations in addition to the classical version, which is probably the most commonly employed. These variations can be used to understand the binding/unbinding mechanism of a ligand over a target of interest. In that sense, **Steered molecular dynamics** (SMD) is becoming a highly used method in drug discovery to describe the process of protein-target binding, giving insights into the binding/unbinding mechanisms (Patel et al., 2014; De Vivo et al., 2016, 2017). As explained before, external time-dependent forces are applied to the ligand in order to accelerate the disassociation of the protein cavity, revealing the force needed to cause the rupture

between the ligand and the receptor (Israelewitz, Baudry, et al., 2001; Israelewitz, Gao, et al., 2001). These forces can be theoretically correlated to the experimental residence time, and also, with its inhibitory capacity (Potterton et al., 2019). Moreover, during this process, it is possible to estimate which interactions are stronger and more necessary to keep the ligand bound.

From MD simulations, in general, a good deal of useful information can be extracted regarding the dynamics and thermodynamics of the studied system. One of the properties that can be measured is the **binding free energy** of target-ligand complexes. This energy is estimated, according to the thermodynamic cycle shown in Figure 14, as by subtracting the free energies of the ligand and the protein in aqueous solution to the free energy of the complex (protein-ligand) (Miller et al., 2012).

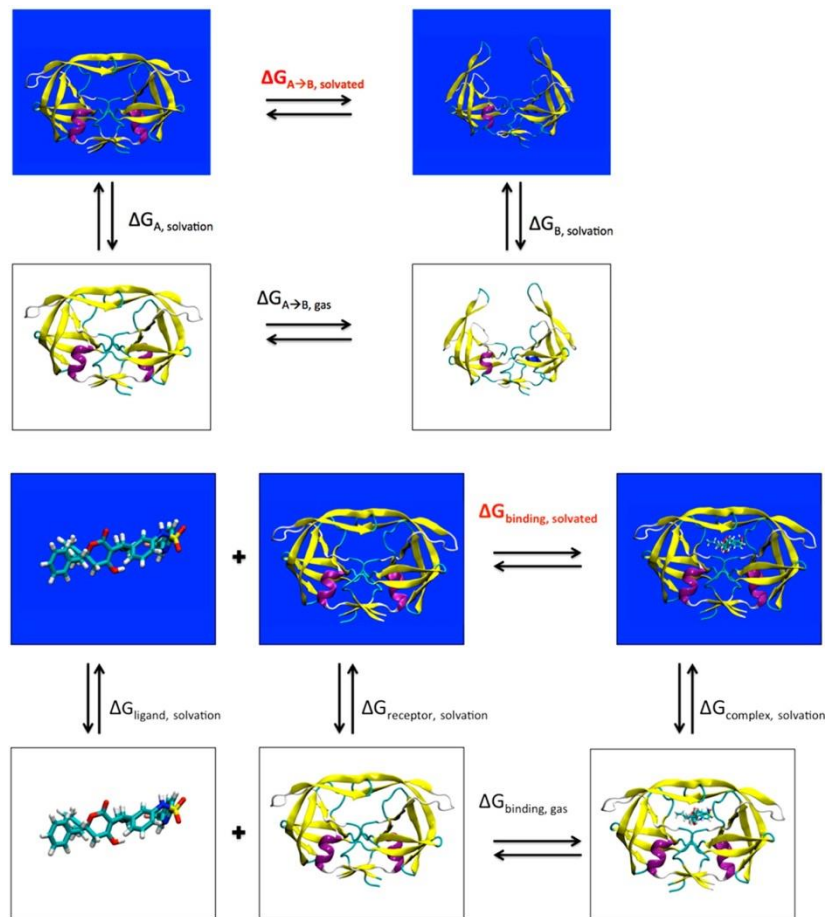


Figure 14. Thermodynamic cycle for free binding energy calculation. Adapted from Miller et.al (2012).

This calculation is done for each frame of the MD simulation and then averaged (Miller et al., 2012), with the aim of taking into consideration all the dynamics of the system, in agreement with the

induced-fit theory mentioned before. To this purpose, there are several methods (each of them with its advantages and drawbacks) with different accuracy and computational cost. Within them, the so-called end-point techniques are a widely used option because of their good balance between accuracy, computational cost and speed. Among them, the **Molecular Mechanical/Generalized Born Surface Area** (MM/GBSA) method, is a popular technique that has been widely employed along this thesis, as it can be seen in **Chapters 1-4** (Kollman et al., 2000; Massova & Kollman, 2000). The binding energy resulting from MM/GBSA is more realistic than the energy obtained from rigid docking calculations, because the dynamic behaviour of the protein-ligand complexes can be taken into account (Mulakala & Viswanadhan, 2013; Genheden & Ryde, 2015). Thus, a better ranking (based on the binding energy) of the analysed compounds can be obtained, allowing for a better prioritization of them, although the obtained binding energies can be far from being experimentally comparable.

Alzheimer's Disease

Alzheimer's disease (AD) is the most common cause of irreversible **dementia** worldwide, representing 60-80% of the total cases. It is estimated to be over 45 million people globally. Its prevalence grows constantly, mostly because of the progressive aging of the population and the long asymptomatic initial stages of the pathology (Crous-Bou et al., 2017). In addition, limitations on current treatments which may slightly improve the symptoms but do not cure the disease, do not help to reduce the high incidence; thus, nowadays AD is one of the major world's socioeconomic and health problems (Citron, 2010). AD is a neurodegenerative disorder resulting in a gradual loss of cognitive function and memory deterioration. Alzheimer's pathologies are characterized by the presence of neurofibrillary tangles (NFT), which are intraneuronal insoluble aggregations mainly composed of abnormal phosphorylated **tau protein**, and senile plaques (SP), principally composed by **beta-amyloid** peptides (A β). Tau protein was discovered in the 70s and it is responsible for the structural morphology of the neurons by stabilizing the microtubules (Kosik, 1993). Tau binding is regulated by its phosphorylation state, a regulated balance between tau kinase and phosphatase activities, which at a time is coordinated by the action of some kinase proteins (Mandelkow et al., 1995). In pathological conditions, such as those provoked by AD, the binding decreases and the neuronal microtubules lose their organization leading to their aggregation and the formation of NFT (Billingsley & Kincaid, 1997; Kolarova et al., 2012) (Figure 15).

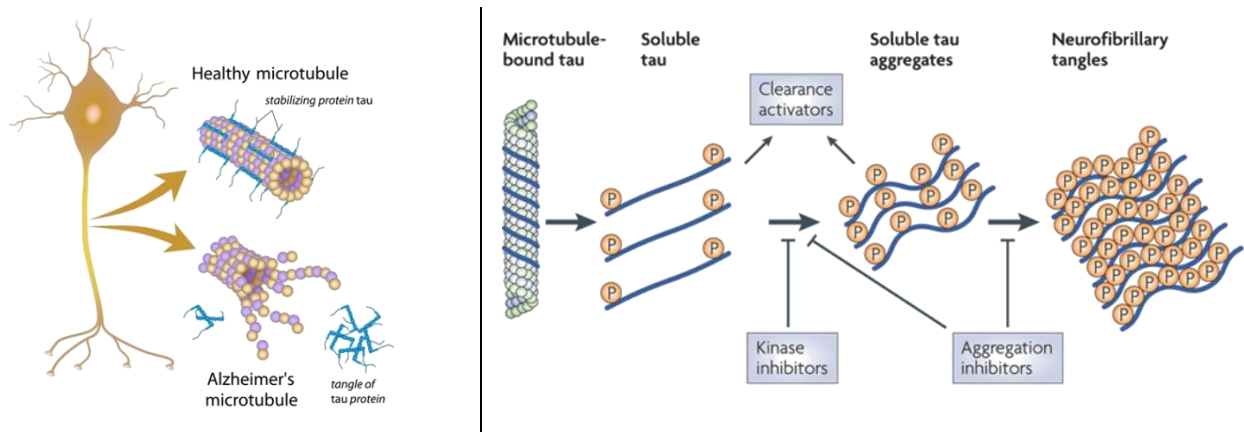


Figure 15. Microtubule-bounded by tau in health conditions and the hyperphosphorylation and consequently aggregation till the formation of neurofibrillary tangles (NFT) in Alzheimer's Disease (AD). Source: Nature Reviews Drug Discovery 9, 387-398, 2010. doi:10.1038/nrd2896 (Citron, 2010) and Alzheimer's news, 2014, Tau Protein Leads To Neuronal Death in Alzheimer's by Patricia Inacio.

$A\beta$ is a peptide of 40 or 42 aminoacids essentially involved in AD as a main component of the SP found on Alzheimer's patients brains (Hamley, 2012). The amyloid precursor protein (APP) is cut by two proteases, beta (β) secretase (also known as beta-site APP cleaving enzyme 1 (BACE1)) and gamma (γ) secretase to yield $A\beta$. In health conditions $A\beta$ is found in a monomeric form, while in pathology conditions it is generally believed that the formation of $A\beta$ oligomers, which are toxic and cause a synaptic dysfunction, starts to aggregate to finally form an amyloid plaque (Shankar et al., 2008; Zhao et al., 2012) (Figure 16). The inhibition on the production of $A\beta$ preventing APP cleaving, remains in the central focus of the research to find a cure for AD, but, the function of APP is still controversial and not well understood yet (Hiltunen et al., 2009). This should make us raise the need to first understand how the pathology works, in order to further proceed in the design of drugs to treat AD.

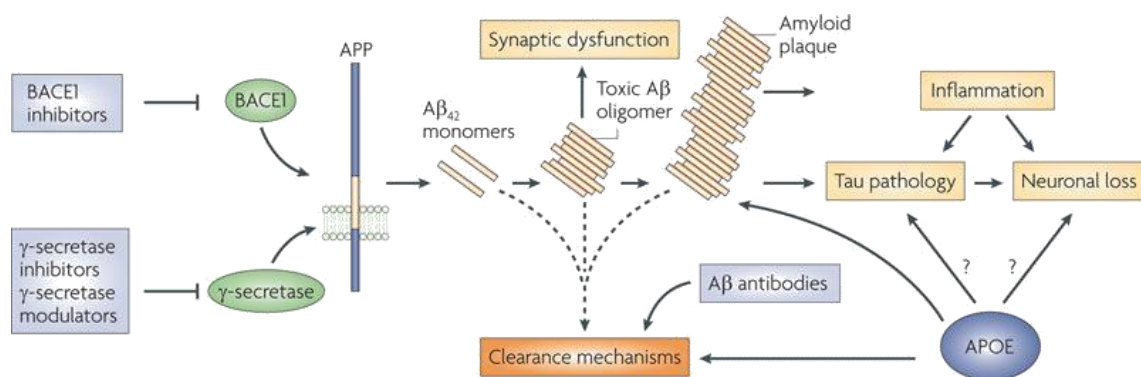


Figure 16. Graphical representation of the amyloid cascade theory where beta secretase (β -secretase, also known as beta-site APP cleaving enzyme 1 (BACE1)) and gamma secretase (γ -secretase) yield beta amyloid ($A\beta$) and its normal form in monomers, and after its aggregation, the formation of toxic oligomers and amyloid plaques directly linked to Alzheimer's Disease (AD) takes place. Source: Nature Reviews Drug Discovery 9, 387-398, 2010. doi:10.1038/nrd2896 (Citron, 2010).

Regarding therapeutic approaches, the most direct approximation is concentrated on the reduction of $A\beta$ production by the inhibition of β -secretase/BACE1 activity, which is the responsible of the proteolysis of APP, precursor of $A\beta$ (Yan & Vassar, 2014) (Figure 16). However, the lack of promising results has created reasonable doubts about the amyloid hypothesis, which never was generally accepted (Doig et al., 2017; Kametani & Hasegawa, 2018). These doubts have ended up in the need to look for new therapeutic options. A strategy oriented to reduce tau hyperphosphorylation and thus, reducing the NFT formation, is conceptually more tempting. In addition to that, there is a general consensus about its damaging effects (Citron, 2010). It is believed that the inhibition of specific tau kinases could reduce the aggregation and now, it is considered a promising approach for the treatment of AD (Martin et al., 2013; Tell & Hilgeroth, 2013; Llorach-Pares et al., 2019).

In this thesis, based upon the results obtained on the elucidation of possible targets from a set of marine molecules, some of the compounds collected on expeditions to Antarctica and the Mediterranean Sea from the BlueBio team (University of Barcelona), as well as other related molecules described in the literature, in **Chapter 1**, we found it interesting to study the relation obtained between meridianin A and the evolutionarily conserved group of dual specificity kinases cdc2-like kinases (CLKs). In fact, one of its isoforms, **CLK1**, is known to be involved in the pathology of AD by the phosphorylation of serine and arginine-rich (SR) proteins responsible for the regulation of the alternative splicing of microtubule-associated tau (Jain et al., 2014). As a consequence of that, we decided to perform a deep study evaluating the possible inhibitory activity of meridianins A-G, the whole family, against the principal kinases involved in tau hyperphosphorylation and thus, AD pathology. Between these proteins,

glycogen synthase kinase-3 beta (**GSK3 β**), a proline-directed serine/threonine kinase that phosphorylates tau at different sites (specifically from 42 sites, 29 of have been found phosphorylated in AD brains), is considered one of the main responsables of tau phosphorylation (Wagner et al., 1996; Hooper et al., 2008; Martin et al., 2013). Interestingly, the fact that more studies suggest a relation between GSK3 β and the production of A β , makes more relevant, if possible, the importance of targeting GSK3 β for the cure of AD (Hernández et al., 2010; Hernandez et al., 2012). Casein kinase 1 isoform delta (**CK1 δ**) also regulates microtubule dynamics through tau phosphorylation (at 46 sites, 25 are found in AD brains) (Singh et al., 2002; G. Li et al., 2004; Hanger et al., 2007). In AD brains, CK1 is co-localized with NFT and isoforms α , δ and ϵ , and their levels are increased respect to normal brains (Schwab et al., 2000; Knippschild et al., 2005). The last kinase studied in this thesis was the dual specificity tyrosine phosphorylation-regulated kinase isoform 1 A (**DYRK1A**). It was first related to the phosphorylation of threonine 212, but today the list of phosphorylation sites rises up to eleven (Wegiel et al., 2011). There are also insights that DYRK1A hyperphosphorylation is related to the inhibition of the ability of tau to enable the microtubule assemble (Ryoo et al., 2007), and its involvement on the formation of NFT (Wegiel et al., 2008).

The inhibition of these four targets is proposed to be key in the treatment of AD, as can be seen in the literature, and during the last decades there have been several studies aiming to prove this (Bhat et al., 2004; Perez et al., 2011; Jain et al., 2014; Branca et al., 2017). The link between NPs and MNPs on the inhibition of kinases involved on AD, and even targeting the four aforementioned kinases, is not new, as several examples can be found in the literature (Haefner, 2003; Liu et al., 2012). For instance, indirubins or phenylmethylene hydantoinis have been proposed to be inhibitors of the GSK3 β (Meijer et al., 2003; Khanfar et al., 2009). For the two dual specificity kinases, DYRK1A and CLK1, harmine, a well-known alkaloid was reported to be a potent inhibitor (Göckler et al., 2009; Adayev et al., 2011; Grabher et al., 2012). Also, lamellarins, pyrrole alkaloids isolated from different marine invertebrates, have been predicted to be CK1 δ inhibitors (Bharate et al., 2013).

Objectives of this thesis

The main purpose of this thesis is **to find possible therapeutic activities and to establish the capability to modulate protein functions in pathogenesis-related pathways from marine molecules by using different CADD tools and techniques**. In order to achieve this goal, the present thesis is divided into three sections which attempt to illustrate these achievements. In **Section I**, I show how a computational approach could **improve the drug discovery pipeline** (Chapter 1). **Section II** focuses on the **elucidation of different pharmacophoric features of marine compounds and a precise *in silico* binding study** that ends with the elucidation of the capability of different marine compounds to act as inhibitors of tau kinases (GSK3 β and CK1 δ) and dual-specificity (DYRK1A and CLK1) protein kinases, all of them related to AD, which constitute a promising starting point for the development of novel anti-AD drugs (Chapter 2 and Chapter 3). **Section III** presents a **computational study and an experimental validation** of the inhibitory activity of meridianins and lignarenones as possible GSK3 β adenosine triphosphate ATP and/or substrate inhibitors, which allows to propose them as drug-like candidates for the treatment of AD pathologies (Chapter 4).

The specific objectives for each chapter are summarized below:

- Chapter 1. ***In silico* Studies to Find New Therapeutic Indications for Marine Molecules**. The main aims of this study are (I) to establish the possible therapeutic potential of a set of marine molecules by using different computational techniques; (II) to predict and validate the marine molecule-target complex binding, (III) to elucidate a list of possible targets, (IV) to evaluate the adverse health effects by performing a preliminary toxicology prediction study; and (V) to estimate the drug-like properties of each studied molecule.
- Chapter 2. **Computer-Aided Drug Design Applied to Marine Drug Discovery: Meridianins as Alzheimer's Disease Therapeutic Agents**. In this chapter, the aims are (I) to highlight the power of CADD techniques in marine molecules, and natural products in general, to find possible therapeutic uses; (II) to evaluate and report the inhibitory activity found in the marine tunicate *Aplidium*: Meridianins A-G, acting as ATP competitive inhibitors of GSK3 β , CK1 δ , DYRK1A, and CLK1; (III) to evaluate the possible adverse health effects of meridianins by performing a

Objectives

preliminary pharmacokinetic study; and (IV) to analyse their pharmacological properties as well as the effect that the presence of halogen atoms in their structure may have.

- Chapter 3. **Kororamides, Convolutamines, and Indole Derivatives as Possible Tau and Dual-Specificity Kinase Inhibitors for Alzheimer's Disease: A Computational Study.** The objectives are (I) to discover the possible therapeutic activity of kororamides and convolutamines against AD by the inhibition of GSK3 β , CK1 δ , DYRK1A, and CLK1; (II) to determine the importance of the indole scaffold for the inhibition of the four studied kinases and the importance and effect of the halogen substituents; (III) to design new possible inhibitors of the four kinases starting from meridianin and kororamide indole scaffolds; and (IV) to evaluate the adverse health effects of kororamides, convolutamines and its derivatives by performing a preliminary ADMET study.
- Chapter 4. **Meridianins and lignarenones as potential GSK3 β inhibitors and inductors of structural synaptic plasticity.** The aims here are (I) to elucidate the possible ATP and/or substrate inhibitory activity of meridianins and lignarenone against GSK3 β , a key target on the AD pathway; (II) to explore druggable binding sites on GSK3 β on the search of new allosteric cavities; (III) to ascertain the pharmacokinetic properties; and (IV) to experimentally validate the inhibitory activity of meridianins and lignarenones comparing Ser9 phosphorylation levels to total levels of GSK3 β as an indication of inhibition.

Supervisor's Report

Dr. Conxita Avila, Director of the PhD Thesis entitled "Computer-Aided Drug Design applied to marine drug discovery", together with Dr. Melchor Sánchez Martínez as co-Director, certify that the dissertation presented here is the result of the work carried out by Laura Llorach Parés under our guidance and supervision. The contribution of the PhD candidate to each one of the manuscripts included in her Thesis is detailed below:

Chapter 1. *In silico* Studies to Find New Therapeutic Indications for Marine Molecules.

Llorach-Pares, L; Nonell-Canals, A; Sanchez-Martinez, M; Avila, C.

(In prep)

LL-P: data analysis, carry out the experiments, results interpretation, and manuscript elaboration.

Chapter 2. Computer-Aided Drug Design Applied to Marine Drug Discovery: Meridianins as Alzheimer's Disease Therapeutic Agents.

Llorach-Pares, L; Nonell-Canals, A; Sanchez-Martinez, M; Avila, C.

Marine Drugs 15(12), 366 (2017)

5-Year Impact Factor: 4.379 (2017)

LL-P: data analysis, carry out the experiments, results interpretation, and manuscript elaboration.

Chapter 3. Kororamides, Convolutamines, and Indole Derivatives as Possible Tau and Dual-Specificity Kinase Inhibitors for Alzheimer's Disease: A Computational Study.

Llorach-Pares, L; Nonell-Canals, A; Avila, C; Sanchez-Martinez, M.

Marine Drugs 16(10), 386 (2018)

5-Year Impact Factor: 4.379 (2017)

LL-P: design of the computational study, data analysis, carry out the experiments, results interpretation, and manuscript elaboration.

Chapter 4. Meridianins and lignarenones as potential GSK3 β inhibitors and inductors of structural synaptic plasticity.

Llorach-Pares, L; Rodriguez, E; Giralt, A; Nonell-Canals, A; Sanchez-Martinez, M; Alberch, J; Avila, C.

(In prep)

LL-P: design of the computational study, data analysis, carry out the computational experiments, results interpretation, and manuscript elaboration.

From all the co-authors of the different chapters, ER has not been awarded a PhD degree yet. We hereafter guarantee that none of the information contained in the chapter co-authored with her will be used to elaborate any other part of someone else's PhD thesis.

For all the above, we consider that the work developed by the PhD candidate grants her the right to defend her Thesis in front of a scientific committee.

Barcelona, September 30th, 2019.

Dra. Conxita Avila

Dr. Melchor Sánchez-Martínez

Section I

Target elucidation from a computational approach

Section I

Chapter 1

In silico studies to find new therapeutic indications for marine molecules.

Chapter 1

In silico studies to find new therapeutic indications for marine molecules.

Laura Llorach-Pares ^{1,2}, Alfons Nonell-Canals ¹, Conxita Avila ² and Melchor Sánchez-Martinez ¹

1. Mind The Byte S.L., 08028 Barcelona, Catalonia.
2. Department of Evolutionary Biology, Ecology and Environmental Science, Faculty of Biology and Biodiversity Research Institute (IRBio), Universitat de Barcelona, 08028 Barcelona, Catalonia.

Abstract

Identifying small molecules that fit well into a binding cavity is one of the first steps in the drug discovery pipeline. In this study, we try to elucidate a list of possible targets and the therapeutic potential of a set of selected marine molecules employing different computational tools. Molecular docking is one of the most common computer-aided drug design (CADD) tools which allow the study of protein-ligand interactions, predicting at the same time the binding molecule orientation or pose. Capturing protein motions is key to understand these molecule-target interactions, and Molecular dynamics (MD) simulation is the best computational tool to do so. By the combination of these computational tools and others, in this work we have established the link between a group of Antarctic marine molecules and some neurodegenerative and cardiovascular pathologies. Moreover, we evaluated the adverse health effects through toxicology predictions, and the drug-likeness properties of this set of marine molecules, providing some insights of the prediction on marine molecule-target complex binding. In this study, we exemplify how the use of computational tools can be applied on the marine drug discovery field, establishing a pipeline to be followed on future studies.

Keywords: Marine natural products, Computer-aided drug design, Neurodegenerative diseases, Molecular docking.

Resum

Identificar petites molècules que s'adaptin bé a una cavitat activa és un dels primers passos a seguir en el descobriment de fàrmacs. En aquest estudi pretenem dilucidar una llista de possibles dianes, i el potencial terapèutic, d'un conjunt de molècules marines utilitzant diferents eines computacionals. L'acoblament molecular és un dels instruments del disseny de fàrmacs assistit per ordinador (DFAO) més comuns i que permet l'estudi de les interaccions proteïna-ligand, predint a la vegada, tant l'orientació com la postura de la molècula acoblada. La captura dels moviments de les proteïnes és clau per entendre aquestes interaccions proteïna-ligand, i la simulació de dinàmica molecular (DM) és la millor eina computacional per fer-ho. Mitjançant la combinació d'aquestes eines computacionals i d'altres, en aquest estudi hem pogut dilucidar el vincle entre un grup de molècules marines i algunes patologies neurodegeneratives i cardiovasculars. A més, hem avaluat els possibles efectes adversos en la salut mitjançant prediccions de toxicologia, i les propietats farmacològiques d'aquest conjunt de molècules marines, proporcionant algunes idees sobre la predicció dels vincles d'unió dels complexos molècules marines-proteïna. En aquest estudi exemplifiquem com es pot aplicar l'ús d'eines computacionals en el camp del descobriment de fàrmacs marins, establint un procediment que es pot seguir en futurs estudis.

Paraules clau: Productes Naturals Marins, Disseny de Fàrmacs Assistit per Ordinador, Acoblament Molecular, Malalties Neurodegeneratives.

Introduction

Molecular docking is a very popular computer-aided drug design (CADD) tool used in molecular biology to evaluate ligand-target complementarity [1]. This method allows the study of protein-ligand interactions at an atomic level and the prediction of preferred binding orientation or poses (binding mode) of one molecule (typically a small organic compound) to another molecule (generally a biological target such as a protein). Also, by the use of scoring functions, this powerful technique assesses and predicts the binding affinity of the complex formed by a receptor and a ligand. So, essentially, docking can be accomplished through two complementary steps: first, by sampling ligand conformations in the active site of the protein and then, ranking these conformations using a scoring function [2]. Two types of docking can be described, a) the classical (rigid) docking, where the crystallographic ligand's pose and the binding site of a given protein is established and can be used, and b) the blind docking, where the binding site is not known, and it is first necessary to search the protein surface to find possible cavities that lodge the active binding sites before performing the classical docking. Both methods are widely used in the drug discovery process.

The first and critical step in small molecule drug design is the identification of ligands that fit well into the binding pocket of the protein target [3]. Looking back at the history, in 1894, Emil Fischer first postulated the specific action of an enzyme with single substrate using the lock and key analogy [4]. First reported dockings using this principle, where both ligand and receptor were treated as rigid entities, were published at the end of 1982 [5]. As this lock-key hypothesis does not fully account for the nature of the enzymatic actions, another theory has been proposed, called the induce-fit theory which refuses that the substrate would do more than simply fit into the active site and assumes a continuous change in the conformation and shape of the enzymes in response to substrate binding, in other words, when a substrate binds to an enzyme, it will change its conformation [6]. After the introduction of this principle, more accurate predictions of binding poses and binding affinities (minimum binding energies) could be resolved. This fact can be traduced into a best candidates selection as active compounds, with higher true positive rates of success, and also, at the same time, can considerably reduce expensive experimental efforts.

From now on, the flexibility of enzymes has been considered, taking into account that is extremely important because of these dynamic movements provide a mechanism for regulating enzymatic activity. Consequently, and in agreement with the theory of the induce-fit, both, the ligands and receptors (proteins) should be treated as flexible during docking. Due to computer limitations resources, docking has been performed with flexible ligand and rigid receptor until today, and remains

one of the most used and popular methods in drug discovery [2]. To deal with this drawback, and with the aim of incorporating protein flexibility, many efforts have been made and new methods are now being used to validate docking techniques, such as Molecular dynamics (MD) simulations which can optimize the predicted binding mode, allowing the study of movements (so-called induce-fit effects) of atoms along the time by integration of Newton's equations of motions disclosing the adaptation of the ligand to the target [7–10]. Recognizing the mechanisms of actions for the protein-ligand complexes formation and understanding its binding, will help at the discovery, design and development of new drugs. Protein-ligand interactions, thus, play an important role in many scientific areas and more concretely, knowledge on these interactions is central for understanding biology at the macromolecular level.

In this project, by the use of docking calculations and MD simulations, a chemical library of ten marine molecules with marine origin was explored (Figure 1).

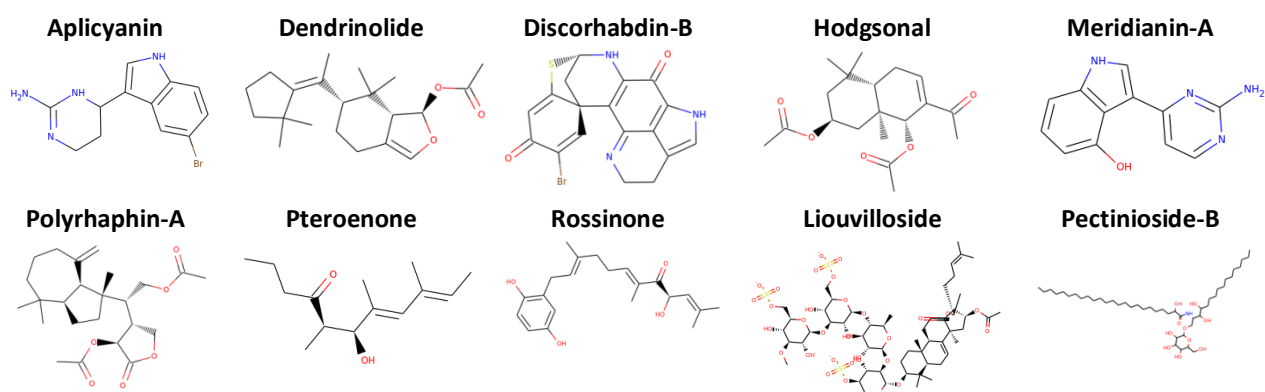


Figure 1. Structures of the ten marine molecules selected for this study.

Specifically, several of these molecules were collected from benthic marine invertebrates from the South Shetland Islands and the Weddell Sea (Antarctica), but also from some other areas of the planet, like the Mediterranean Sea, and the Sea of Japan. Antarctica, in particular, because of its biological and geographical characteristics, is considered a “biodiversity mine” where new species and new marine molecules are constantly being discovered [11–13]. Oceans cover about 70% of the Earth's surface, and up to 80% of life inhabits in the marine ecosystem. Marine organisms constitute amazing resources of enzymes and bioactive compounds. Marine natural products have always been a rich source of drug-like compounds [14,15]. Biodiversity is related to chemodiversity as well, giving wider opportunity for discovering novel therapeutics with novel mechanisms of action. Currently, there are

seven marine-derived products approved as drugs by the United States Food and Drug Administration (US FDA), supporting the importance of marine drug-like compounds [16].

The identification of natural products that are capable of modulating protein functions in pathogenesis-related pathways is one of the most promising lines followed in drug discovery [17]. In the present study, with the aim of knowing the possible therapeutic potential of a set of selected marine molecules, we divided the study into four parts. **First** we aim to determine a list of possible targets. To do so, two dimensions (2D) and three dimensions (3D) ligand-based virtual profiling (VP) software tools were used. Once the targets were known, we decided to focus only on neurodegenerative and cardiovascular pathologies. According to the “Health at a Glance 2013” report, in 2009, there were 14 million people estimated of suffering from dementia in the Organization for Economic Cooperation and Development countries (OECD). The same study, considered that in 2011, cardiovascular diseases were the main cause of mortality in most OECD countries representing 33% of all death [18]. To achieve the relation between the targets found on the first step and relate them to the pathologies of interest, we used DisGeNET, a database that integrates information on gene-disease associations [19]. The **second** part is dedicated to toxicology prediction. Due to the importance on these predictions during drug discovery, several computational toxicology tools are used to evaluate the adverse health effects of the studied molecules. These software tools integrate information and data for a wide kind of sources, and they allow developing predictive mathematical and computer-based models. Toxicology methods can be used to reduce the dependence on experimental testing in general, and animal testing in particular, and this means saving money and time [20]. Moreover, these techniques, given their inexpensiveness and expeditious results, can be used in an early phase of drug discovery, before the synthesis of the molecules, to prevent future problems and helping to rationally focus the drug development process. Computational toxicity could be assessed based on ligand or target. In this study we performed both kinds of predictions, in order to ensure as much as possible whether the studied molecules are toxic or not. In the **third** part, a drug-likeness evaluation was carried out. Drug-likeness is a qualitative concept used in drug design based on structural properties of compounds and it is used to estimate drug-like properties of molecules. From molecular properties and structural features, it is possible to determine whether a particular molecule is similar to a known drug [21,22]. The **fourth** and final part, is mainly focused on the computational techniques previously mentioned as protein-ligand docking and blind docking, which offer the capability to predict quickly and cheaply the binding mode and the affinity of a ligand-receptor complex, and the induce-fit MD simulations and binding free energy calculations, specifically, using reweighting techniques. These methods are popular approaches to estimate the free

energy of the binding of small ligands, in these particular case marine molecules, to biological macromolecules and proteins. They are typically based on MD simulations of the receptor-ligand complexes and may be useful to improve the results of docking calculations or to understand observed affinities and trends [23]. Despite the fact that MD simulations are more time-consuming approaches, computing the free energies of the complex systems based on the thermodynamic principles are more powerful methods. Also, a detailed analysis of data generated by MD simulations (key structural ligand-target interactions, energy and temperature terms, root-mean-square deviation (RMSD) and fluctuation (RMSF), and radius of gyration) are performed as this is a crucial step to understanding the nature of the complex binding.

Results and discussion

Virtual Profiling

Molecular similarity is an important concept in drug discovery. Its principle is based on the assumption that structurally similar molecules generally have similar properties, as they share similar physical properties and biological functions. On drug discovery process, 2D similarity approaches are widely used due to their simplicity, accuracy and efficiency [24]. Using Cabrakan software tool, we performed 2D VP on the initial set of 10 molecules. At this point, two molecules, Liouvilloside and Pectinoside-B, had to be discarded because in both cases no target was found for them. This means that the database did not contain enough molecules similar to the query molecule. Sometimes, those findings were consequence of the complex chemical structures that marine molecules have. As the profiling experiments allow for the exploration of the molecular targets as well as their activities, the remaining eight molecules were classified according to the activity shared and the protein families they interact with.

Recently, the focus of these kind of studies has moved to 3D similarity methods. Those functions are now gaining attention for their application in molecular target prediction. So, a 3D VP was performed, by the use of Hurakan software tool, over the initial set of selected molecules. With this technique, Liouvilloside and Pectinoside-B, had to be discarded because they were too large to be analyzed by Hurakan, since this software can not load molecules with atomic masses higher than 900 Da or more than 32 rotamers. At this point, Discorhabdin-B was discarded too because no targets were found, probably due to the highly complex chemical nature of these marine molecules. The selection

criteria of these software is different, and that is way there are molecules that have a suitable size for Cabrakan but not for Hurakan.

Using DisGeNET data could be crossed and this allowed us to select those targets related to cardiovascular and neurodegenerative diseases. Target selection was done following the criteria explained in methods section. Importantly, we used all targets presenting more than three hits in both of our VP analysis. By “hits” we mean that the database searching found at least three similarity matches for the selected target. Therefore, at the end of the target selection process, targets associated to neurodegenerative and cardiovascular pathologies, but also associated to other pathologies, were included. From all the targets found, it was possible to relate 12 to neurodegenerative diseases, six to cardiovascular diseases and nine to both neurological and cardiovascular pathologies. Six of them were related with other disorders, specifically orphan diseases (digital clubbing, pituitary-dependent Cushing’s disease, and mental retardation X-linked), peripheral nervous system disease, prostatic and lung neoplasm (Figure 2).

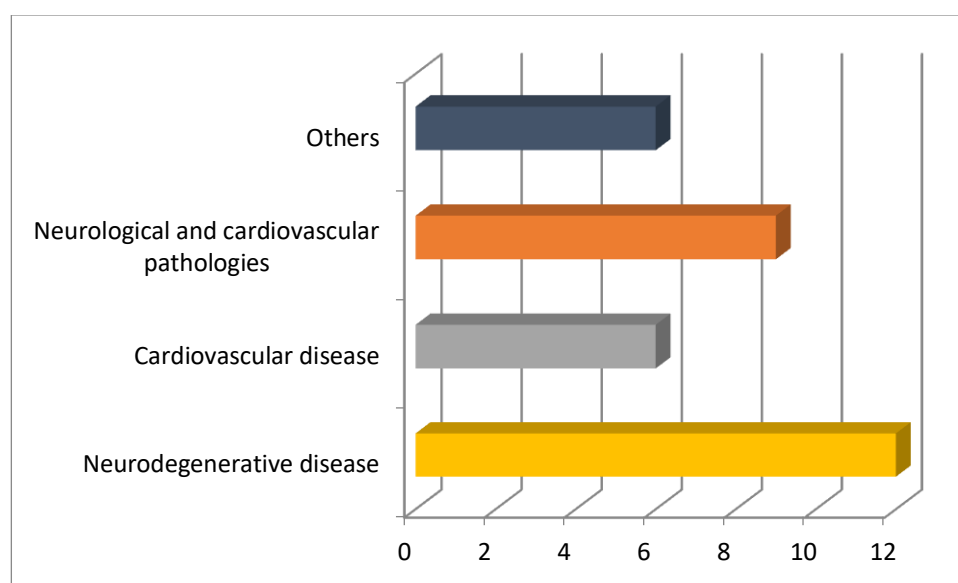


Figure 2. Relation between targets and pathologies. Yellow: neurodegenerative diseases, Grey: cardiovascular diseases, Orange; Neurodegenerative and cardiovascular pathologies, Purple: Other pathologies.

Therefore, after the two VPs for the seven remaining molecules, a total of 33 targets were found. Some of them were shared between different molecules of the set. Specifically, there were 9 targets that interacted with more than one marine molecule. At the end, we had 75 marine molecule-target complexes.

Toxicology prediction

Due to the importance of toxicology studies during the process of drug discovery, the whole set of 10 molecules was studied. Although three of them were already discarded, it was considered important to include them in the prediction because it could give very interesting information on marine molecules toxicology. The toxicology study based on ligand was focused on the study of carcinogenicity, toxicity, mutagenicity, and skin sensitization, and was performed using VEGA software tool [25]. From all the set of molecules, Liouvilloside showed no results in any model probably because it is too different from the molecules in the models we tried. For the rest of the molecules, the results are as reported in Figure 3.

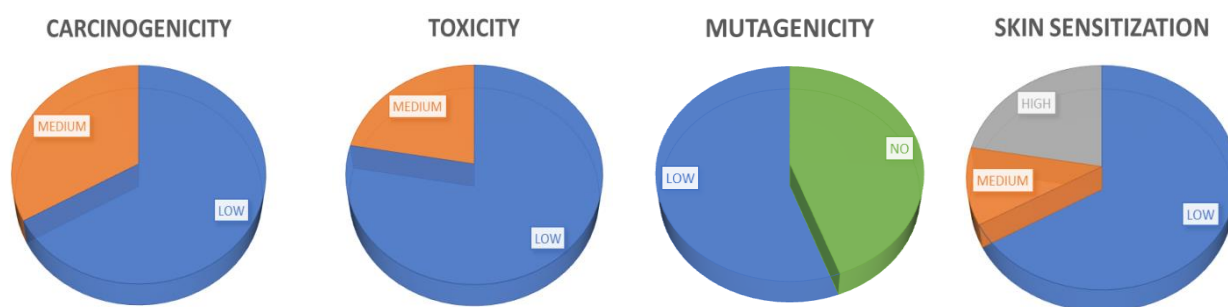


Figure 3. The above graphics show the four different toxicology prediction models obtain using VEGA software. The toxicity is divided into for models: carcinogenicity, toxicity, mutagenicity and skin sensitization.

- Carcinogenicity: Six marine molecules with low probability and three with medium.
- Toxicity: Seven molecules with low probability and two with medium.
- Mutagenesis: No mutagenesis was found on four molecules while in five we found low probabilities.
- Skin sensitization: Six marine molecules were found with low probabilities, one with medium and two with high.

Overall, the results obtained show that most marine molecules studied here had some toxicology effects, even if low. These results may be easily explained by the fact that these marine molecules come from marine invertebrates, such as sponges, molluscs, echinoderms, and tunicates, which in fact, use toxic metabolites as chemical defenses against different species of fish and other small animals, like amphipods [26]. There is a similar case in the literature, where didemnin B, a cyclic depsipeptide isolated from a tunicate collected in the Caribbean Sea, was in clinical trials that had to be suspended due to significant neuromuscular toxicity, thus highlighting the importance of these kind of studies in preclinical drug discovery phases [27].

In addition to ligand-based toxicity prediction, we also performed calculations based on target. To that extent, The toxin and Toxin Target Database (T3DB), which contains toxins linked to their corresponding targets, showing toxin-target association, was used as a reference database [28]. We found that from the 10 studied molecules, none of them had a Tanimoto score higher than 0.65, what is interpreted as no toxicity for any of the molecules of the set (Figure 4).

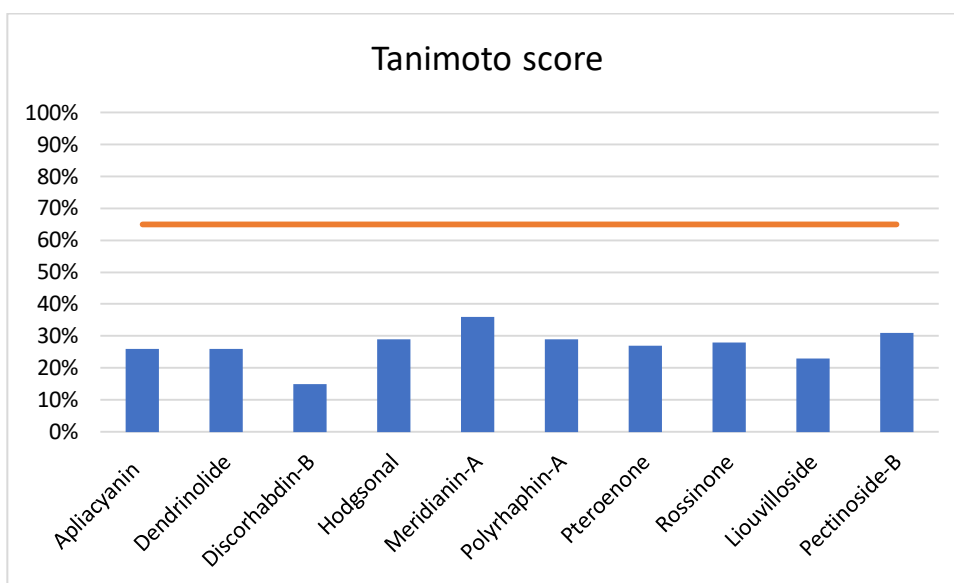


Figure 4. 2D Tanimoto based similarity results. The orange line represents the Tanimoto score (0.65).

The Tanimoto index used here is one of the best metrics for similarity calculations [29]. Performing these kind of analysis in early stages on the drug discovery process could help to predict drug side effects and adverse effects, which is crucial for the efficiency, because the early identification of any potential toxicity can save an enormous investment of money in a drug that will later be found unfeasible [30].

Drug-likeness evaluation

Using this test we evaluate if the selected marine molecules share a similar behavior with known drugs, by comparing their docking binding energy when they are docked to the targets of the study. To do so, for each of the targets found on the VP step, we obtained a list of 190 drugs related. Results were only obtained for 17 of the 33 set of targets, including six of the seven marine molecules tested, excluding Pteroenone. Docking calculations with every drug-target and marine molecule-target were performed against crystallographic structures or homology models. From the binding energies scores

obtained, those values lower than -6.5 kcal/mol were discarded and then, the remaining energies were averaged separately (table 1). This is done, first by its link with the target, and then depending on the type of docking performed, either crystallographic representation with ligand, or blind docking against homology models.

Table 1. Summary of the results obtained performing docking simulations using the PDB structures and blind docking simulations, using homology models (HM) with all the drugs found per target, and the corresponding marine molecule. Drugs binding energies are an average of all the drug energies per target. To avoid false positives, each docking calculation was performed twice (R0/R1). Energy values are on kcal/mol.

	PDB - LIGAND	DRUGS FOUND	DRUGS	MOLECULE
			Binding Energy R0/R1	Binding Energy R0/R1
P11511	3EQM-ASD	5	-7,5 / -7,6	-8,1 / -8,2
	3EQM-HEM	5	-8,4 / -8,4	-8,3 / -8,4
P09874	1UK0-FRM	3	-10,1 / -10,1	-8,2 / -8,2
O15530	2R7B-253	4	-10,0 / -10,0	-7,7 / -7,7
	3QC4-MP7	4	-7,5 / -7,5	-8,1 / -7,6
P31749	3O96-IQO	2	-9,0 / -9,0	-8,6 / -8,6
P00491	1ULB-GUN	7	-7,5 / -7,5	-7,2 / -7,2
P15428	2GDZ-NAD	1	-11,9 / -11,8	-7,9 / -7,9
P00374	1MVS-DTM	15	-8,9 / -8,9	-7,7 / -7,7
P49841	3PUP-OS1	6	-9,2 / -9,4	-7,7 / -7,7
P00352	4WB9-NAI	3	-8,4 / -8,5	-8,8 / -8,5
P07550	4GBR-CAU	33	-8,0 / -8,1	-8,7 / -8,7
Q99714	1U7T-NAD	1	-10,1 / -10,1	-7,6 / -7,6
	1U7T-TDT	1	-10,7 / -10,7	-7,7 / -7,7
Q07343	HM	16	-7,2 / -7,2	-8,6 / -8,6
P14867	HM	46	-6,7 / -6,8	-7,6 / -7,4
P24046	HM	1	-7,9 / -7,9	-7,5 / -7,6
P27815	HM	10	-7,3 / -7,3	-8,8 / -8,7
P46098	HM	10	-8,4 / -8,4	-8,1 / -8,1
Q08499	HM	3	-8,0 / -8,0	-9,2 / -9,1

Summarizing, a total of 392 docking calculations were performed and results (Table 1) show a difference of ± 3 kcal/mol between drug-target and marine molecule-target in those represented by crystallographic structures, while for those represented by homology models, differences were ± 2 kcal/mol. These results, supported by the representation of six molecules from the set of seven (Pteroenone was previously excluded), allow us to suggest that the studied marine molecules behave, and thus could act, as drugs.

Virtual profiling validation

Docking calculations

Once the target and the molecule modeling were performed, it was possible to start doing docking calculations to validate the stability of the complexes found at the first VP analyses. At this point, over the 75 complexes (target-marine molecule), 166 dockings were performed. Only those complexes with binding energies higher than -6.5 kcal/mol were selected. So, we ended up with 30 targets and 52 complexes. From these complexes and based on the different docking techniques used, we obtained that, 32 of them were performed over crystallographic structures with ligand, 16 over homology models, and 4 over crystallographic structures without ligands (Table A1, Table A2, Table A3). Considering the binding energies obtained and after a visual analysis, the best ligand conformation-target complex was selected.

Molecular Dynamics simulations

After the selection of the 54 best complexes and its respective selected poses, all of them were then submitted to a short (1ns) MD simulation to post processing the docking poses, with the aim of adding the target flexibility. This would allow to observe the induce-fit events coming from the accommodation of the target to the ligand and vice versa, compared to the rigid docking procedures where only the ligand is allowed to move, missing the protein flexibility, which is essential to carry out their function [31,32]. After each MD simulation, a trajectory with the positions of the atoms comprising the marine molecule-target complexes was generated as a function of the simulation time. As the aim of this study was to find new indications for the set of marine molecules, from the results obtained, after a visual analysis and based on the binding energies obtained, the best complex per molecules were selected to perform a deeper study. At this point, Pteroenone, had to be discarded and was not further analyzed, due to the fact that during the visual analysis of each simulation, we observed an artefact on the protein and the system could not be validated. For Hodgsonal, Polyrhaphin-A and Dendrinolide only one target was selected (P11511, P04798 and P16662, respectively), while for Rossinone two targets (P15428 and P00352) were chosen, and three targets for Meridianin-A (Q9Y463, P15428 and P49759) and Aplicyanin (O15530, P00491 and P31749) were selected.

To sum up, after these analyses, we focused on 11 marine molecule-target complexes (Table A4). From these trajectories obtained after MD simulations, different features such as total energy (Figure A1), potential energy (Figure A2), kinetic energy (Figure A3) and temperature, Radius of Gyration

(Rg), RMSD/RMSF and hydrogen bonding (HB) were obtained. With all these data, we analyzed each feature per complex with the aim of validating the simulation itself.

- Temperature and energy terms

Temperature is one of the fundamental concepts in physics and represents the intensity of the thermal motions of molecules [33]. Due to its strong influence over molecular simulations, especially molecular dynamics, in which the velocities of the atoms are continuously adjusted according to various temperature-controlled algorithms, it is an important value to check [34]. After the analysis of the all 11 temperatures, one for each simulation, could be observed its stability with an average of 297 Kelvin (K) \pm 2. Thus, our results (Figure A1-A3) confirm the validity of all the simulations performed, and therefore, further particular and specific analysis can be undertaken [35,36].

- Radius of gyration

Radius of gyration (Rg) are related to (and give global account of) the general tertiary structure [37]. Calculating the Rg of the protein system along the trajectory, allow us to analyze the compactness of the protein. We thus obtained the Rg for each system (the set of 11 complexes) (Figure 5).

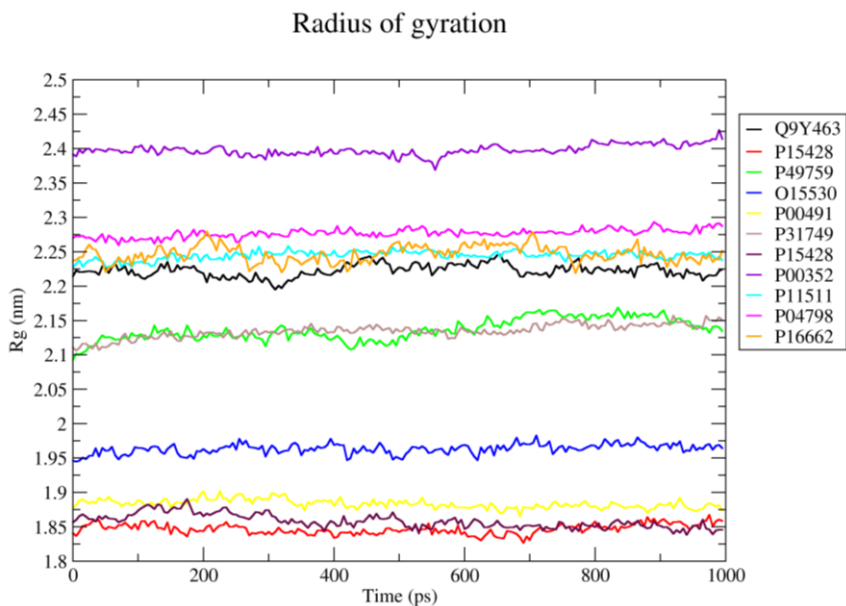


Figure 5. Time evolution of the Radius of Gyration (Rg) obtained for each system. The color code for each system (Uniprot ID) can be seen in the legend box.

As a general trend, the decrease in the average value of R_g as a function of the simulation time suggests contraction of the structure, what is translated into a gain in compactness. P11511 and P04798 fluctuate in a constant way around 2.23-2.26 nm and 2.26-2.29 nm respectively, during the whole simulation, thus indicating stability over time. These results contrast with P16662, where the R_g oscillates from 2.22 to 2.28 nm, and Q9Y463 oscillates from 2.19 to 2.25 nm, which points out an incompactness of the systems. P00352, despite seeming constant on time, shows the highest values compared to the other 10 systems. P49759 and P31749 show a very similar and constant trend; at the time step 0 ps, R_g values are around 2.1 nm and they increase slightly until, at time step 1000 ps, when they reach 2.13 and 2.15 nm. O15530 has lower values than the systems above and the R_g fluctuates from 1.94 to 1.97 nm, indicating stability despite an abrupt steadily fluctuation can be observed. P00491 is one of the most constant systems which fluctuates constantly at 1.88 nm. The two last systems analyzed are P15428, once for Meridianin-A (red) and once for Rossinone (maroon). Both systems have the lower R_g values, indicating stability and compactness. Interestingly, despite the fact that at the time step 0 ps, R_g values are different, 1.84 and 1.85 nm, from the 700 ps they merge and follow the same pattern till the end. These results give reliability over all R_g analysis (Figure 5).

- Root-mean-square deviation

The RMSD of atomic positions is the measure of the average distance between the atoms of superimposed structures [38,39]. To check the stability of the simulations, the RMSDs of the $C\alpha$ atom was calculated and monitored over the course of simulation. In general, this superimposition is performed among the structures extracted from the MD simulation and a reference structure [40]. As a reference, we used the starting structure of each target, and the RMSD value was obtained by comparing it with the structure (all atoms) obtained at each step of the trajectory. This process was performed over the 11 remaining complexes (Figure 6).

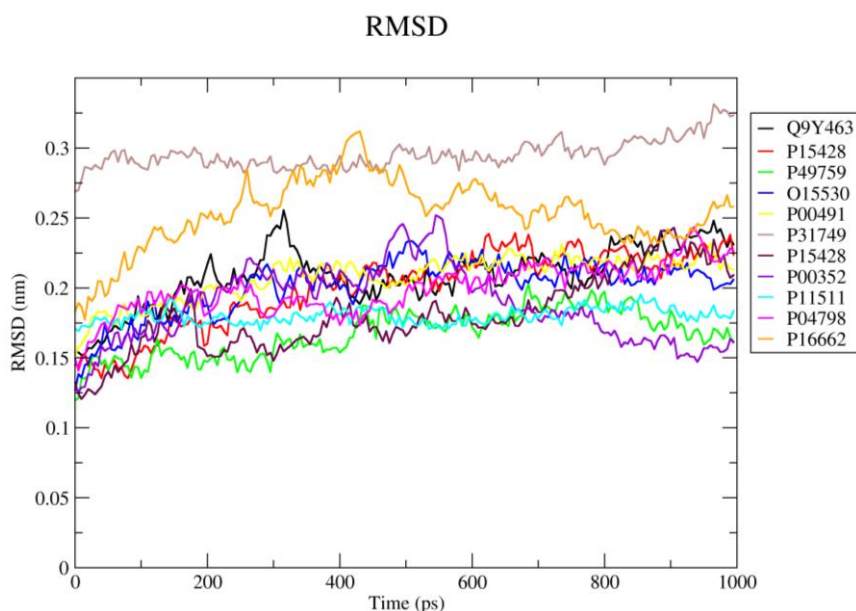


Figure 6. The progress and completion of the MD simulation processes were monitored by plotting a graph of Root-mean-square deviation (RMSD) of the eleven systems obtained by comparing each structure at the starting point and the structure (all atoms) obtained along the trajectory. The color code for each system (Uniprot ID) can be seen in the legend box.

RMSD values increase from the beginning of the simulation to a certain moment in which it remains constant (MD converged). For all the systems, the RMSD remains stable around average values of 0.15-0.25 nm over a considerably time period of the trajectory, indicating that the systems were stable during the simulation (Figure 6). This is the expected behavior during an MD simulation as RMSD is commonly used as an indicator of convergence of the structure towards an equilibrium state.

- Hydrogen bonding

The time-averaged number of hydrogen bonds (HBs) present on each marine molecule-target complex was calculated from the MD trajectories. HBs were defined in such a way that the distance between donor and acceptor was less than the cut-off distance of 3.5 Å (0.35 nm) and the angle donor-H-acceptor was less than the cut-off angle of 20 degrees. Given that this is a dynamical system, the number of HBs is not constant; in fact, these bonds are forming and breaking constantly as the simulation runs. For this reason, we calculated the live time of these HBs throughout the simulation. In this way, we obtained the occupancy of a particular interaction during the 1ns simulation of each complex (Figure 7).

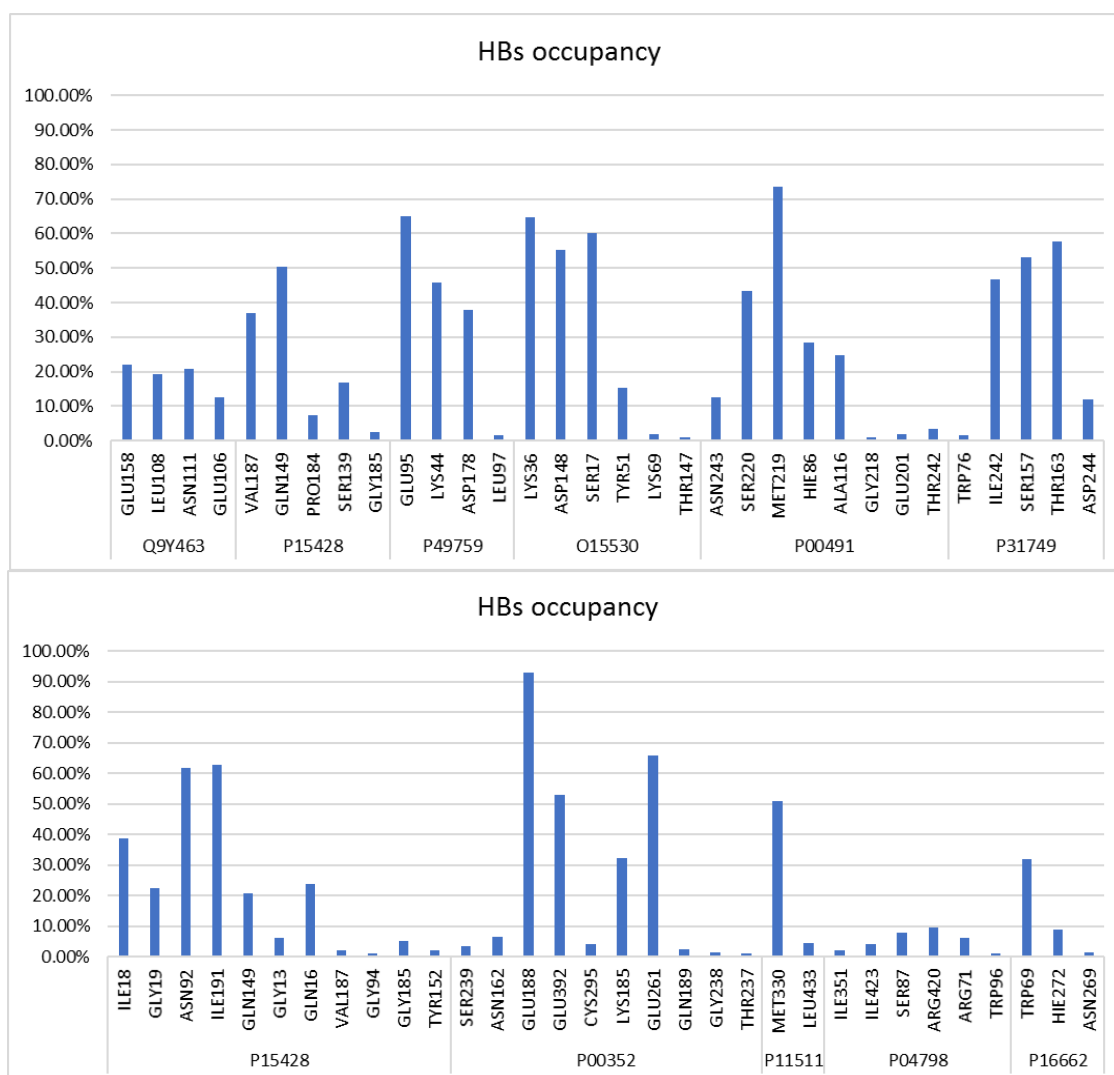


Figure 7. Hydrogen bond (HB) occupancy per target. All occupancies lower than 0.99% were not taking into account and are not shown. Horizontal numbers are the Uniprot ID and vertical letters and numbers refers to the residue involved on the HB of each target.

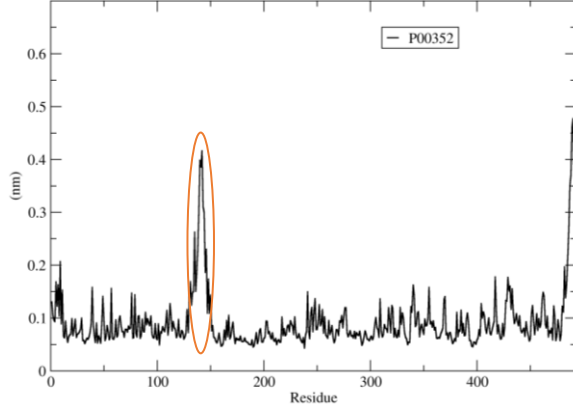
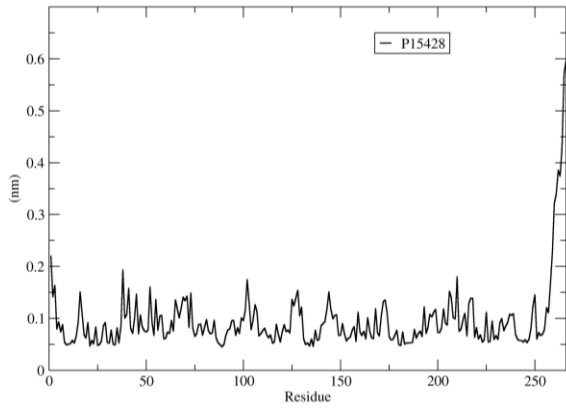
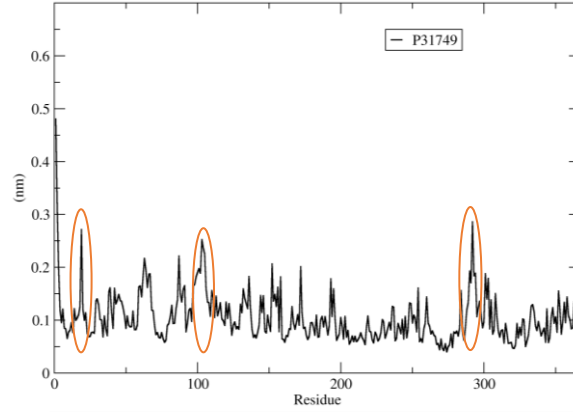
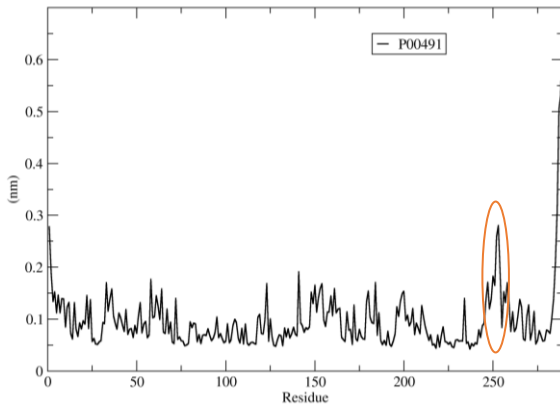
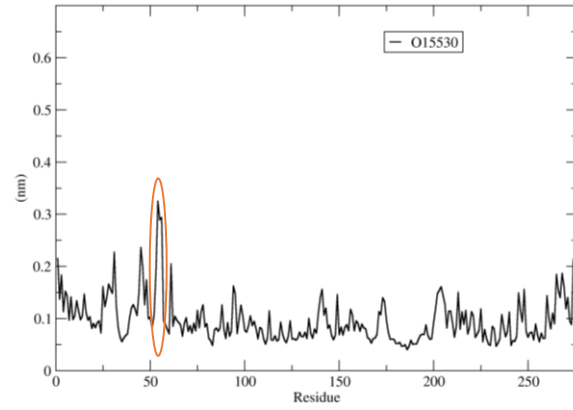
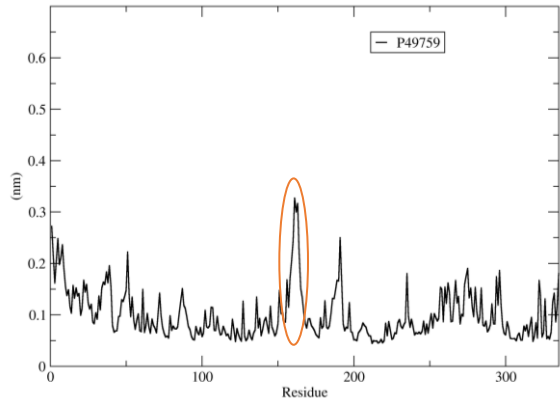
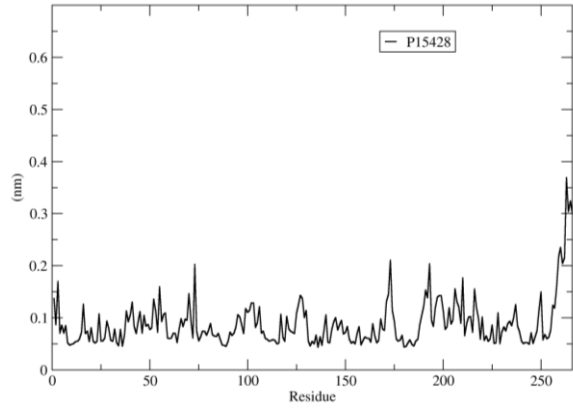
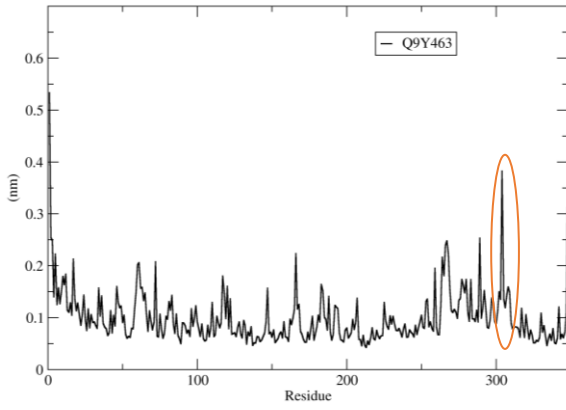
Since this is a dynamical system, the HBs were classified according to their live time as long-lived (present in more than 50% of the simulation), medium-lived (present between 10% and 50%), and short-lived (formed in less than 10% of the simulation). Considering the analysis of the results (Figure 7) and focusing on each target, for Q9Y463 four medium-lived HBs were found. The first P15428 (Meridianin-A) presents one long-lived HB with the residue GLN149, two medium and short-lived HBs. P49759 show four HBs, one long-lived with the residue GLN95 with a live time of 65%, two medium-lived and one short-lived HB. O15530 formed three long-lived HBs with the residues SER17, LYS36 and ASP148 with a live time of 60%, 65% and 55%, respectively, and also,

presents one medium-lived and two short-lived HBs. P00491 formed a total number of eight HBs. Only one of them is a long-lived and is established with MET219, and there are four medium-lived HBs and three short-lived. P31749 presents two long-lived HBs with the residues SER157 and TRH163 with occupancy of 53% and 57%. Also, there are two more HBs but with a medium-lived and one with short-lived. With eleven HBs, P15428 (Rossinone) is the target with more HBs formed despite this only two of them, ILE91 and ASN92 are considered long-lived with occupancies of 62%; four medium-lived are also found, as well as five short-lived HBs. P00352 is the target with the higher occupancy with a value of 93% with the residue GLU188, which means that this HBs is almost maintained during all the MD simulation. Two other residues as GLU261 and GLU392 are also considered long-lived HBs, but with 65% and 52% of occupancy. Only one medium-lived HB was formed while six short-lived were present. P11511 had one long-lived HB with an occupancy of 51% and there was one short-lived HB. P04798 is the only target where any long or medium-lived HB was founded, and the six residues found, were short-lived. P16662 has one medium-lived HB and two short-lived HBs. As said, P15428 is studied twice, and in this case, the residues involved on the formation of HBs are not the same in both cases, except GLN149. This is due to the size of the binding cavity, which is quite large, favoring the different location of the molecules inside the cavity while performing the docking calculations.

- Root-mean-square fluctuation

When a dynamic system such as a protein fluctuates about some well-defined average position, the RMSF of atomic positions can be calculated from the MD trajectory. That is, RMSF measures the amplitude of atom motions during simulation (Figure 8) [41,42].

Marine Drug Discovery



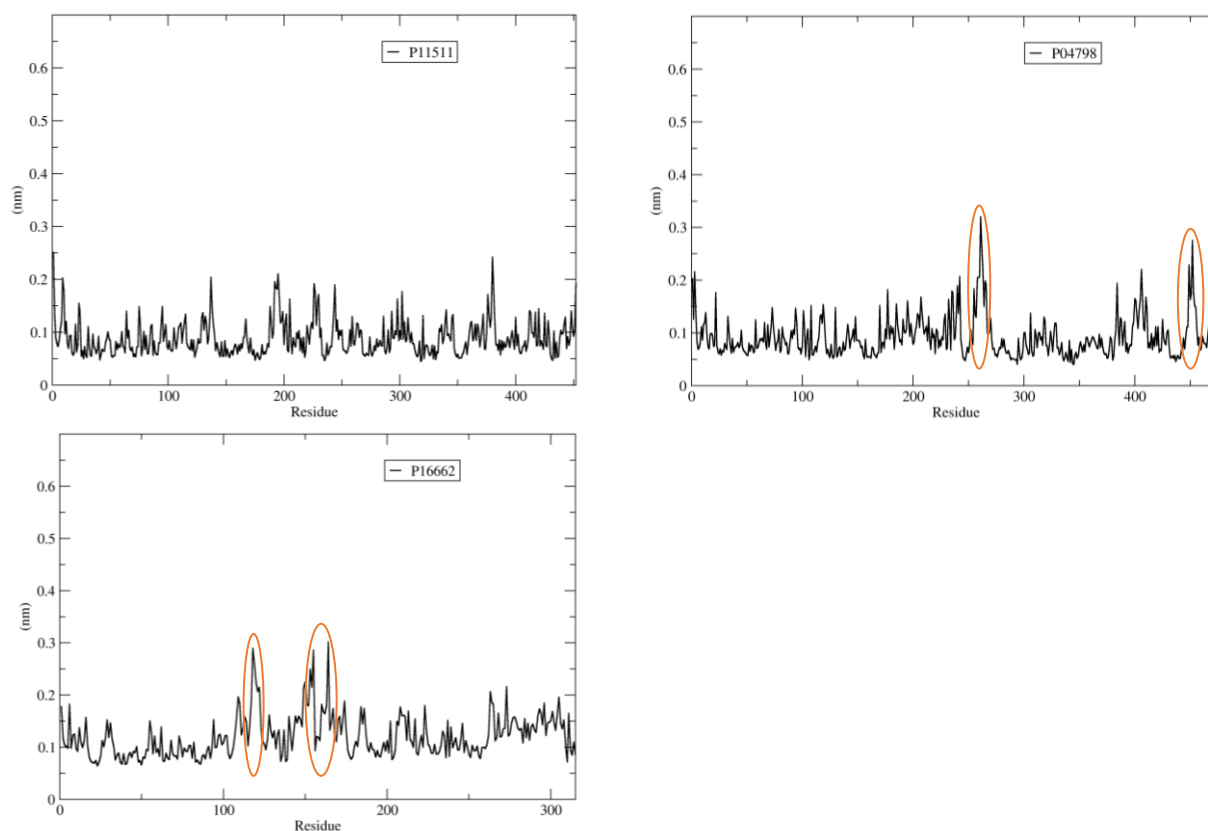


Figure 8. Root-mean-square fluctuation (RMSF) per residue (X-axis) values of each system along the MD simulation. The highest fluctuations (> 0.25 nm) detected have been highlighted with orange circles.

Interestingly, after an exhaustive analysis of each system, the results from the RMSFs of each aminoacid, highlight the flexible regions of the systems. RMSFs values higher than 0.25 nm (orange circles Figure 8) are characteristic of aminoacids residues belonging to flexible regions. Eight out of the 11 analyzed systems display RMSFs values higher than 0.25 nm, which means that remaining there, both P15428 and P11511 are, in general, rigid structures. Q9Y463 has a peak of fluctuation higher than 0.35 nm around residues 305-310. P49759 has fluctuation with values higher than 0.30 nm between residues 160-170. This system has another point of fluctuations around residues 40, 50, 195 and between 260 and 300, despite these has low values (< 0.25 nm). O15530 seems to be a more rigid system, even though before the highest fluctuation observed on residues 52-58, two other lower peaks can be observed indicating that there is a region between residues 25 to 60 which is more flexible than the rest of the protein. P00491 is also a system with low fluctuations were only a fluctuation of 0.28 nm can be seen around residue 255. P31749 is the system with more peaks of fluctuation and they are placed around residue 20, between 98 and 108 and around 297. Also, other lower peaks can be seen between residues 50 and 100, and again between 150 and

175. P00352 is a rigid complex but with one unique and high peak of fluctuation (> 0.4 nm) between residues 130 and 150. P04798 has a first peak (> 0.3 nm) over 250-275 but the previous residues between 230 and 245 show some fluctuations, and another lower fluctuation can be seen near the end, at residue 450-455. The last system is P16662, with three peaks lower than 0.3 nm, between residues 120 and 130, and between 150 to 170, showing that these regions of the protein are quite flexible. In general, for some RMSFs increase abruptly at the N and C-terminal residues of the protein, probably because these terminal parts are more flexible because they are usually more exposed to the solvent.

With all these results, it is possible to validate the binding of each molecule to the binding site selected, as all the residues involved on the binding (mainly those with long and medium-lived HBs occupancies) are placed on the rigid regions of the proteins, while the most flexible regions do not affect the binding. Furthermore, a detailed study of Rg and RMSF indicates a trend. Q9Y463 and P16662, with a clear Rg incompactness also show high values of fluctuations. The same pattern can be seen for P49759 and P31749, where the compactness can be questioned and the RMSF also show elevated values. P00352 shows one high peak in the RMSF, perhaps related to the highest Rg observed from all the systems. On the other side, P11511 and both P15428 do not show any remarkable fluctuation in the RMSF, and the Rg reveals stability and compactness. In between these extreme cases, P04798, P00491 and O15530 also show stability, with P00491 being the most stable system although O15530 shows the highest peak of fluctuation of all the studies systems.

Molecular Mechanics/Generalized Born Surface Area

The molecular mechanics energies combined with generalized Born and surface area continuum solvation (MM/GBSA) are popular computational approaches to estimate free energy of binding of small molecules to proteins [43,44]. These methods are used to predict ligand-binding affinities based on docking or MD simulations to get a more realistic view of the interaction of docked complexes. The obtained energies are more realistic than those obtained after docking calculations, as it is generally accepted that they outperform docking results, allowing a better ranking of the analyzed compounds [45]. These results although improve docking binding energy values, are far to be biological comparable. In our case, and following similar approaches, we applied reweighting techniques, specifically, MM/GBSA over the generated MD trajectories for post-processing docking results and the energy values were obtained (Figure 9).

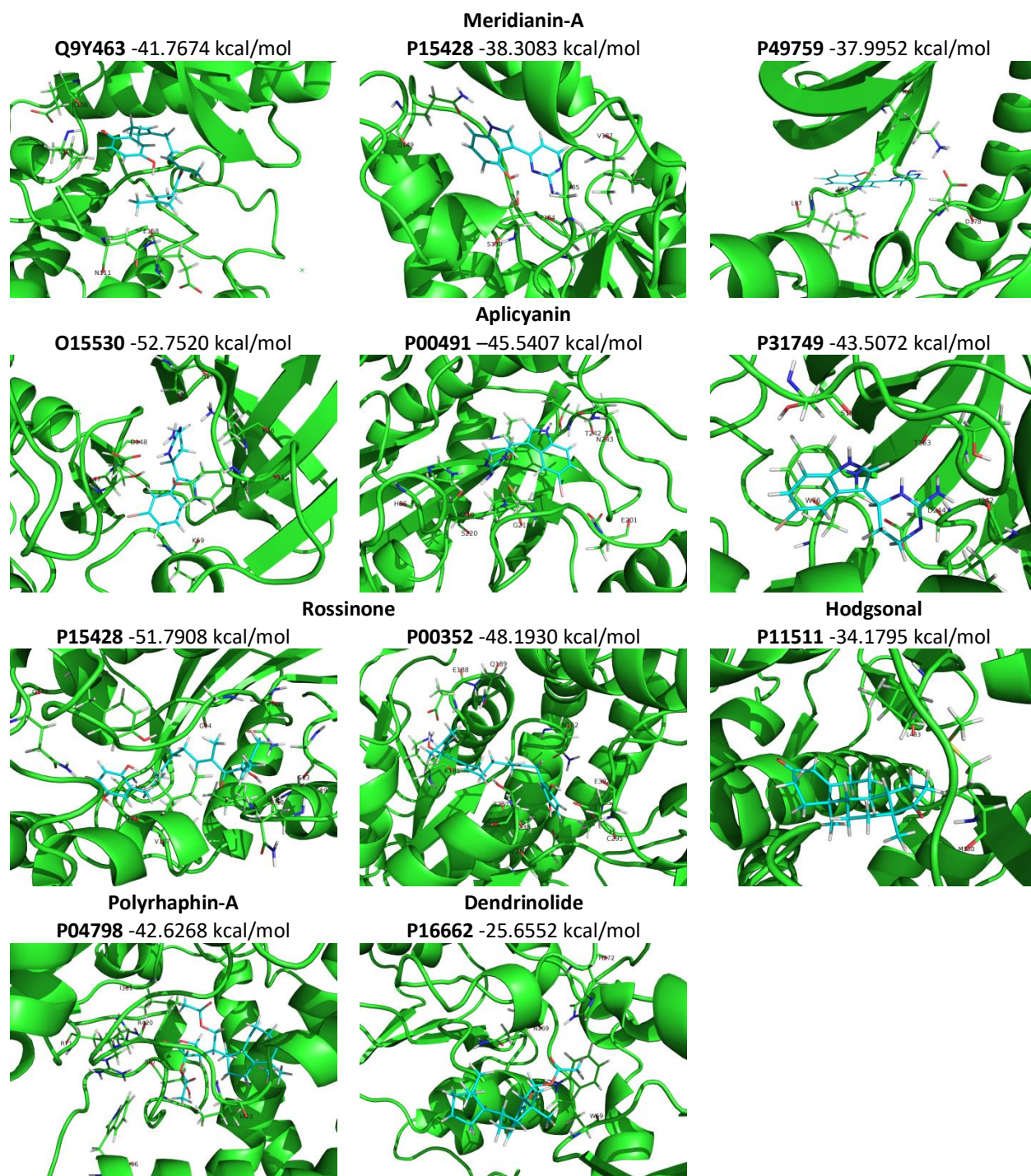


Figure 9. Images of the binding mode of each marine molecule inside the binding cavity of the corresponding target. Also, the summary of binding energy results after MD simulations with MM/GBSA calculations are indicated here. All energies are in kcal/mol.

From the negative total binding energy found on all the studied systems, we conclude that all of them are favorable complex in pure water. However, the computational results will not be equal to the

real (experimental) binding free energy, because we did not estimate the (dis-favorable) entropy contribution to binding, since this is a weak point of these methods. Meridianin-A, related with three different targets, Q9Y463, P15428 and P49759, does not show any significant difference on the binding energy that could suggest any selectivity over one target or another. Aplicyanin, also related to three targets, O15530, P00491 and P31749, seems to show preference for O15530, which is a target related to heart failure. Hodgsonal, Polyrhaphin-A, and Dendrinolide are only related with one target each, P11511, P04798 and P16662, respectively. The complex formed by Dendrinolide and P16662 target related to peripheral nervous system disease, is interesting because it is, by far, the lower binding energy found. However, this target shows incompactness in the Rg analysis and also has elevated fluctuations, which leads to think that this is the less favorable complex. Rossinone, is related with two targets, P15428 and P00352, has good results on the qualitative analysis while no significant differences can be observed on both targets. For P15428, target related to digital clubbing, the binding energy obtained is the highest one and revealed a significant difference (13 kcal/mol) between the binding of Rossinone or Meridianin-A (Figure 9).

Materials and Methods

Virtual Profiling

VP techniques able to automatically evaluate very large libraries of compounds using computer programs to finding targets for an input (query) molecule. In other words, given an initial molecule and by executing a similarity searching against a reference database (compound-target associations database) are able to find similar molecules (restricted by a cut-off) and thus finding plausible target to the input molecule. Basically, the way VP works is that if a molecule A is like a molecule B, the molecule A should interact with the target of molecule B. For that purpose, we employed Cabrakan, a 2D ligand-based VP software tool that compares molecules using 2D fingerprints and the assignment of biological activity, and Hurakan, a 3D VP tool that compares molecules according to their similarity using methods that take into account both physical and chemical properties of the molecules and their behavior with the environment when acting on a target [46,47].

Target selection

To relate the found targets with the specific selected pathologies we employed DisGeNET, a database that integrates information on gene-diseases associations. From these data, we focused on neurodegenerative and cardiovascular diseases [19].

Target modelling

From the selected targets, the search for 3D models was performed by exploring the Protein Data Bank (RCSB PDB) [48]. For those targets without crystallographic structures available or showing poor sequence representation (<30%), homology models were constructed by SWISS-MODEL, an automated protein structure homology model software tool to generate models of targets from the original human aminoacid sequence [49].

Toxicology prediction

Regarding ligand-based toxicology, the predictive models used here allow measuring Structure-Activity Relationship (SAR). The SAR concept means that the biological activity of a chemical can be related to its molecular structure, and when this is quantified, this relationship is known as QSAR. Here we used VEGA software tool, which is integrated in different QSAR models that predict biological toxicity at different levels [25]. The toxicity prediction depends on its reliability in a range that goes from 1 (low) to 3 (high). In this study, to gain statistical significance and reliability, all the available models in VEGA were employed. Because of that, the results of each category were averaged over all models used and then the results were classified according to its probability of being toxic in the following terms; no toxicity, low (<2), medium (2-2.75) or high (2.75-3). Regarding the toxicology based on target, T3DB was used, and over it a 2D Tanimoto based similarity search was performed using an inhouse developed tool [28].

Drug likeness evaluation

To evaluate the drug likeness of marine molecules, we use SuperTarget, a database which provides drug-target relations [50]. Only those relations that come from the well-known database DrugBank were selected [51]. Later, the 3D structure of those drugs selected was downloaded from PubChem compound database [52].

Docking calculations

Ligands were prepared to generate energetically minimized 3D coordinates and then docked into the active site of each target using Itzamna software tool [53]. Kin software tool was also used to perform blind docking calculations [54]. Best binding poses were determined and subsequently ranked based on their calculated binding energies.

Molecular dynamics simulation

The MD simulations were performed using NAMD software through four sequential steps: First, we run an energy minimization of the system, which is required to sort out any bad starting structures caused during the generation of the system [55]. The second and third steps are part of an equilibration simulation. The second step consists of simulating the system at NVT ensemble (at constant number of particles (N), volume (V) and temperature (T)), with position restraints on the solute to get the temperature at 300K. The third step consists of simulating at constant pressure (NPT) to fix the density of the system while the system, is heated incrementally until reaching the desired temperature [56,57]. The fourth and last step consists of running the production simulation for sufficient time so that property/phenomena of interest can be observed in required detail. In this study we performed short (1ns) simulations with a time step of 2 fs. For the modelling of the protein targets we used Amber ff99SB-ILDN, and for the modelling of the marine molecules, the General Amber Force Field (GAFF) set of parameters was used [58,59]. Ligand GAFF parameters were obtained using Antechamber, whereas the receptor structures were modelled using the leap module of Amber Tools [60,61].

Molecular dynamics analysis

Visual inspection of the trajectories and the HBs occupancies were performed using Visual Molecular Dynamics (VMD) [62]. The quality assurance of the thermodynamic parameters (temperature, potential, kinetic and total energy) and the review in terms of structure (Rg, RMSD and RMSF) were performed using GROMACS quality assurance tool [63,64].

MM/Generalized Born Surface Area

MM/GBSA rescoring was performed using the MMPBSA python algorithm contained within the Amber Tools suit. The snapshots generated at the end of MD simulations were used as input into the post-simulation MM/GBSA calculation of binding free energies [65].

Graphical representations

Graphical representations were prepared using PyMOL version 1.7 and the graphs are produced using the program GRaphing, Advanced Computation and Exploration of data (xmgrace) version 5.1.22 [66,67]. 2D marine molecules were prepared using RDKit python library [68].

Conclusions

Molecular docking and the techniques that derive from it, as MD simulations and MM/GBSA have demonstrated to be solid tools to be used during the process of drug discovery. Along this study, we employed different computational tools to validate, even from different points of views, the results obtained after each step of the process of elucidating the possible therapeutic potential of this set of marine molecules. In the first step, after the VP techniques elucidated a list of possible targets related to neurodegenerative and cardiovascular diseases, the marine molecule-target complexes were confirmed using docking and blind docking calculations. Following that, the toxicology predictions based on ligand and target revealed a tendency of these molecules to be toxic. Due to the importance of this kind of studies during drug discovery, further studies are required to completely assess the safety of these molecules. Maintaining the perspective and approach given to this study, and with the objective of comparing the possible drug-likeness properties of these marine molecules, a total number of 392 docking calculations were performed over known drugs and marine molecules against the targets selected on the first step and the results revealed that no significant differences were observed between them. This indicates that all the molecules of the set could act as drugs. As previously explained here, docking calculations have its limitations, and to give more reliability to our results, the induce-fit MD simulations and MM/GBSA techniques were employed. A deep study of the time averaged structural properties, such as Rg, RMSD, RMSF and HBs occupancies analyzed after the simulations allows us to suggest that in general, the eleven systems studied can be validated, despite some of them, as Q9Y463, P16662, P49759 and P31749, are less favorable than the others in terms of a qualitative analysis. In terms of energy values, it is worth to mention that the lower energy was obtained for the complex formed by Dendrinolide and P16662.

This study is a clear example on how the use of different CADD tools could help on the elucidation of different marine molecules with a potential therapeutic activity, in this case against cardiovascular, neurodegenerative, and some orphan diseases. Along this work, we identified, not only the diseases but the targets and the regions of the target sequence, where the marine molecules bind,

proving invaluable insights for further studies. Also, the introduction of the induce-fit MD simulation events allowed an improvement of docking results and an exhaustive qualitative analysis of each system could be performed. Finally, we believe that with our study, a general pipeline on the field of drug discovery has been established, which can be applied on the elucidation of therapeutic agents from marine molecules in future studies.

Acknowledgments: This research was partially supported by an Industrial Doctorate grant from the Generalitat of Catalonia to L.L.-P (DI 2016-051).

Author Contributions: M.S.-M., L. L.-P- and C.A. conceived the study and designed the experiments. L.L.-P. carried out the experiments whereas C.A. and M.S.-M. supervised them. All the authors analyzed and discussed the results, as well as wrote the manuscript.

Funding: Industrial Doctorate grant from the Generalitat of Catalonia to LL-P (DI 2016-051).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

Table A1. Summary of the results obtained performing docking simulations and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of marine molecules against each target (UniProt) with crystallographic structures (PDB) with ligand. To avoid false positives, each docking calculation was performed twice (R0/R1). All the energy values are in kcal/mol.

UniProt	PDB - Ligand	Docking Binding Energy R0/R1	MM/GBSA Binding Energy	UniProt	PDB - Ligand	Docking Binding Energy R0/R1	MM/GBSA Binding Energy
Hodgsonal				Aplicyanin			
P11511	3EQM-ASD	-8,1 / -8,2	-34,1795	P09874	1UK0-FRM	-8,2 / -8,2	-32,3788
	3EQM-HEM	-8,3 / -8,4	-22,4488	O15530	2R7B-253	-7,7 / -7,7	-22,8350
P15428	2GDZ-NAD	-6,9 / -6,9	-6,3093	P31749	3QC4-MP7	-8,1 / -8,1	-52,7520
Q99714	1U7T-NAD	-7,2 / -7,2	-0,0562	P00491	3O96-IQO	-8,6 / -8,6	-43,5072
	2O23-NAD	-6,7 / -6,7	-14,0999		1ULB-GUN	-7,2 / -7,2	-45,5407
Meridianin-A				Rossinone			
P00374	1MVS-DTM	-7,7 / -7,7	-28,9913	P00352	4WB9-NAI	-8,8 / -8,5	-48,1930
P48730	4KBK-1QG	-7,3 / -7,2	-34,4325	O00255	3U88-CHD	-6,8 / -6,8	-25,5673
Q13976	3OGJ-CMP	-6,7 / -6,7	-28,4003	P07550	4GBR-CAU	-9,7 / -9,7	-43,2079
	4QX5-CMP	-7,9 / -7,9	-13,5921		1U7T-TDT	-8,9 / -9	-39,1248
P49841	3PUP-OS1	-7,7 / -7,7	-29,1106	Q99714	1U7T-NAD	-8,1 / -8	-40,7381
	1U7T-TDT	-7,7 / -7,7	-25,4862		2O23-NAD	-8,6 / -8,5	-39,0733
Q99714	1U7T-NAD	-7,5 / -7,4	-21,3118	P15428	2GDZ-NAD	-9,2 / -9,3	-51,7908
	2O23-NAD	-7,6 / -7,6	-23,5035	P04637	5AB9-92O	-6,4 / -6,5	-31,0615
Q13627	4AZE-3RA	-8,2 / -8,2	-32,8447				
P15428	2GDZ-NAD	-8,4 / -8,4	-38,3083				
Pteroenone							
P15428	2GDZ-NAD	-7 / -7,1	-30,7877				
P07550	4GBR-CAU	-7,7 / -7,6	-36,9769				
Q99714	1U7T-TDT	-6,5 / -6,5	-23,3863				
	2O23-NAD	-6,5 / -6,4	-22,0587				

Table A2. Summary of the results obtained performing blind docking simulations and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of marine molecules against each target (UniProt) represented by homology models. Pocket means the cavity chosen to perform MD simulations. To avoid false positives, each docking calculation was performed twice (R0/R1). All the energy values are in kcal/mol.

UniProt	Pocket	Docking Binding Energy R0/R1	MM/GBSA Binding Energy	UniProt	Pocket	Docking Binding Energy R0/R1	MM/GBSA Binding Energy
Hodgsonal				Rossinone			
Q96KQ7	0	-7,8 / -7,9	-20,4079	Q96KQ7	1	-8,1 / -8,1	-26,9966
Meridianin-A				Q07343	0	-8,6 / -8,6	-41,0540
P49759	0	-9,3 / -9,3	-37,9952	Q16236	0	-8,1 / -8	-36,1075
Q9Y463	0	-8,8 / -8,1	-41,7674	P27815	0	-8,8 / -8,7	-37,3115
Q96KQ7	0	-7,3 / -7,3	-19,5406	Q08499	0	-9,2 / -9,3	-32,3433
Aplicyanin				Polyrhaphin-A			
Q96KQ7	0	-7,5 / -7,5	-21,8578	P24046	0	-7,5 / -7,6	-27,6724
Q16236	0	-6,6 / -6,6	-22,3084	P46098	1	-8,1 / -8,1	-30,8251
Pteroenone				P14867	0	-7,6 / -7,5	-14,0650
Q16236	0	-6,2 / -6,5	-15,8940	P04798	3	-9,5 / -9,2	-42,6268

Table A3. Summary of the results obtained performing blind docking simulations and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of marine molecules against each target (UniProt) represented by crystallographic structures (PDB) without ligand. Pocket means the cavity chosen to perform MD simulations. To avoid false positives, each docking calculation was performed twice (R0/R1). All the energy values are in kcal/mol.

UniProt	PDB	Pocket	Docking Binding Energy R0/R1	MM/GBSA Binding Energy	UniProt	PDB	Pocket	Docking Binding Energy R0/R1	MM/GBSA Binding Energy
Hodgsonal					Rossinone				
P04637	3Q01	1	-6,1 / -6,6	-11,3246	P04637	3Q01	1	-7 / -6,5	-23,5341
Meridianin-A					Dendrinolide				
Q13976	4KU8	0	-7,1 / -7,1	-19,9742	P16662	2O6L	0	-8,1 / -8,1	-25,6552

Table A4. Summary of the best affinities obtained after Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of marine molecules against each target (UniProt) and also, pathologies listed by target. Each target can be represented by crystallographic structures with ligands, without ligands, or homology models (HM). To avoid false positive, each docking calculation was performed twice (R0/R1). All the energy values are in kcal/mol.

UniProt	PDB-Ligand	Docking Binding Energy R0/R1	MM/GBSA Binding Energy	Pathologies
Hodgsonal				
P11511	3EQM-ASD	-8.1 / -8,2	-34,1795	Autism
Meridianin-A				
Q9Y463	HM	-8.1 / -8,8	-41,7674	Alzheimer
P15428	2GDZ-NAD	-8.4 / -8,4	-38,3083	Digital clubbing
P49759	HM	-9.3 / -9,3	-37,9952	Alzheimer
Aplicyanin				
O15530	3QC4-MP7	-8.1 / -8,1	-52,7520	Heart failure
P00491	1ULB-GUN	-7.2 / -7,2	-45,5407	Alzheimer
P31749	3O96-IQO	-8,6 / -8,6	-43,5072	Cardiovascular, schizophrenia
Rossinone				
P15428	2GDZ-NAD	-9.2 / -9,3	-51,7908	Digital clubbing
P00352	4WB9-NAI	-8.5 / -8,8	-48,1930	Parkinson
Polyrhaphin-A				
P04798	HM	-9.2 / -9,5	-42,6268	Cardiovascular disease
Dendrinolide				
P16662	2O6L	-8.1 / -8,1	-25,6552	Peripheral nervous system disease

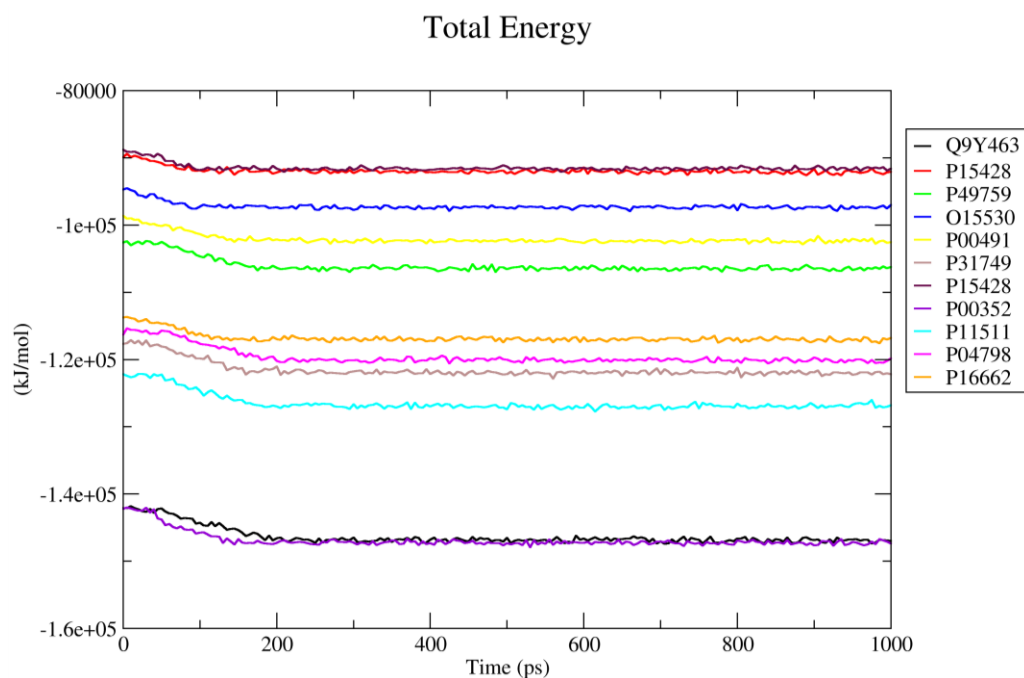


Figure A1. Total energy during the progress of the MD simulation of the eleven target-marine molecule systems. The color code of each target can be seen in the legend box.

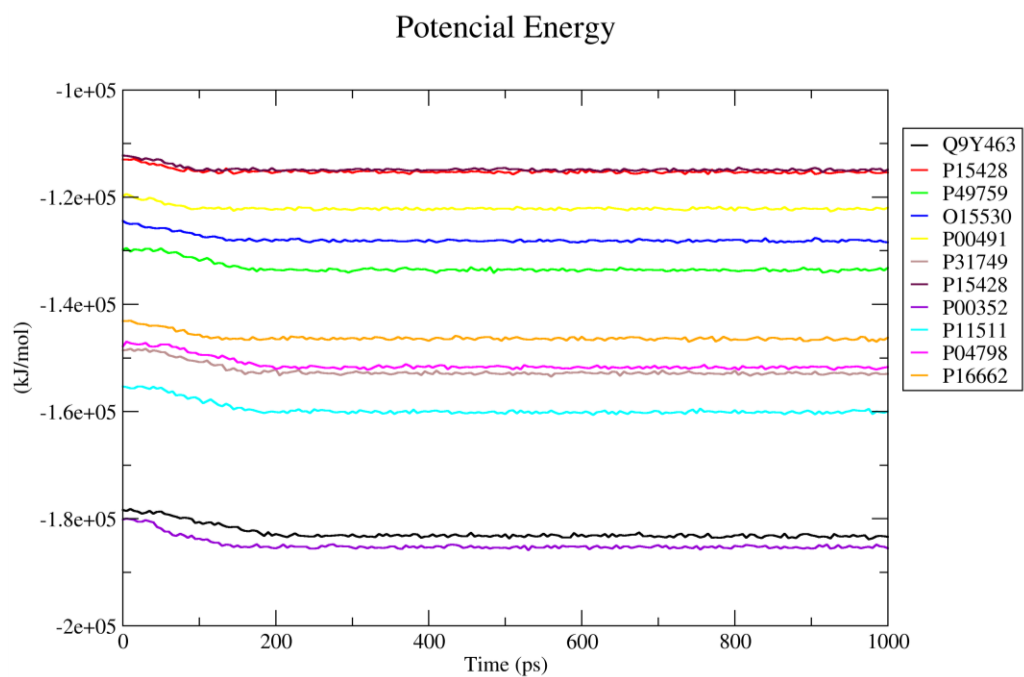


Figure A2. Potential energy during the progress of the MD simulation of the eleven target-marine molecule systems. The color code of each target can be seen in the legend box.

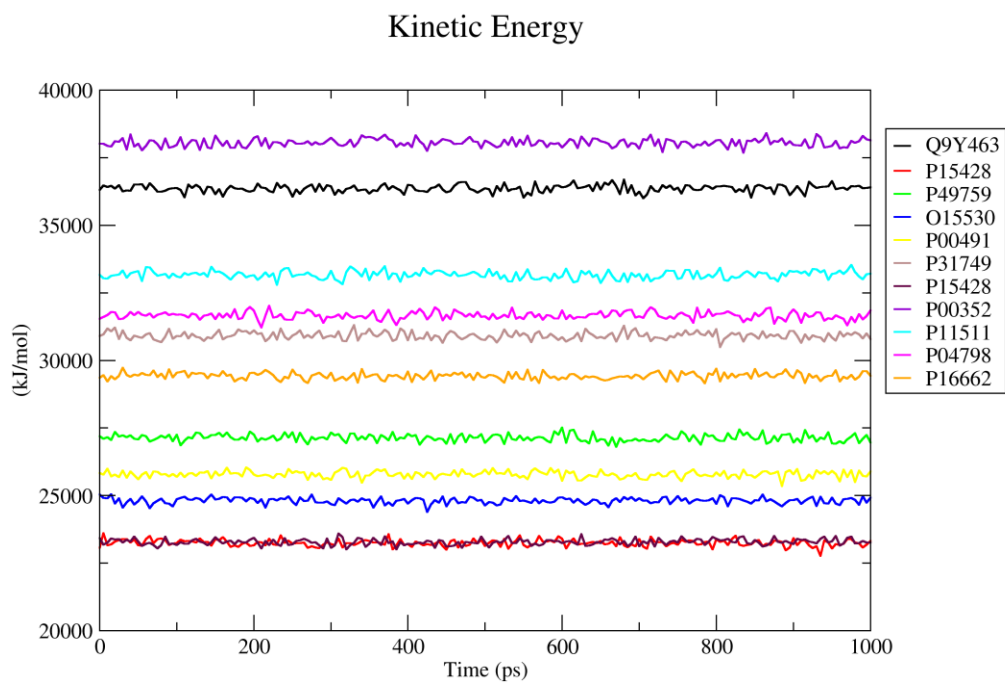


Figure A3. Kinetic energy during the progress of the MD simulation of the eleven target-marine molecule systems. The color code of each target can be seen in the legend box.

References

1. De Vivo, M.; Cavalli, A. Recent advances in dynamic docking for drug discovery. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2017, 7, e1320.
2. Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comput. Aided. Drug Des.* 2011, 7, 146–57.
3. Jorgensen, W. L. The many roles of computation in drug discovery. *Science* 2004, 303, 1813–1818.
4. Fischer, E. No Title; 1894; Vol. 27.
5. Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* 1982, 161, 269–288.
6. Koshland, D. E. Correlation of structure and function in enzyme action. *Science* 1963, 142, 1533–41.
7. Aravindhan, G.; Coote, M. L.; Barakat, K. Molecular dynamics-driven drug discovery: leaping forward with confidence. *Drug Discov. Today* 2017, 22, 249–269.
8. Lill, M. A. Efficient Incorporation of Protein Flexibility and Dynamics into Molecular Docking Simulations. *Biochemistry* 2011, 50, 6157–6169.
9. Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* 2006, 26, 531–568.
10. Sherman, Woody; Day, Tyler; P. Jacobson, Matthew; A. Friesner, Richard and; Farid, Ramy. Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. 2005.
11. Taboada, S.; Wiklund, H.; Glover, A. G.; Dahlgren, T. G.; Cristobo, J.; Avila, C. Two new Antarctic *Ophryotrocha* (Annelida: Dorvilleidae) described from shallow-water whale bones. *Polar Biol.* 2013, 36, 1031–1045.
12. Taboada, S.; Junoy, J.; Andrade, S. C. S.; Giribet, G.; Cristobo, J.; Avila, C. On the identity of two Antarctic brooding nemerteans: redescription of *Antarctonemertes valida* (Bürger, 1893) and description of a new species in the genus *Antarctonemertes* Friedrich, 1955 (Nemertea, Hoplonemertea). *Polar Biol.* 2013, 36, 1415–1430.
13. Taboada, S.; Bas, M.; Avila, C. A new *Parougia* species (Annelida, Dorvilleidae) associated with eutrophic marine habitats in Antarctica. *Polar Biol.* 2015, 38, 517–527.
14. Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on natural products for drug design. *Nat. Chem.* 2016, 8, 531–541.
15. W-H Li, J.; Vederas, J. C. Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science* (80-.). 2009, 325, 161–165.
16. Ebrahim, H.; El Sayed, K. Discovery of Novel Antiangiogenic Marine Natural Product Scaffolds. *Mar. Drugs* 2016, 14, 57.
17. Prachayasittikul, V.; Worachartcheewan, A.; Shoombuatong, W.; Songtawe, N.; Simeon, S.; Prachayasittikul, V.; Nantasenamat, C. Computer-Aided Drug Design of Bioactive Natural Products. *Curr. Top. Med. Chem.* 2015, 15, 1780–800.
18. Health at the glance-OCDE.
19. Pinero, J.; Queralt-Rosinach, N.; Bravo, A.; Deu-Pons, J.; Bauer-Mehren, A.; Baron, M.; Sanz, F.; Furlong, L. I. DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database* 2015, 2015, bav028-bav028.
20. Muster, W.; Breidenbach, A.; Fischer, H.; Kirchner, S.; Müller, L.; Pähler, A. Computational toxicology in drug development. *Drug Discov. Today* 2008, 13, 303–310.
21. Nature Reviews Drug Discovery. Nature Publishing Group November 1, 2007, pp. 853–853.
22. Schneider, G. Prediction of Drug-Like Properties. 2013.

23. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 2015, 10, 449–61.
24. Kumar, A.; Zhang, K. Y. J. Advances in the Development of Shape Similarity Methods and Their Application in Drug Discovery. *Front. Chem.* 2018, 6, 315.
25. VEGA-QSAR <http://ceur-ws.org/Vol-1107/paper8.pdf> (accessed May 13, 2016).
26. Proksch, P. Defensive roles for secondary metabolites from marine sponges and sponge-feeding nudibranchs. *Toxicon* 1994, 32, 639–655.
27. Nuijen, B.; Bouma, M.; Manada, C.; Jimeno, J. M.; Schellens, J. H.; Bult, A.; Beijnen, J. H. Pharmaceutical development of anticancer agents derived from marine sources. *Anticancer. Drugs* 2000, 11, 793–811.
28. Wishart, D.; Arndt, D.; Pon, A.; Sajed, T.; Guo, A. C.; Djoumbou, Y.; Knox, C.; Wilson, M.; Liang, Y.; Grant, J.; Liu, Y.; Goldansaz, S. A.; Rappaport, S. M. T3DB: the toxic exposome database. *Nucleic Acids Res.* 2015, 43, D928-34.
29. Bajusz, D.; Rácz, A.; Héberger, K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J. Cheminform.* 2015, 7, 20.
30. Valerio, L. G. *In silico* toxicology for the pharmaceutical sciences. *Toxicol. Appl. Pharmacol.* 2009, 241, 356–370.
31. Tzeng, S.-R.; Kalodimos, C. G. Protein activity regulation by conformational entropy. *Nature* 2012, 488, 236–240.
32. Fenwick, R. B.; Orellana, L.; Esteban-Martín, S.; Orozco, M.; Salvatella, X. Correlated motions are a fundamental property of β -sheets. *Nat. Commun.* 2014, 5, 1–9.
33. Liu, B.; Xu, R.; He, X. Kinetic Energy-Based Temperature Computation in Non-Equilibrium Molecular Dynamics Simulation. 2009.
34. Crehuet, R.; Field, M. J. A temperature-dependent nudged-elastic-band algorithm. *J. Chem. Phys.* 2003, 118, 9563–9571.
35. Merz, P. T.; Shirts, M. R. Testing for physical validity in molecular simulations. *PLoS One* 2018, 13, e0202764.
36. Jepps, A; Ayton, J; Evans, E. Microscopic expressions for the thermodynamic temperature. *Phys. Rev. E. Stat. Phys. Plasmas. Fluids. Relat. Interdiscip. Topics* 2000, 62, 4757–63.
37. Pirolli, D.; Sciandra, F.; Bozzi, M.; Giardina, B.; Brancaccio, A.; Rosa, M. C. De Insights from Molecular Dynamics Simulations: Structural Basis for the V567D Mutation-Induced Instability of Zebrafish Alpha-Dystroglycan and Comparison with the Murine Model. *PLoS One* 2014, 9, e103866.
38. Kufareva, I.; Abagyan, R. Methods of protein structure comparison. *Methods Mol. Biol.* 2012, 857, 231–57.
39. Maiorov, V. N.; Crippen, G. M. Significance of Root-Mean-Square Deviation in Comparing Three-dimensional Structures of Globular Proteins. *J. Mol. Biol.* 1994, 235, 625–634.
40. Yu, W.; Mackerell, A. D. Computer-Aided Drug Design Methods. *Methods Mol. Biol.* 2017, 1520, 93–94.
41. Fuglebakk, E.; Echave, J.; Reuter, N. Measuring and comparing structural fluctuation patterns in large protein datasets. *Bioinformatics* 2012, 28, 2431–2440.
42. Bornot, A.; Etchebest, C.; de Brevern, A. G. Predicting protein flexibility through the prediction of local structures. *Proteins* 2011, 79, 839–52.
43. Steinbrecher, T.; Labahn, A. Towards Accurate Free Energy Calculations in Ligand Protein-Binding Studies. *Curr. Med. Chem.* 2010, 17, 767–785.
44. Murcko, M. A. Computational Methods to Predict Binding Free Energy in Ligand-Receptor Complexes. *J. Med. Chem.* 1995, 38, 4953–4967.

45. Zhang, X.; Perez-Sanchez, H.; Lightstone, F. C. A Comprehensive Docking and MM/GBSA Rescoring Study of Ligand Recognition upon Binding Antithrombin. *Curr. Top. Med. Chem.* 2017, 17, 1631–1639.
46. Santamaria-Navarro, E.; Felix, E.; Nonell-Canals, A. Cabrakan <https://www.mindthebyte.com/>.
47. Santamaria-Navarro, E.; Nonell-Canals, A. Hurakan <https://www.mindthebyte.com/>.
48. Berman, H.; Henrick, K.; Nakamura, H.; Markley, J. L. The worldwide Protein Data Bank (wwPDB): Ensuring a single, uniform archive of PDB data. *Nucleic Acids Res.* 2007, 35, 2006–2008.
49. Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T. G.; Bertoni, M.; Bordoli, L.; Schwede, T. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 2014, 42, 252–258.
50. Günther, S.; Kuhn, M.; Dunkel, M.; Campillos, M.; Senger, C.; Petsalaki, E.; Ahmed, J.; Urdiales, E. G.; Gewiess, A.; Jensen, L. J.; Schneider, R.; Skoblo, R.; Russell, R. B.; Bourne, P. E.; Bork, P.; Preissner, R. SuperTarget and Matador: resources for exploring drug-target relationships. *Nucleic Acids Res.* 2008, 36, D919–22.
51. Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A. C.; Liu, Y.; Maclejewski, A.; Arndt, D.; Wilson, M.; Neveu, V.; Tang, A.; Gabriel, G.; Ly, C.; Adamjee, S.; Dame, Z. T.; Han, B.; Zhou, Y.; Wishart, D. S. DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Res.* 2014, 42, 1–7.
52. Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H. PubChem Substance and Compound databases. *Nucleic Acids Res.* 2016, 44, D1202–D1213.
53. Felix, E.; Santamaría-Navarro, E.; Sanchez-Martinez, M.; Nonell-Canals, A. Itzamna <https://www.mindthebyte.com/>.
54. Felix, E.; Nonell-Canals, A. Kin <https://www.mindthebyte.com/>.
55. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 2005, 26, 1781–1802.
56. Jorgensen, W. L.; Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *J. Comput. Chem.* 1998, 19, 1179–1186.
57. Andersen, H. C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* 1983, 52, 24–34.
58. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 2010, 78, 1950–8.
59. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* 2004, 25, 1157–1174.
60. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Antechamber, An Accessory Software Package For Molecular Mechanical Calculations. Natl. Institutes Heal.
61. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* 2005, 26, 1668–88.
62. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Molec. Graph.* 1996, 14, 33–38.
63. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* 2008, 4, 435–447.

64. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 2013, 29, 845–854.
65. Miller III, B. R.; McGee Jr., T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* 2012, 3314–3321.
66. Yuan, S.; Chan, H. C. S.; Hu, Z. Using PyMOL as a platform for computational drug design. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2017, 7, e1298.
67. 67.Grace User's Guide (for Grace-5.1.22) <http://plasma-gate.weizmann.ac.il/Grace/doc/UsersGuide.html> (accessed Apr 26, 2019).
68. RDKit: Open-source cheminformatics.

Section II

Elucidation of different pharmacophoric features of marine compounds and a precise in silico binding study

Section II

Chapter 2

Computer-aided drug design applied to marine drug discovery: Meridianins as Alzheimer's disease therapeutic agents

Chapter 2

Computer-aided drug design applied to marine drug discovery: Meridianins as Alzheimer's disease therapeutic agents

Laura Llorach-Pares^{1,2}, Alfons Nonell-Canals², Melchor Sánchez-Martínez² and Conxita Avila¹

1. Department of Evolutionary Biology, Ecology and Environmental Sciences, Faculty of Biology and Biodiversity Research Institute (IRBio), Universitat de Barcelona, 08028 Barcelona, Catalonia.
2. Mind the Byte S.L., 08028 Barcelona, Catalonia, Spain.

Published in: *Marine Drugs* 15(12), 366 (2017) doi: 10.3390/md15120366.

Abstract

Computer-aided drug discovery/design (CADD) techniques allow the identification of natural products that are capable of modulating protein functions in pathogenesis-related pathways, constituting one of the most promising lines followed in drug discovery. In this paper, we computationally evaluated and reported the inhibitory activity found in meridianins A–G, a group of marine indole alkaloids isolated from the marine tunicate *Aplidium*, against various protein kinases involved in Alzheimer's disease (AD), a neurodegenerative pathology characterized by the presence of neurofibrillary tangles (NFT). Balance splitting between tau kinase and phosphate activities caused tau hyperphosphorylation and, thereby, its aggregation and NFT formation. Inhibition of specific kinases involved in its phosphorylation pathway could be one of the key strategies to reverse tau hyperphosphorylation and would represent an approach to develop drugs to palliate AD symptoms. Meridianins bind to the adenosine triphosphate (ATP) binding site of certain protein kinases, acting as ATP competitive inhibitors. These compounds show very promising scaffolds to design new drugs against AD, which could act over tau protein kinases Glycogen synthase kinase-3 Beta (GSK3 β) and Casein kinase 1 delta (CK1 δ , CK1D or KC1D), and dual specificity kinases as dual specificity tyrosine phosphorylation regulated kinase 1 (DYRK1A) and cdc2-like kinases (CLK1). This work is aimed to highlight the role of CADD techniques in marine drug discovery and to provide precise information regarding the binding mode and strength of meridianins against several protein kinases that could help in the future development of anti-AD drugs.

Keywords: Computer-aided drug discovery/design; meridianins; Alzheimer disease; protein kinases; tau protein kinases; dual specificity kinases; marine natural products.

Resum

Les tècniques de descobriment o disseny de fàrmacs assistits per ordinador (DFAO) permeten la identificació de productes naturals que són capaços de modular les funcions de proteïnes que estan relacionades amb una determinada patologia, constituint una de les línies a seguir més prometedores en el descobriment de fàrmacs. En aquest treball, hem avaluat computacionalment l'activitat inhibidora trobada en les meridianines A – G, un grup d'alcaloides indòlics marins aïllats del tunicat marí *Aplidium*, contra diverses proteïnes quinases implicades en la malaltia d'Alzheimer (MA), una patologia neurodegenerativa caracteritzada per la presència de cabdells neurofibril·lars. La pèrdua de l'estabilitat entre les activitats de la tau quinasa i el fosfat és la causa de la hiperfosforilació de la tau i, per tant, la seva agregació i formació de cabdells neurofibril·lars. La inhibició de quinases específiques implicades en la seva via de fosforilació podria ser una de les estratègies clau per revertir la hiperfosforilació de la tau i representaria una aproximació per desenvolupar fàrmacs per pal·liar els símptomes de la MA. Les meridianines s'acoblen al lloc d'unió del trifosfat d'adenosina (ATP) de determinades proteïnes quinases, actuant com a inhibidors competitiu del ATP. Aquests compostos mostren esquelets molt prometedors per dissenyar nous fàrmacs contra la MA, els quals podrien actuar sobre les proteïnes tau, com ara la glycogen synthase kinase-3 beta (GSK3 β) i la casein kinase 1 delta (CK1 δ , CK1D o KC1D), i les quinases de doble especificitat, com la dual specificity tyrosine phosphorylation regulated kinase 1 (DYRK1A) i les quinases cdc2-like (CLK1). Aquest treball té com a objectiu destacar el paper de les tècniques de DFAO en el descobriment de fàrmacs marins i proporcionar informació precisa sobre el mode i la força d'unió de les meridianines contra diverses proteïnes quinases que podrien ajudar en el futur desenvolupament de fàrmacs contra la MA.

Paraules clau: Disseny de fàrmacs assistit per ordinador; malaltia d'Alzheimer; proteïnes quinases, proteïna quinasa tau, proteïnes de doble especificitat, productes naturals marins.

Introduction

Drug discovery is the process of identifying new molecules with a certain therapeutic activity. This process is very expensive in terms of money and time. Translating basic research to the market (going through drug discovery, preclinical and clinical studies) takes tens of years and costs billions of dollars. The average cost to develop a new molecular entity is estimated to be \$1.8 billion and requires about 13.5 years [1]. However, the usage of computational techniques at various stages of the drug discovery process could reduce that cost [2]. Hence, computer-aided drug discovery/design (CADD) methods are becoming very popular and during the last three decades have played a major role in the development of therapeutically important molecules [3,4]. CADD techniques cover several aspects of the drug discovery pipeline, ranging from the selection of candidate molecules to the optimization of lead compounds. For instance, virtual profiling (VP) methods can predict the biological profile as well as mechanisms of action (MoA) of a certain molecule; molecular modelling techniques, such as docking and molecular dynamics (MD), can predict ligand–target interactions in terms of binding mode and/or binding strength, allowing discrimination between candidate compounds [5,6]; virtual screening (VS) methods are able to find analogues (similar molecules) for a given compound(s) and/or build compound libraries from an input molecule(s); hit to lead (H2L) optimization techniques are used to design new molecules, improving an existing compound; absorption, distribution, metabolism, excretion and toxicity (ADMET) prediction techniques are able to predict the physicochemical properties of a given compound, i.e., information that can be coupled to H2L techniques in order to design better and safer drugs before synthesizing them.

A common classification of these techniques is based on the nature of the input molecule. In this sense, there are two general types of CADD approaches: structure-based drug design (SBDD) and ligand-based drug design (LBDD). In SBDD, macromolecular three-dimensional (3D) target structures, usually proteins, are analysed with the aim of identifying compounds that could interact (block, inhibit or activate) with them. In LBDD, chemical compounds are analysed in order to, for instance, find chemical analogues, explore their biological and/or toxicological profile, or improve their physicochemical and pharmacological characteristics with the aim of developing drug-like compounds (Figure 1) [7,8].

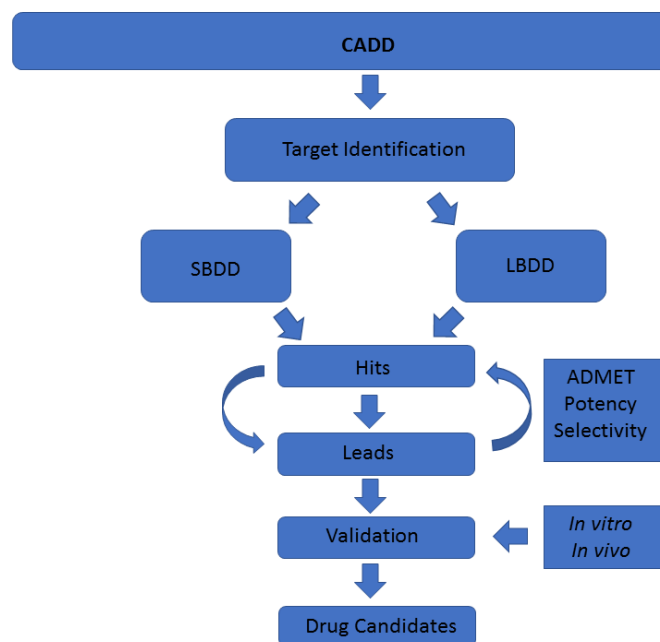


Figure 1. Schematic representation of the computer-aided drug discovery/design (CADD) techniques depicting a drug discovery pipeline. Original from the author.

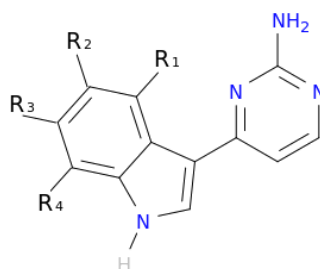
Historically, most new drugs have been designed from natural products (secondary metabolites) and/or from compounds derived from them [9]. Natural products have thus been a rich source of compounds for drug discovery, and often, feature biologically relevant molecular scaffolds and pharmacophore patterns that have evolved as preferred ligand–protein binding motifs. The United States Food and Drug Administration (US FDA) revealed that between 1981 and 2010, 34% of those medicines approved were based on small molecules from natural products or direct derivatives of them [10,11]. The identification of natural products that are capable of modulating protein functions in pathogenesis-related pathways is one of the most promising lines followed in drug discovery [12]. Therefore, natural products constitute a huge source of inspiration in drug design [13].

An example is Alzheimer’s disease (AD), a neurodegenerative pathology that constitutes the most common type of dementia (60–80% of the total cases), characterized by the presence of neurofibrillary tangles (NFT) primarily composed of abnormal phosphorylated tau and senile plaques (SP). Nowadays, despite its high incidence, there is still no specific treatment approved to cure this disease. Tau phosphorylation is regulated by a balance between tau kinase and phosphate activities. Splitting of this balance was considered to cause tau hyperphosphorylation and thereby its aggregation and NTF formation [14,15]. Due to that fact, inhibition of specific tau kinases or kinases involved in tau

phosphorylation pathway, could be one of the key strategies to reverse tau phosphorylation and, ultimately, fight AD [16].

The main relevant protein kinases involved in tau phosphorylation have been grouped into two classes: tau protein kinases and dual specificity kinases. The first group contains proteins such as glycogen synthase kinase-3 beta (GSK3 β), that phosphorylates tau at different sites (specifically at 42 sites, 29 of them phosphorylated in AD brains) and casein kinase 1 delta (CK1 δ), a non-proline-directed protein kinase (non-PDPK) that regulates the microtubule dynamics through tau phosphorylation at 46 sites (25 of them phosphorylated in AD brains). The second group contains proteins such as dual specificity tyrosine phosphorylation regulated kinase 1 (DYRK1) that self-catalyse their autophosphorylation and behave as serine/threonine kinase that phosphorylates tau and the transcription factor cyclic adenosine monophosphate-response element binding (cAMP-CREB), and an evolutionarily conserved group of dual specificity kinases cdc2-like kinases (CLKs), which play an important role in the regulation of ribonucleic acid RNA splicing and are involved in the pathology of AD by phosphorylating the serine residues in arginine-rich (SR) proteins [14,15,17–19].

Among natural products, those of unexplored marine world origin are of great interest in the discovery of novel chemical structures, since they harbour most of the biodiversity of the world [20,21]. For instance, compounds from marine invertebrates may possess interesting pharmacological activities. Examples include Porifera, Cnidaria, Bryozoa, Mollusca and Tunicata [22,23]. However, although very interesting and useful from a pharmacological point of view, obtaining these compounds is difficult, both from technical and biological points of view; technically, because specimens have to be collected by hand using scuba diving or by trawling (both expensive, logistically difficult, and time consuming), and biologically, due to their marine habitats and due to the fact that they are usually unculturable [23]. All these factors, together with the adequate implementation of the Nagoya Protocol and the bioavailability of marine natural products, result in CADD contributions being highly relevant, since no biological sample is needed to perform an *in silico* analysis [24]. This also alleviates some of the marine drug discovery difficulties, such as the quantity of natural product necessary to be used in further clinical studies.



Meridianin A	R ₁ = OH, R ₂ = H, R ₃ = H, R ₄ = H
Meridianin B	R ₁ = OH, R ₂ = H, R ₃ = Br, R ₄ = H
Meridianin C	R ₁ = H, R ₂ = Br, R ₃ = H, R ₄ = H
Meridianin D	R ₁ = H, R ₂ = H, R ₃ = Br, R ₄ = H
Meridianin E	R ₁ = OH, R ₂ = H, R ₃ = H, R ₄ = Br
Meridianin F	R ₁ = H, R ₂ = Br, R ₃ = Br, R ₄ = H
Meridianin G	R ₁ = H, R ₂ = H, R ₃ = H, R ₄ = H

Figure 2. Structures of meridianins A–G.

To exemplify and highlight the power of CADD techniques in marine drug discovery, as part of an ongoing study of bioactive marine molecules from benthic invertebrates, in this paper we evaluated and reported the inhibitory activity found in meridianins A–G (Figure 2), a group of marine indole alkaloids consisting of an indole framework connected to an aminopyrimidine ring, isolated from specimens of the tunicate genus *Aplidium*, against various protein kinases involved in AD.

Results

Virtual Profiling

In a previous VP study (unpublished data not shown here), we observed that meridianins could bind to diverse targets involved in different diseases associated with aging or neurodegenerative pathologies, such as AD and Parkinson's disease, cancer and cardiovascular diseases (Figure 3). The found targets are of special interest as they are involved in several diseases that affect millions of people worldwide, having a huge social incidence and also, in most cases, there is no cure for them. Regarding AD, the most common disease in which meridianins could have a therapeutic role according to our results (Figure 3), GSK3 β , CK1 δ , DYRK1A and CLK1 (four kinases involved in it) could be targeted. This finding can be easily checked in the literature, confirming that meridianins can bind to these kinases. Moreover, it can also be confirmed that the target analysis results are trustworthy, since not only is the involvement of meridianins in AD disease found in the literature, but the role of meridianins as anti-cancer agents can also be easily checked [25,26].

With these results in hand, the four kinases GSK3 β , CK1 δ , DYRK1A and CLK1 were selected for further analysis due to the prevalence of AD as the most common meridianin therapeutic target.

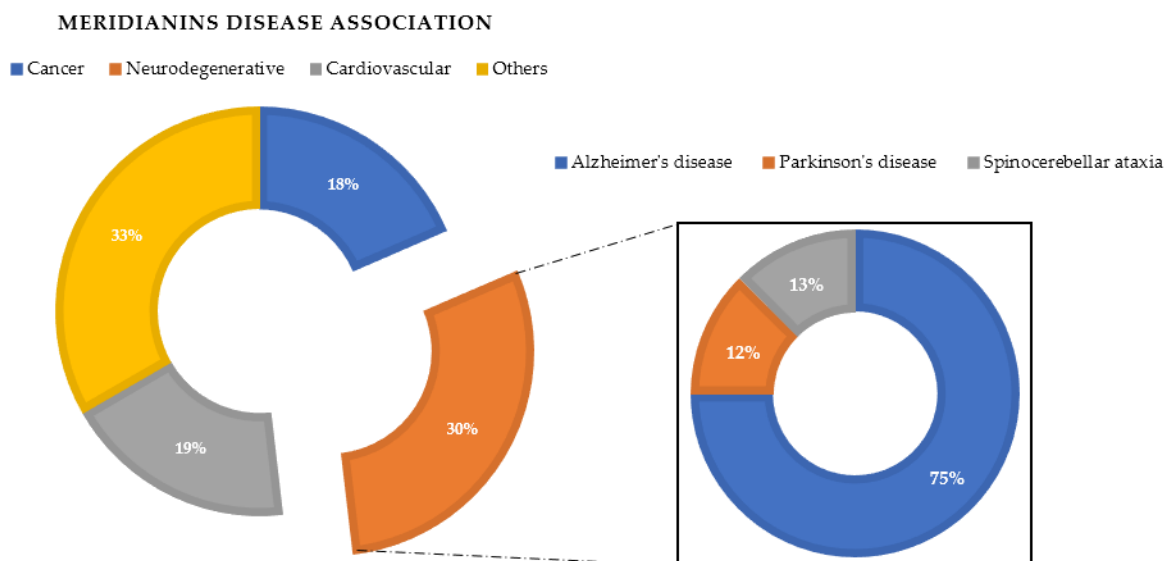


Figure 3. Disease association performed after a virtual profiling (VP) using meridianin A as a seed. Original from the author.

Structural and Sequence Analysis

Despite that the structural catalytic domains of most kinases are highly conserved, sequence alignment studies revealed some differences (Figure 4). The kinase catalytic domain, referred to as the hinge region, is divided into two lobes; the N-terminal mostly consists of β -sheets, whereas the C-terminal lobe is mainly helical. According to various authors, the adenosine triphosphate (ATP)-binding pocket of protein kinases can be divided into five regions: adenine region, sugar pocket, hydrophobic regions (I and II) and the phosphate-binding region [27–29]. In addition to this division, some recent studies have identified another important region: the glycine-rich loop, which is defined by the GxGxxG motif and is highly conserved among the protein kinase family. This region is suggested to significantly contribute to the potency and selectivity of binding inhibitors [29–31]. The glycine-rich loop and the hydrophobic pocket are placed in the so-called N-terminal region, while the sugar pocket and phosphate-binding region are located over the so-called C-terminal region. The adenine region is situated in the middle of these regions. We have found that meridianins are able to bind to all these regions, with a different binding strength depending on their chemical structure.

			GRR	HP						
P49841	GSK3B_HUMAN	35	SKVTTVVATPGQGPDRPQEVSYDTKVV	NGNSFGV	VVYQAKLCD-SGELV	IKKVLQDKRF	93			
P48730	KC1D_HUMAN	1	-----MELRVGNRYRLGRK	GSGSFGD	IYLGTDIA-AGEEVA	ILECVKTKH	46			
Q13627	DYR1A_HUMAN	142	DDNYDYIVKNG---EKWMDRYEIDSLIGKGS	GG	VKAYDRV-EQEWA	IKI IKKKAF	196			
P49759	CLK1_HUMAN	144	DEEGHLCIQSG---DVLRSARYEIVDTL	GEGA	SK	VECIDHKAGGRHV	IVKKNVDRY	199		
			AR							
P49841	GSK3B_HUMAN	94	KNRELQIMRKL DHCNIVRLRYFFYS	SGEKD---	EV-YLNLVLDYV	PETVYRVARHYSRA	149			
P48730	KC1D_HUMAN	47	-PQLHIESKIYKMMQGG-----VG	IPTIRWCGAEGDYNVM	MELLGPSLEDLF---	NFC	96			
Q13627	DYR1A_HUMAN	197	LNQAQI	VRLELMNKHDT	EMKYI	HLKRHF	MFRN-HLCLV	MM	SYNLYDLLRNTNFR	255
P49759	CLK1_HUMAN	200	CEAARS	IQVLEHLN	TTDPNSTFR	QMLEWFEH	HG-HICIV	GL	GLSTYDFIKENGLF	258
			SP							
P49841	GSK3B_HUMAN	150	KQTLFVIYVKLYMYQLFRSLAYIHSF--	GICHRDIK	PNL	LDPD-----	192			
P48730	KC1D_HUMAN	97	SRKFSCLKTVLLADQMISRIEYIHSKN--	FIHRDV	KPDNF	MGLGK-----	141			
Q13627	DYR1A_HUMAN	256	--GVSLNLRKFAQQMCTALLFLATPELS	IIHCDL	KPENI	LCNPK-----	299			
P49759	CLK1_HUMAN	259	--PFRLDHIRKMAYQICKSVNLFHSNK--	LTHTDL	KPENI	IFVQSDYTE	EA	YNPKIKRDER	314	
			BBP							
P49841	GSK3B_HUMAN	193	---TAVLKLC	FGSAKQL	VLRGEPN--V-----	SYICSRYYRAPELIFGATDYTSSIDVW	241			
P48730	KC1D_HUMAN	142	---GNLVYII	FGLAKKYR	DARTHQHI	PYRENKNLTGTARYASINTHLGI-EQSRDDLE	197			
Q13627	DYR1A_HUMAN	300	---RSATKI	FGSSCQL--	GQRI--Y-----	QYIQRFYRSPEVLLGM-PYDLAIDMW	345			
P49759	CLK1_HUMAN	315	TLINPDIKV	FGSATYD--	DEHH--S-----	TLVSTRHYRAPEVILAL-GWSQPCDVW	363			

Figure 4. Amino acid sequence alignment of GSK3 β , CK1 δ , DYRK1A and CLK1. In the image, only the ATP-binding pocket residues are shown. In blue, the key residues are conserved between all kinases. Green shows those conserved residues between tau protein kinases GSK3 β and CK1 δ , and red shows those conserved in dual specificity kinase DYRK1A and CLK1. Key residues refer to the residues implied in the binding of all the meridianins shared by the different targets and that are evolutionary conserved. The orange boxes represent the diverse region of the adenosine triphosphate (ATP) binding pocket. GRR: glycine-rich region; HP: hydrophobic pocket; AR: adenine region; SP: sugar pocket; BBP: phosphate binding pocket.

As explained above, we analysed two classes of protein kinases, specifically four members of them. The core catalytic regions are conserved among all as they belong to the same enzyme subclass (EC 2.7) and protein family (protein kinase). However, this protein family is divided into subfamilies: serine-threonine protein kinases (EC 2.7.11), dual-specificity kinase (EC 2.7.12), protein-histidine kinases (EC 2.7.13) and other protein kinases (EC 2.7.99). Thus, it seems logical that the binding site may be more conserved among subfamilies, and even more so in lower classifications (sub-subfamilies such as tau protein kinases (EC 2.7.11.26) and dual specificity kinase (EC 2.7.12.1)) than among the whole family. Analysing our results, we have confirmed this trend. Several key residues (associated with the substrate/inhibitor binding mode and/or conforming the pocket(s)) are conserved between the four studied proteins (Figure 4), but a higher identity is observed by pairs. GSK3 β and CK1 δ share more catalytic residues between them than with DYRK1A and CLK1, and vice versa. This observation agrees with the finding of a common binding pattern between the four protein kinases plus another pattern per each subfamily.

***In silico* Binding and Interaction Analysis**

Meridianins bind to the ATP binding pocket of each of the selected targets, acting as ATP competitive inhibitors. Binding energies obtained after docking and MD simulations (summarized in Table A1) show a reasonably similar binding strength between the diverse meridianins and even among the four kinases. Despite that fact, it could be observed that meridianin F tends to show higher energies than the rest of the compounds. Moreover, in general, meridianins present better binding interaction energies against CK1 δ , DYRK1A and CLK1 than GSK3 β . It must be said that these differences are hardly noticeable and cannot constitute a unique and definitive prioritization tool.

The binding mode per meridianin and target (that slightly changes between each complex) is summarized in Tables A2–A5. Comparing the interacting residues with the identified binders (summarized in Table A6), it is clearly observed that meridianins could behave as inhibitors of the analysed kinases. Moreover, analysing the observed binding mode together with the identified binders and the conserved residues (Figure 4, Tables A6 and A7), as mentioned above, some patterns of the general binding of meridianins to protein kinases could be extracted. It has to be highlighted that the majority of the residues found in these patterns are identified as binders.

For tau protein kinases, GSK3 β and CK1 δ , 5 binding residues are shared between each of them, whereas for dual specificity kinases, DYRK1A and CLK1, 12 are conserved. Moreover, there are four residues conserved along the four analysed targets (Figure 4 and Table A7). Concretely, these residues are an alanine and a lysine placed in the hydrophobic pocket, a leucine in the sugar pocket and an aspartic acid in the phosphate binding region. Regarding tau protein kinases, there is also an isoleucine shared by GSK3 β and CK1 δ . In the case of dual specificity kinases, there are eight other shared binders, specifically, two phenylalanine, three valines, two glutamic acids and one leucine conserved and identified as binders. Analysing the meridianin binding mode by focusing on the conserved amino acids also identified as binders, we have found that two of them, A83 and K85 placed in the hydrophobic pockets, are present in all meridianin binding modes over GSK3 β and CK1 δ (in the latter case, numbered A36 and K38). For DYRK1A, three of the conserved residues are identified as key residues for the binding of all meridianins, specifically V173, L241 and L294, in the same way as for CLK1 (in this case numbering as V175, L244 and L295). In addition to these residues, others were found implicated in the general binding of meridianins not conserved through all the targets (Table A6), specifically, for GSK3 β I62, V70, L132 and D200, for CK1 δ I23, M82, L85, L135 and I148, and for DYRK1A K188, V222, F238, V306 and D307. Finally, CLK1 residues L167 and A189 were identified as key meridianin binders.

Besides the above-mentioned residues, there are other important residues per meridianin and target not present in the observed patterns that have a key role (Table A7), not dependent in a general behaviour but dependent on the particular nature of each meridianin and target (Tables A2-A5).

Glycogen synthase kinase-3 Beta

Meridianins (Figure 5) tend to be placed within adenine (LDYV motif) and the hydrophobic regions, formed by the conserved residues A83 and K85, in the catalytic cleft. The indole scaffold of the meridianins is wrapped by N-terminal I62, F67, V70, A83, K85 and C-terminal T138, Q185, L188, D200 residues together with the LDYV motif in the hinge adenine region. Core interaction residues stabilize meridianins by establishing hydrophobic contacts with I62, V70, A83, K85, L132, D200 and hydrogen bonds with I62, K85, D200. The observed results further suggest that meridianins establish interactions over the glycine-rich loop on GSK3 β , defined by the GNGSFG motif, as well as with D200, a residue present in the phosphate pocket. The fact that meridianins bind to I62, V70, A83, K85, L132, L188 and D200, previously identified as binders, highlights meridianins inhibitory nature against GSK3 β .

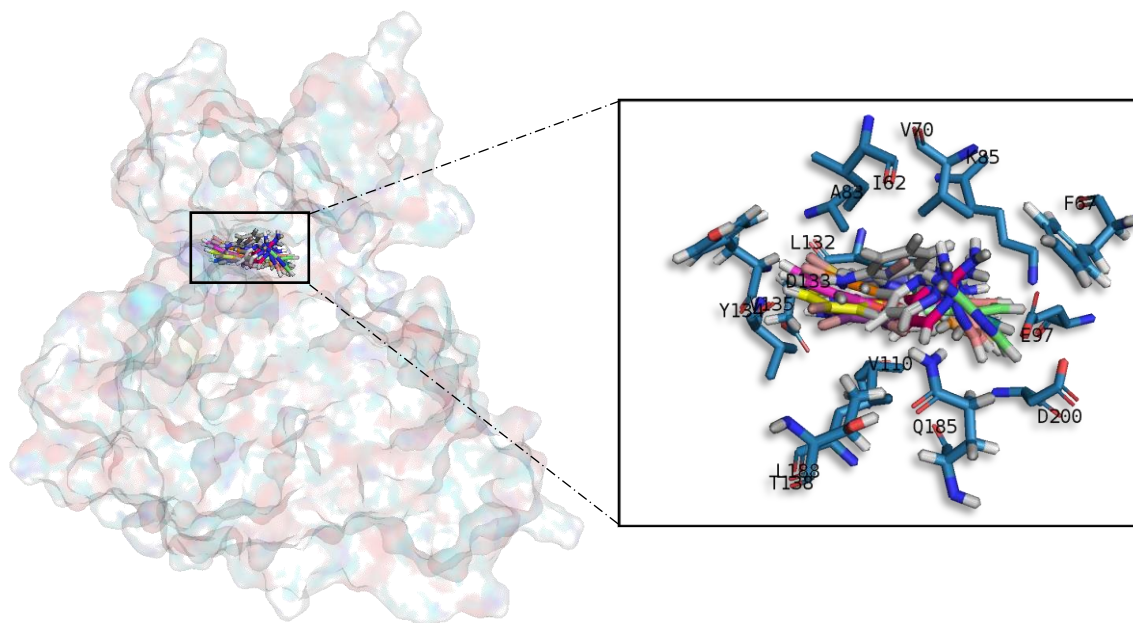


Figure 5. Meridianins A–G superposition over GSK3 β . Labelled ligand-active site amino acid residues involved in binding and the binding position of each meridianin models are enlarged. Original from the author.

Casein kinase 1 delta

All meridianin structures (Figure 6) share common interactions occupying the adenine region formed by the MELL motif. Meridianins are stabilized in the hinge catalytic region, establishing hydrogen bonds with A36, K38, M82, L85 and hydrophobic contacts with I23, K38, M82, L85, L135, and I148. Interestingly, it has also been observed that the indole group of the higher ranked poses has additional interactions with N-terminal I15, Y24, A36 and C-terminal D149 residues. It is important to remark that meridianins bind to the previously identified binder residues I23, A36, K38, M82, L85, L135 and I148, a fact that highlights meridianins inhibitory nature against CK1 δ .

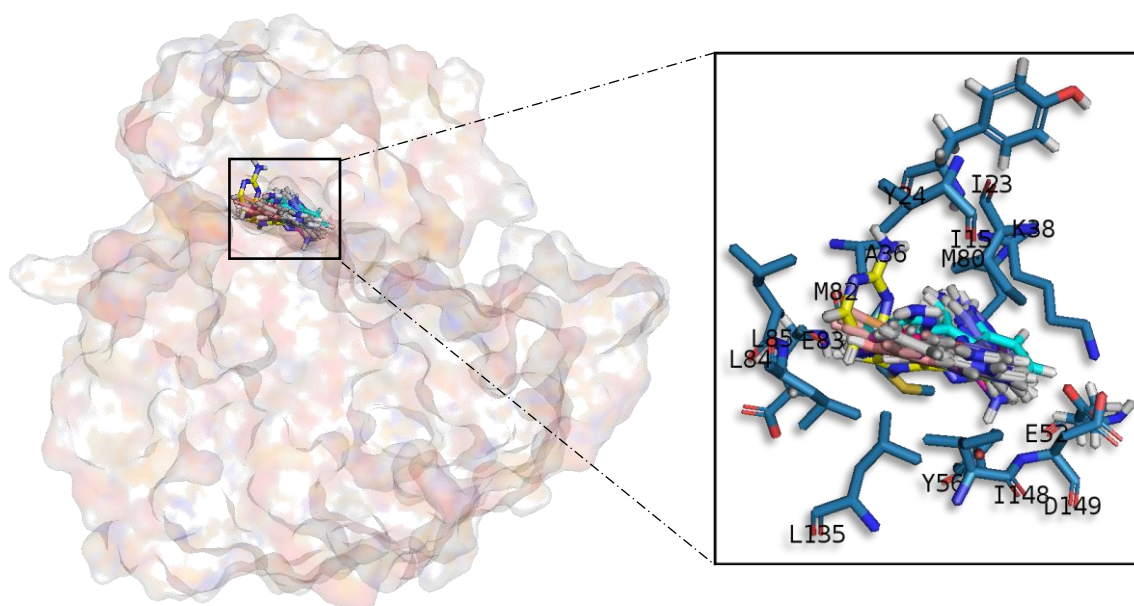


Figure 6. Meridianins A–G superposition over CK1 δ . Labelled ligand-active site amino acid residues involved in binding and the binding position of each meridianin model are enlarged. Original from the author.

Dual specificity tyrosine phosphorylation regulated kinase 1

Meridianins are placed on the C-terminal region over the phosphate and sugar pockets as well as the adenine motif FEML (Figure 7). Despite the fact that meridianins seem to interact with the N-terminal residue V173 and the hydrophobic pocket residue K188, the rest of the key interactions are established with residues placed over the C-terminal side. Meridianins establish hydrogen bonds with K188, L241 and V307 as well as hydrophobic contacts with V173, K188, V222, F238, L241, L294, V306 and V307. Moreover, they perform π -cation and π - π stacking interactions with F238, which belongs to the adenine motif. The inhibitory effect of meridianins against DYRK1A is confirmed by the fact that all

of them bind to V173, K188, V222, F238, L241, L294, V306 and V307, i.e., residues previously identified as binders.

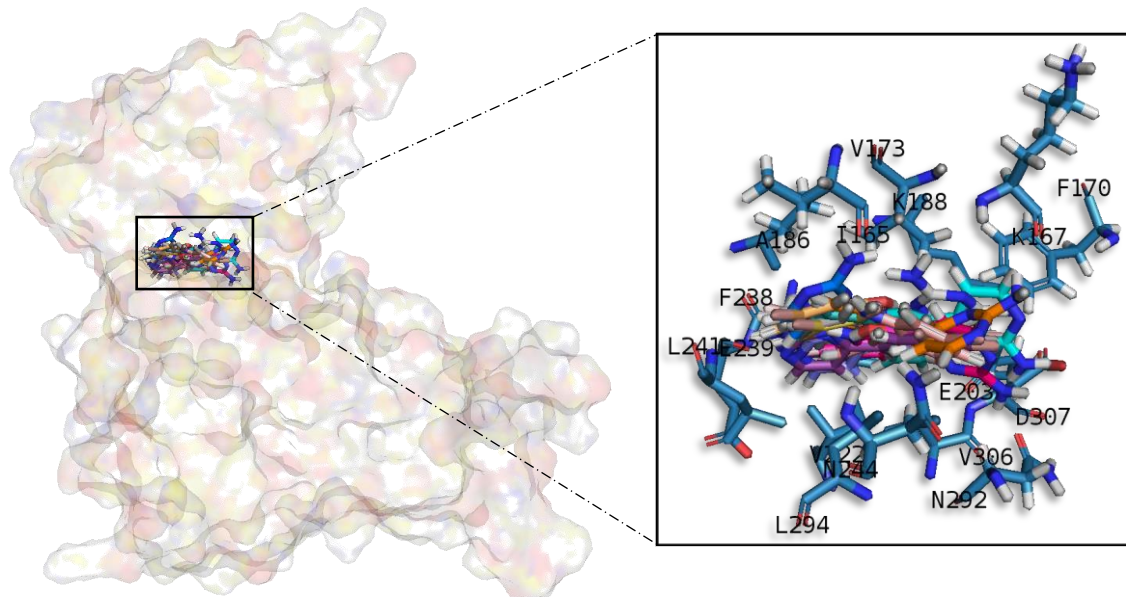


Figure 7. Meridianins A–G superposition over DYRK1A. Labelled ligand-active site amino acid residues involved in binding and the binding position of each meridianin model are enlarged. Original from the author.

Cdc2-like Kinases

Meridianin A–G conformations against CLK1 differ by pose, as can be observed in the superimposition shown below (Figure 8). In fact, over this target is where meridianins displayed a more different conformation between the family members. In general, all poses tend to be located near the glycine rich loop and the hydrophobic pocket, interacting with the adenine motif FELL through L244 by a hydrogen bond interaction. The different poses were well stabilized into the hinge catalytic pocket by establishing hydrogen bonds interactions with L167 and L244 and hydrophobic contacts with L167, V175, A189, L244 and L295, all of them previously identified as binders, a fact that underline their inhibitory nature against CLK1.

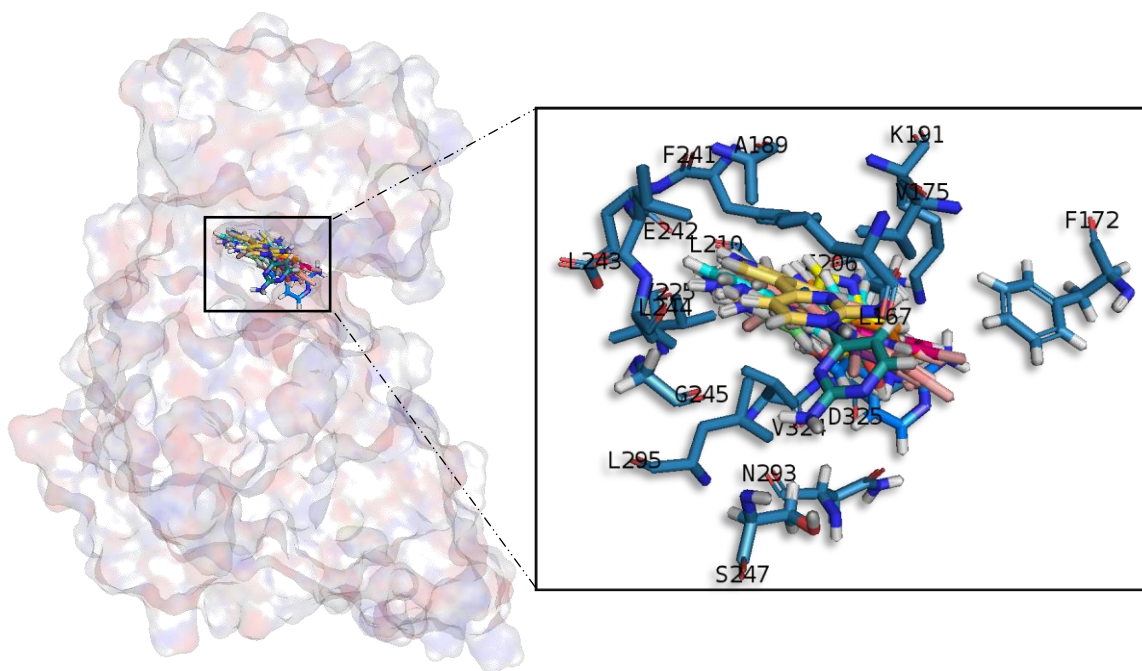


Figure 8. Meridianins A–G superposition over CLK1. Labelled ligand-active site aminoacid residues involved in binding and the binding position of each meridianin model are enlarged. Original from the author.

Selectivity

Since the results of *in silico* binding showed good interactions against the four studied targets, we wanted to know whether meridianins could be selective inhibitors of the studied protein families. Thus, we conducted a selectivity test consisting of analysing the meridianins binding over eight kinases (seven protein kinases and one non-protein kinase) with the aim of observing meridianin binding preference. This test included meridianins and three compounds derived from them, previously proposed as kinase inhibitors with a good selectivity for DYRK1A and CLK1 [25,32]. Our results show that meridianins and the derived compounds are able to bind to all the studied protein kinases, suggesting that they are not selective among them, although, for isocitrate dehydrogenase cytoplasmic (IDH1) and cGMP-dependent protein kinase 1 (PRKG1), slightly lower binding energies can be observed. Moreover, although compounds 1-3 tend to better interact with DYRK1A and CLK1, large differences are not observed in binding affinity between meridianins and their derived compounds (Table A8). In that sense, the derived compounds show a selectivity for DYRK1A and CLK1 respect to GSK3 β and CK1 δ , but not to all the tested kinases. Together, our results revealed the necessity to increase the selectivity of the meridianins and their, herein analysed, derived compounds.

Pharmacokinetic Properties

Due to the importance of pharmacokinetics (PK) studies during drug discovery, the whole set of meridianins and the three meridianins derived compounds were analysed, studying the ADMET properties for each molecule (Tables A9- A11).

Absorption properties

In Caco-2 permeability, two different models were used as in the first one (ML model), compounds 1 and 2 cannot be evaluated because they are out of the applicability domain (OAD). All the analysed molecules have high permeability according to our proprietary model; while using pkCSM meridianin G and compounds 2 and 3 show low permeability values, but they are almost considered as high (>0.9). LogS values confirm good solubility in water and good bioavailability for each compound. Intestinal absorption shows quite good percentages (absorbance >88%) for all the studied compounds, as molecules showing values lower than 30% would be considered to be poorly absorbed. Both the P-glycoprotein (Pgp) substrate and I/II inhibitor models show good concordance, and all of the studied molecules have been predicted to be Pgp substrates, and any of them could act as an inhibitor. The last absorption property studied was skin permeability, and results show values >-2.76 , which means reasonable low skin permeability.

Distribution properties

Log P values were calculated. The steady state volume of distribution (VD_{ss}) show by the studied molecules is low, as all are above 2.81 L/Kg, Log VD_{ss} > 0.45. For plasma protein binding (PPB) property, all the studied compounds have a probability of binding > 90%. Blood-brain barrier (BBB) permeability results show poor permeability for all meridianins and the three derived compounds. Compounds with a blood-brain permeability-surface area product (logPS) >-2 are considered to penetrate the central nervous system (CNS), and in that sense, compounds 2 and 3 could be considered as penetrants as they show slightly better results, i.e., logPS values of -1.88 and -1.99 , respectively. However, they are on the border, and the general tendency of all of them is to show poor penetration.

Metabolism properties

Cytochrome P450 interaction reveal that all the molecules in the studied sets are likely to be metabolised. All of the analysed compounds are able to inhibit the CYP1A2 isoform. Besides meridianin F and compounds 1 and 2 can also inhibit the CYP2C19 isoform, and compounds 1 and 3 the CYP3A4 isoform. Moreover, compound 2 can act as a substrate of the CYP3A4 isoform (Table A11).

Excretion properties

None of the analysed compounds is a substrate of organic cation transporter 2 (OCT2), which means that non-clearance problems and adverse interactions with co-administrated OCT2 inhibitors are expected. Moreover, total clearance was measured.

Toxicology properties

Regarding the maximum recommended tolerated dose (MRTD), our results show that only meridianins A, B and E have high (greater than 0.477 log (mg/kg/day) MRTD values, which means that a higher dose could be administrated, while the other compounds show lower values. AMES toxicity predicts mutagenic and carcinogenic characteristics and the results reveal that only meridianins A, B and E have no apparent toxicity. The human ether-a-go-go gene (hERG) I and II inhibitor method show that any of the studied molecules is likely an hERG inhibitor. Hepatotoxicity results point out that meridianins B and F may be associated with disrupted normal function of the liver. Skin sensitisation results show no adverse effects for dermally applied products. In summary, based on all analysed compounds, only meridianins A and E seem to be non-toxic and administrable with a possible high dose without presenting adverse toxic effects.

Discussion

CADD techniques have an enormous potential in drug discovery, especially when they originate from marine natural products, as they do not waste natural resources. As mentioned, there are numerous different methodologies enclosed within the term CADD [2,4]. Usually the methodology is chosen based on its applicability, advantages/drawbacks, previous studies in the field, and also the expertise of the authors. In that sense, general methods such as docking, MD or ligand similarity searches have been developed, as well as more specific techniques such as disease or target models [33–44]. Each technique requires a specific input and gives a specific output, aiming to solve one step of the drug discovery pipeline (Figure 1). However, although individual CADD methods can provide insight and solve many questions, their power is their strength when combined, as we show here. With the techniques employed in this study, we have mostly covered the drug discovery process able to be coped computationally. The methodologies we show in this work, as well as the way and the order in which we have used them, are addressed to cover a plausible general pipeline, which in our opinion is of general interest regarding marine molecules discovery. In previous years, many resources have been invested in biodiscovery (for instance, European funded projects such as PharmaSea, MaCuMBA, SeaBiotech,

BlueGenics or MicroB3) and some lead compounds have been designed, but a lot of information remains stored [45–49]. Using CADD techniques, this information could be easily analysed and, potentially, employed to find drug candidates. In summary, we have shown how starting from a molecule, we were able to provide lead compounds (although in this case we provide insights to construct them instead of fully designed compounds) against a certain disease. In that sense, and as we have commented above, we exemplified the role of CADD tools applied to marine drug discovery in general, and in this particular case, analysing the role of meridianins in AD, even more specifically, against four protein kinases involved in its pathology.

The four protein kinases studied here were previously described by other authors as meridianin targets [25,32,50,51]. This constitutes an excellent validation of our computational, blind, approach to identify the biological profile of meridianins. However, although in the literature the possible anti-AD activity of meridianins was reported and several compounds have been designed from them [25,32,50,51], several aspects have not been taken into account and analysed, from a target-based (structural) perspective, as we have done here.

A common observed feature of protein kinases inhibitors is that most of them usually interact with the phosphate binding groove, in the innermost part of the pocket. This is a rich polar region, with groups such as arginine or aspartate, that consequently can create hydrogen bonds with small molecules acting as inhibitors [52]. We observed that meridianins also show this trend, supporting their already mentioned general kinase inhibitory capacity. This, in addition to the fact that most of the meridianin binding residues are previously described as binders of known inhibitors, as well as the enzymatic assays that validated meridianin binding against the four studied kinases, also reinforce their tau protein and dual specificity kinase inhibitory capacity. As mentioned above, to exert this inhibitory capacity, meridianins show general binding trends against protein kinases in general and the studied targets in particular, but also specific features related to the nature of each of the targets. The understanding of these interactions (meridianin–target) and the identification of which of these characteristics are the most important to obtain good interactions is key in the design of meridianin-derived kinase inhibitors.

It was observed that for GSK3 β , the best scored meridianins C, D, E, and F (Table A1) establish hydrophobic contacts within the aminopyrimidine ring, revealing that this scaffold could be important in having optimal interactions. This highlights the fact that the most important interactions between GSK3 β and meridianins were on the glycine rich loop and the hydrophobic and phosphate pockets. For CK1 δ , analysing our *in silico* binding results, we observed that for the best scored meridianins C, D and F

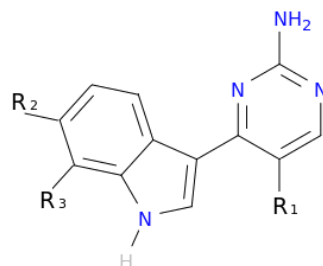
(Table A1), it seems that to increase the affinity of the ligand on this receptor, the aminopyrimidine moiety should be oriented towards the top of the hydrophobic pocket at the N-terminal region. Also, key interactions were observed in the adenine and sugar-phosphate pockets. Regarding DYRK1A, meridianins mostly tend to be located over phosphate and sugar pockets as well as the adenine motif FEML rather than the glycine rich loop. Best scored meridianins B, C, E, and F (Table A1), share similar conformations but with different orientation with respect to the rest of the analysed meridianins, a fact that could be exploited for future developments together with meridianins preferential placement over the phosphate and sugar pocket. For CLK1, our molecular modelling studies have revealed that the best interacting meridianins B, C, D and F (Table A1) tend to be located near the glycine rich loop and the sugar pocket.

In general, the orientation of meridianin indole scaffolds differs from one complex to another. Its preferential positioning is directed by hydrophobic interactions and steric effects, due to the aminopyrimidine ring position. In some models, it occupies hydrophobic region I, similar to many potent serine/threonine or tyrosine kinase inhibitors [27]. It must also be mentioned that for GSK3 β and CLK1, the preferred meridianin binding zones were located over the glycine rich loop (N-terminal). Nevertheless, over CK1 δ and DYRK1A, meridianins tend to be located over the sugar and phosphate region (both over the C-terminal region), correlating this fact with the slightly highest interacting energy observed after *in silico* binding experiments (Table A1). This could establish a new insight into future development of inhibitors.

Another interesting feature observed with respect to the meridianin binding mode is the presence of bromine. When present, interaction energies seem to be higher. The perfect example is meridianin F, which has two Br at R2 and R3, and has the best interaction energies for each of the studied targets with respect to the rest of meridianins. Emphasizing this issue, a pattern was observed within the two classes of kinases. For CK1 δ meridianins C (Br = R2), D (Br = R3) and F present the best interaction energies. In GSK3 β , meridianins D and F are among the three best interacting compounds. On DYRK1A, meridianins B (Br = R3), C and F are three of the four best interacting compounds and in CLK1, meridianins B, C, D and F are the ones that show the best energies. All these facts led us to hypothesize that Br on R2 and R3 on meridianins could be synonymous with potency and has to be taken into account for the design of new lead compounds against tau and dual-specificity kinases, in particular, and protein kinases in general. Interestingly, the most promising meridianin-derived compounds already designed (by Bharate and co-workers and Giraud and co-workers), are bromine-iodo derivatives (compounds 2 and 3) and non-iodinated bromine analogues (compound 1) (Figure 9) [25,32].

This fact supports our hypothesis about the influence of Br in the potency of binding showed by meridianins. According to our binding results, the derived compounds do not interact with target kinases stronger than do the meridianins. Therefore, we hypothesized that to design more potent inhibitors, the presence of Br atoms is key, but it is not enough. Playing with the different orientations and binding residues implicated in the observed patterns in meridianins-kinase binding should be also taken into account.

As protein kinases are a wide family of proteins involved in many cellular events, being selective against the desired ones is key, probably even more important than having a potent inhibitor, to avoid undesired effects. In that sense, our results show that both meridianins and the compounds reported by Bharate and co-workers, as well as Giraud and co-workers, could bind to different protein kinases with a similar strength [25,32]. In addition to that, the reported selectivity of the derived compounds for DYRK1A and CLK1 respect to GSK3 β and CK1 δ is observed, but it is not extensible to all the tested kinases. Going deeply into the results (Table A8), it could be observed that for IDH1 and PRKG1, the binding energies are slightly lower in comparison with the other targets. This fact is very relevant and could be explained because IDH1 is not a protein kinase. We put it in the pool of tested targets to see if out of the studied family, some selectivity could be observed. Regarding PRKG1, despite that it is a protein kinase member, the employed 3D structure contains an amino acid sequence that does not cover the kinase region. It was included to see what happened if despite being a protein kinase family member, the catalytic hinge region was not present. These findings allowed us to hypothesize that, despite meridianins do not show specific selectivity against any of the protein kinases tested, they do have a preferred binding to protein kinases. Moreover, this study validates the hypothesis that meridianins can act as protein kinase inhibitors. However, the low selectivity observed with respect to meridianins indicate that none of them is selective enough to properly act as AD therapeutic agent, even if able to inhibit the desired kinases. Although they could be a good starting point to design new drugs against AD, their selectivity should still be improved. To achieve that improvement, the presence of Br atoms is not enough. A rational design based on the structural differences and binding patterns observed along all meridianins should be carried out to obtain selective compounds that could have options to become an anti-AD drug. In that sense, the analysed derived compounds constitute an excellent example of how to improve meridianins to become therapeutic agents, but a new design is needed to overcome broader selectivity issues.



Compound 1	R ₁ =H, R ₂ =H, R ₃ =Br
Compound 2	R ₁ =I, R ₂ =Br, R ₃ =H
Compound 3	R ₁ =I, R ₂ =H, R ₃ =Br

Figure 9. Structures of the three compounds. Selected from Bharate et al. Table 4—Compounds 68–70 and Giraud et al. Figure 2 —Compounds 30,33,34 [25,32].

Potency and selectivity are important characteristics of a drug, but fulfilling certain ADMET requirements is also very important. The characterization of ADMET for the molecules being pursued as potential drug candidates is essential, as clinical failures of about 50% of the drugs under investigation are due to their inadequate ADMET attributes. In this regard, we have analysed the behaviour of all the studied meridianins and also the three compounds designed by Giraud and co-workers to evaluate if the implemented modifications improve the properties of the meridianins (Tables A9–A11) [25,32].

Meridianins and the three derived compounds show a potentially high, oral and intestinal, absorbance as well as reasonable low skin permeability. Probably one of the most relevant findings is that any of the studied compounds is able to cross the BBB by itself, which is essential for a drug that should act in the brain. Good penetration was not shown in the CNS in general. In addition to CNS entrance, the Pgp that seems to play a role in amyloid beta (A β) transport across the BBB and its modulation (inhibition) has been designed as a mechanism to improve CNS pharmacotherapy [53–56]. Unfortunately, any of the studied compounds has been predicted as an inhibitor, but as a substrate, which reinforces their inability to cross the gate into the CNS. Also, in relation to distribution properties, high PPB probabilities were observed as well as a low VD_{ss}, which means these compounds will have a lot of difficulties in diffusing or traversing cell membranes.

These compounds are also able to interact with cytochrom P450, acting as inhibitors and even substrates of some isoforms, as described in the results. As it is well known that CYP450 drug metabolism can induce clinical effects, these properties should be carefully analysed in order to design lead compounds from the herein studied molecules [57]. Moreover, toxicology predictors show that the studied molecules tend to have bad toxic effects, except meridianins A and E, for which no toxicity was predicted and the maximum tolerated dose increases with respect to the rest of the studied compounds.

Together, the obtained results suggest the necessity of performing a H2L optimization, in order to improve the absorption, distribution, metabolism and toxicity of the studied compounds, as well as their selectivity, with the aim of obtaining lead compounds able to become effective anti-AD drugs.

Materials and Methods

Virtual Profiling

VP techniques are computational tools aimed to elucidate the biological profile of a given molecule, for instance, therapeutic indications or targets of a chemical compound could be estimated. These techniques can be ligand- or target-based. Ligand-based approaches are able to automatically evaluate very large libraries or databases of compounds containing diverse information, for example, compound–target–bioactivities associations, using a chemical structure as a seed. As a result, similar molecules (restricted by a cut-off) are found and for instance, plausible targets to the input molecule selected. In this study, meridianin A was used as a seed. To run LBDD experiments, Cabrakan and Hurakan (Mind the Byte SL, Barcelona, Spain) software tools were employed [58,59]. Cabrakan is a two-dimensional (2D) ligand-based VP tool that compares molecules, through the use of 2D fingerprints, over a reference database and the assignment of biological activity. It allows the identification of similar chemical compounds (analogues) to the input molecule. Hurakan is a three-dimensional (3D) VP tool that compares a query molecule with the structures present in a reference database using Comparative Molecular Similarity Indices Analysis (CoMSIA) fields on a 3D grid. Hurakan can compare molecules according to their relationship with their environment, thus obtaining biomimetic compounds with different chemical structures. ChEMBL, which contains around 1,300,000 chemical compounds with detailed information including target data, was employed as the reference database [60]. A target was counted once when it appeared as both 2D and 3D hit during ligand-based VP experiments.

Here, we have employed similarity search based techniques, as they are simple, fast and accurate. However, they have the limitation imposed by the reference database employed. If there are no similar molecules to the input compound in the database, no results will be returned. This limitation is shared with other LBDD techniques such as quantitative structure–activity relationship (QSAR) or quantitative structure–property relationship (QSPR). The choice of these software tools and not another ones is based basically on the deep knowledge we have about the algorithm, the database and their performance.

Target-based approaches are able to, through knowledge of the 3D structures, evaluate huge databases that contain cavity information of these structures and after a binding site identification, docking calculations can be performed. As a result, the binding energy of every possible interaction is returned, which allows the classification and prediction of the best targets. In this study, meridianin A was used as a seed. Ixchel (Mind the Byte SL, Barcelona, Spain) is a structure-based VP tool that performs docking calculations of a molecule (spatial data file (SDF) or simplified molecular input line entry specification (SMILE) file) against an in-house developed database comprising almost 9000 protein cavities (binding-sites) curated from Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) according to UniProt Knowledgebase (UniProtKB) human entries [61–63].

To run target (or virtual) profiling experiments related to SBDD, docking is the most used technique. MD simulations or related techniques could be also employed, but they are much too computationally expensive for these kinds of techniques, with docking the preferred option. There are several variants of the docking techniques, but as we have commented for LBDD, the main limitation is the reference database. In our case, we have selected a technique whose algorithm is well known and it also incorporates a curated database of which we have a deep understanding. A deep knowledge of the employed techniques is basic and based on that, we have selected Ixchel to run our experiments.

Structure Modelling

The meridianin structures were modelled from the 2D chemical structure published by Núñez-Pons, Avila and co-workers [26]. The three meridianins derived compounds used for the selectivity test were modelled from Giraud and co-workers and Bharate et al. [25,32].

Prior to any calculation, all the structures of the selected targets, for the binding and the selectivity analysis, were modelled from their crystal structures available from the Protein Data Bank (RCSB PDB). All of them represent human targets. As obtaining good structures is crucial, the best 3D structures were selected; the structures and chains that cover the maximum amino acid region sequence, in general, and the binding region of each of the selected targets in particular.

GSK3 β was modelled from the crystallographic 3D structure with a PDB ID 3PUP that contains the crystallographic ligand OS1. It is stored in the PDB database as a homodimer, but only chain B was considered for further studies since GSK3 β biological assembly is in monomeric form [31]. CK1 δ was modelled from the 3D crystallographic structure corresponding to the entry 4KBK that contains the crystallographic ligand 1QG. Only chain B, since it is naturally a monomer, was considered for further studies [64]. DYRK1A was modelled from the crystal 3D structure with a PDB ID 4AZE that contains the

crystallographic ligand 3RA. In the PDB database, we found 3 chains (A, B and C), but only chain A was considered for further studies as DYRK1A biological assembly is in a monomeric form [52]. CLK1 was modelled from the crystallographic 3D structure with a PDB code 2VAG with V25 as a crystallographic ligand. As this protein is naturally a monomer, there is only one chain in the PDB database, so further studies were performed against chain A [52].

To test selectivity, for all the PDB crystallographic structures selected, chain A was used in all cases. Structures were modelled from their respective crystallographic 3D structure: Fibroblast growth factor receptor 1 (FGFR1); 1AGW containing SU2 as a ligand, cAMP-dependent protein kinase catalytic subunit alpha (PRKACA); 2GU8 containing 796 as a ligand, hexokinase-2 (HK2); 2NZT containing BG6 as a ligand, dual specificity mitogen-activated protein kinase 1 (MAP2K1); 3DY7 containing ATP as a ligand, phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit gamma isoform (PIK3CG); 3IBE containing L64 as a ligand, PRKG1; 3OCP containing CMP as a ligand, serine/threonine-protein kinase N1 (PKN1); 4OTI containing MI1 as a ligand and one non protein-kinase IDH1; 4I3K containing NDP as a ligand.

To test the binding of meridianins and their selectivity, molecular modelling experiments were performed using the 3D structural models of meridianins A–G, and the models generated from the crystallographic structures available in the PDB (PDB ID 3PUP, 4KBK, 4AZE and 2VAG, respectively) and the PDB ID structures 1AGW, 2GU8, 2NZT, 3DY7, 3IBE, 3OCP, 4OTI and 4I3K, respectively.

Docking Calculations

Docking calculations constitute a simulation method, which predicts the preferred orientation of one molecule (ligand) to a second (target). When only the movements of the first molecule are allowed, the docking is considered classical or rigid; when both molecules are allowed to move, docking is considered flexible. Generally, docking, without any other specification, refers to classical (rigid) docking [7]. Docking, in the context of small-molecule drug discovery, concerns the study of binding process of small molecules (ligands) and targets (proteins), i.e., a candidate binding mode (pose) is predicted when ligand and receptor bind to each other. Scoring functions allow us to classify and rank, based on their calculated binding energies, the most favourable pose. In that sense, flexible docking has advantages over the rigid version of the technique. The dynamics is an intrinsic characteristic of proteins, necessary to carry out any of their functions. Flexibility incorporation within the binding mode prediction is key to obtain results capable of being correlated with experimental data. However, not all are advantages, as the predicted binding energies could worsen. The inclusion of additional degrees of freedom to simulate

protein flexibility could increase the difficulty of accurately predicting the free energy of binding. This complication could arise because more contributions to the free energy must be considered, for instance, the interaction between flexible residues and the core of the protein, and typically, these additional contributions also introduce additional inaccuracies [65].

Another option to add flexibility is the post-processing of docking results, which means, for instance, docking validation and/or refinement by MD simulations. Rigid docking can predict the optimal placement of a ligand within the binding site of a receptor, but not all the key interactions between the ligand and receptor are usually depicted accurately. Hence, MD simulations can optimize the predicted binding mode and also check the stability of the docked complex, as a bad docking pose will generate an unstable MD trajectory, during which the ligand could even leave the binding site [34,36]. In this study, we have employed a pipeline aimed to simulate a flexible docking protocol in a similar way to other studies reported in the literature, in that we post-processed the obtained docking poses [66]. We selected this approximation as this two-step protocol constitutes a (probably the most) practical and convenient approach to address the docking problem [67]. It is in general less computationally expensive and provides the results that we need in an accurate way, comparable to “real” flexible docking methodologies (such as ensemble-based or flexible induced-fit docking). In general, using MD as a post-processing tool, a smaller fraction of the conformational space is usually covered, but without the several limitations that affect sampling and scoring algorithms for docking.

All docking calculations were performed using Itzamna and Kin software tools (Mind the Byte SL, Barcelona, Spain) [68,69] to perform classical and blind docking calculations, respectively. Itzamna is used to carry out docking calculations and needs the structure of the molecule to dock, as well as the cavity where it should be placed as an input. Kin is a software tool designed to perform blind docking calculations. It involves a cavity search and a (best) cavity selection prior to performing the binding calculation; a difference of Itzamna is that the docking cavity is given as an input to the calculation. When the employed crystal structures were co-crystallized with a ligand, the cavity defined by the ligand was employed. As mentioned above, the modelled structures of the meridianins and the selected targets were used. Two runs were carried out for each calculation to avoid false positives.

Results obtained from docking calculations were ranked based on their calculated binding affinities, and the best poses summarized in Tables A1 and A8.

Molecular Dynamics Simulations

One of the principal tools in the computational studies of biomolecules are MD simulations, a theoretical method for studying the physical movements of atoms and molecules. MD calculates the time dependent behaviour of a molecular system, which means that atoms and molecules are allowed to interact for a fix period of time, giving a view of the dynamic evolution of the system.

Short (1 nanosecond (ns)) MD simulations were performed using NAMD program version 2.11 over the best-docked complexes, which were selected based on ΔG_{bind} [70]. The Amber ff99SB-ILDN and the General Amber Force Field (GAFF) set of parameters were employed for modelling receptors and ligands, respectively [71,72]. The election of these force-fields was based on the fact that both have been extensively tested, being two of the most used for protein and protein-ligand simulations [71–74]. It has been shown that ff99SB-ILDN correlates consistently well with experimental data, and the GAFF force-field can conveniently and quickly produce reasonable ligand (especially organic molecules) parameters. Moreover, as amber force-fields, both are compatible, giving combined satisfactory results in several studies. Ligand GAFF parameters were obtained using Antechamber, whereas the receptor structures were modelled using the leap module of Amber Tools [75,76]. Simulations were carried out in explicit solvent using the TIP3P water model with the imposition of periodic boundary conditions via a cubic box [77]. Electrostatic interactions were calculated by the particle-mesh Ewald method using constant pressure and temperature conditions. Each complex was solvated with a minimum distance of 10 Å from the surface of the complex to the edge of the box. Temperature was kept at 300 Kelvin (K) using a Langevin Piston barostat. The time step employed was 2 femtoseconds (fs). Bond lengths to hydrogens were constrained with the SHAKE algorithm [78]. Before production runs, the system was energy minimized. Next, the solvent surrounding the protein was equilibrated at the target temperature using harmonic position restraints on the heavy atoms. Finally, the system was submitted to a slow heating-up phase, from 0 to 300 K. For the production run, all position restraints were removed.

Molecular Mechanics/Generalized Born Surface Area (MM/GBSA)

The so-called reweighting techniques are computational approaches to estimate the alchemical free energy of interaction (ΔG_{bind}) between small ligands and biological macromolecules. In the literature, MM/GBSA is usually employed to estimate ligand-binding affinities based on docking or MD simulations to get a more realistic view of the interaction of docked complexes. The obtained energies are more realistic than the docking interaction values, allowing a better ranking of the analysed

compounds, although they cannot be biologically comparable. In our case and following similar approaches, we applied reweighting techniques, specifically MM/GBSA, over the generated MD trajectories for post-processing docking results [34,66,79].

MM/GBSA rescoring was performed using the MMPBSA python algorithm contained within the Amber Tools suite [80]. The snapshots generated in the 1 ns MD simulation were imputed into the post-simulation MM/GBSA calculation of binding free energy. MM/GBSA was chosen over other techniques such as molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA), linear interaction energy (LIE), thermodynamics integration (TI) or free energy perturbation (FEP) because of its good balance between accuracy and computational cost.

Rigorous thermodynamic pathway approaches, such as TI or FEP, provide more accurate predicting binding free energies, whereas LIE, MM/GBSA and MM/PBSA constitute the so-called end-point methods that in general are less accurate. Each of these methods has its own strengths and limitations, and their computational requirements and speed are inversely correlated with their accuracy. TI and FEP, which outperform end-point approaches, are very useful, especially for ranking molecules inside a chemical series. Consequently, and regardless of their computational cost but given the computational advances, these techniques are gradually being more frequently used in the drug discovery pipeline, especially in guiding lead optimisation. However, in this study, our aim is not to provide a detailed library of lead compounds, and thus we have employed a less rigorous, but very popular approach in SBDD, alternative as the MMGBSA approach. The main problem of these techniques could be that the efficacy of the method is usually system dependent. However, it is generally accepted that they outperform docking results, so a better ranking of the analysed compounds will be always obtained, although, as commented above, the obtained binding energies could be far from being experimentally comparable.

Interaction Analysis

To analyse the key residues of the active site involved in the inhibitor binding, we examined the obtained binding modes after molecular modelling studies with already known binders of each of the targets. These binders (residues that have been revealed as necessary for the binding of known substrates/inhibitors) were identified through an evidence-based interaction analysis. It was carried out through a bibliographical search plus a database analysis. The bibliographical search was conducted using several studies in which inhibitors against the selected kinases were identified describing each compound binding mode [25,31,32,50–52,64,81–83]. The database search was done using an in-house,

recently constructed database. It was built by crossing ChEMBL and the RCSB PDB [62], and it contains all PDB structures per UniProtKB ID with active compounds (by now there are only PDBs with compounds not competing against cofactors). Moreover, the database also contains the residues to which each active compound (per PDB) is bound. Thus, it allows the user, after docking or an MD calculation, to easily check whether the analysed molecules behave as a binder.

Sequence Analysis

The four targets were aligned using the UniProtKB clustal omega interface from the amino acid sequence associated with each UniProtKB entry.

Selectivity Analysis

Docking calculations of meridianins, as well as the three selected compounds (derived from them and described in the literature), against twelve protein kinases were performed. These meridianins derived molecules were obtained from the papers of Bharate et al. and Giraud et al. [25,32], and have shown interesting inhibitory concentration (IC50) values in the micro and sub-micromolar range, and a good selectivity for DYRK1A and CLK1. We selected them to see how the selectivity was taken into account in the design of these compounds as they strongly resemble the original meridianins scaffolds that we suspect are not selective enough. To test the selectivity, we choose seven protein kinases, specifically, FGFR1, PRKACA, HK2, MAP2K1, PIK3CG, PRKG1 (for which the selected crystal structure do not contain the catalytic hinge), PKN1 and one non-protein kinase, IDH1. Thus, we tested if the selected compounds are selective between different protein kinases, belonging to different subfamilies, and between protein and non-protein kinases. Moreover, we explored if without the catalytic hinge, binding could be produced.

ADMET Properties Prediction

For the meridianins and the three derived compounds, ADMET properties prediction was carried out using proprietary machine-learning (ML) models and the pkCSM webserver [84,85]. The proprietary ML models covered logS (molecular aqueous coefficient), logP (octanol/water partition coefficient), Caco2 permeability, BBB penetration and PPB. The first two models were generated by super vector regression (SVR) techniques and the last three employed supper vector machines (SVM). For training and testing the models, Chembl (logS, logP, Caco2) and Huuskonen (logS) datasets were employed, and

for BBB and PPB, the datasets described by Muehlbacher et al. and Zhu and coworkers [86–88]. The pkCSM webserver allows the prediction of PK properties based on (I) compound general properties (including molecular properties, toxicophores and pharmacophore) and (II) distance-based graph signatures. Given an input molecule, both sources of information are used to train and test machine learning-based predictors. The webserver is composed of 28 (not all employed in this work) regression and classification ML models that have been generated and trained against 30 datasets (described at Pires et al.) [85].

The use of proprietary models, some of which are also covered by pkCSM, is because these methods, similar to other such as VS or VP, strongly rely on the employed reference dataset. As we have a deeper knowledge of our methods, we prefer to use them when possible. Only for Caco2 did we employ both models, ours and the pkCSM model, because for two compounds, our model is not good enough to make a reliable prediction (they are out of the applicability domain as they are too different with respect to the molecular fragments contained in the dataset employed to generate and train the model. If less than 90% of the molecular fragments in that the input molecule can be decomposed are not in the database, the prediction is not done). pkCSM predicted properties for all the compounds; however, it does not indicate if a prediction is out of the applicability domain.

In summary, we have analysed 21 ADMET properties, 5 of which were studied with our proprietary ML models and 17 with pkCSM. One of these properties, Caco2, was analysed twice using both our proprietary model and the pkCSM model.

Absorption properties

Caco2 permeability, LogS, intestinal absorption (human), P-glycoprotein substrate, P-glycoprotein I/II inhibitor and skin permeability. Caco-2 permeability is used to predict the absorption of orally administered drugs. A high permeability is assessed when the predicted value is >0.90 for the pkCSM model, or high (H), in the proprietary model. LogS reflects the solubility of the molecule in water at 25 °C and also reflects the bioavailability of a given compound; it is represented by the logarithm of the molar concentration (log mol/L). Intestinal absorption indicates the portion of compounds absorbed through the human intestine; a molecule with an absorbance (intestinal absorption) of less than 30% is considered to be poorly absorbed. Pgp acts as a biological barrier by extruding toxins and xenobiotics out of cells, although it could have other, transport mediated, functions in certain tissues and organs. The predictor assesses whether a given compound is likely to be a substrate of Pgp. Pgp I and II inhibitors have significant PK implications for Pgp substrate, and the predictor will determine the inhibitory effect of a given compound against Pgp I/II, which could have advantages that can be

exploited therapeutically, or result in contraindications. Skin permeability predicts if a given compound is likely to be skin permeable ($\log K_p > -2.5$).

Distribution

LogP, VD_{ss}, PPB, BBB and CNS permeability. LogP allows us to estimate the distribution of a drug within the body (lipophilicity). VD_{ss}, which is the theoretical volume that the total dose of a drug would need to be uniformly distributed to give the same concentration as in blood and plasma, is considered low if $\log VD_{ss} < -0.15$ and high if > 0.45 (the higher the VD, the greater the drug distribution in tissue rather than plasma). PPB estimates the probability (>90% is considered high) that a given molecule binds to a plasma protein, the less bound a drug is, the more efficiently it can traverse cell membranes or diffuse. BBB permeability describes the ability of a drug to cross into the brain. The predictor describes whether a compound is able to cross the BBB. CNS permeability measures blood brain permeability surface-area ($\log PS$), and it is similar to BBB but more direct, as it lacks the systemic distribution effects that may distort brain penetration. Compounds with a $\log PS > -2$ are considered to penetrate CNS, while those with $\log PS < -3$ are considered unable to penetrate.

Metabolism

Cytochrom P450 (CYP450) isoforms are important detoxification enzymes in the body and are essential for the metabolism of many medications. Drugs can be inhibitors of CYP450, blocking its metabolic activity, or can be metabolised (substrate) by them. CYP metabolism predictor assess whether a given molecule is likely to be metabolised or not and act as inhibitor of specific isoforms of CYP450; a specific inhibitor of CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4 and/or substrate of CYP2D6 and CYP3A4.

Excretion

Renal OCT2 substrate and Total Clearance. OCT2 is a renal uptake transporter that plays an important role in disposition and renal clearance of drugs and endogenous compounds. The OCT2 substrate predictor indicates if a given molecule is likely to be an OCT2 substrate, which provides not only clearance-related information but potential contraindications. Total clearance is related to bioavailability and is also important for determining dosing rates to achieve steady-state concentrations, and the predictor measures their value in $\log(\text{mL}/\text{min}/\text{kg})$.

Toxicology

MRTD, AMES toxicity, hepatotoxicity, skin sensitization, hERG I/II inhibitors. MRTD provides an estimated of the toxic dose threshold of chemicals in humans, and results less than or equal to 0.477

log(mg/kg/day) are considered low, and high when greater than 0.477 log(mg/kg/day). AMES toxicity indicates if a compound could be mutagenic and therefore may act as a carcinogen. hERG I and II inhibitor predictors determine if a given compound is likely to be a hERG I/II inhibitor as the inhibition of potassium channels encoded by hERG could result in fatal pathologies (for instance it is the principal cause of the development of acquiring long QT syndrome, fatal arrhythmia) and the withdrawal of many substances from the pharmaceutical market. Hepatotoxicity predicts if a given molecule is likely to be associated with disrupted normal function of the liver. Skin permeability predicts if a given compound is likely to be associated with skin sensitisation.

Graphical Representations

Graphical representations of protein-ligand complexes were prepared using PyMOL version 1.7 [89] and PLIP version 1.3.0 [90].

Conclusions

Meridianins can be classified as kinase inhibitors and can be used as a starting point to design and develop novel anti-AD drugs. It has been demonstrated, *in silico* and *in vitro*, that they are able to bind specific tau (GSK3 β and CK1 δ) and dual-specificity (DYRK1A and CLK1) protein kinases. However, they are not selective enough to constitute a therapeutic treatment against AD by themselves. In fact, as they are demonstrated to be protein kinase inhibitors, they could probably inhibit several kinases involved in different diseases[91]. In any case, they could serve as a starting scaffold to design new anti-AD drugs. To achieve that, a rational design taking advantage of the differences found in the binding patterns against different protein-kinases subfamilies, has to be carried out. In that sense, the presence of Br on R2 and the R3 position over the meridianin indole scaffold could be synonymous with potency. Besides, it seems that exploiting the C-terminal region (sugar and phosphate pocket) rather than the N-terminal side, could increase the strength of the interactions exerted by meridianins, and probably the potency shown by the designed compounds. However, although potency is important, and maintaining the presence of Br seems to be fairly accomplished [25,32], the selectivity between protein-kinase subfamilies is a crucial point to design proper anti-AD drugs, and even anti-cancer drugs. Meridianins are not selective enough and should be improved to gain functionality and applicability. In addition, their measured ADMET properties indicate that they should be optimized in order to become a drug or at least a drug-lead compound. Therefore, the above-mentioned rational design in order to improve the potency and selectivity of meridianins should include H2L optimization cycles. The showed toxicity

should be removed, and compounds interaction with Cytochrom P450 carefully analysed and, given the case, eliminated or modulated. Moreover, their distribution properties should be improved, lowering the PPB and VDss, to be able to diffuse and penetrate into cells easily. Besides, a mechanism to cross the BBB should be found and in that sense, modifying each compound to be Pgp inhibitors could be a possible strategy, although there are other mechanisms to overcome the BBB, including other protein binding and nanodelivery, that could be also exploited [92–94].

Regarding meridianins specifically and CADD methods in general, we can conclude that these techniques, despite their drawbacks, are very helpful in drug discovery, constituting a powerful tool that could save time and money in experiments. Our study with meridianins is an example of this, since we have been able to find plausible targets, that in the case of AD and cancer we have already validated through the literature. The key role that these techniques could have in drug discovery is even higher for the discovery and development of marine drugs, since no sample is needed to run these virtual experiments. Moreover, since these methods could point out the best direction to follow and in which targets expand the low sample amount that usually is available, these are crucial technologies to maximize the success of marine prospection, as well as to protect biodiversity.

Acknowledgments: This research was partially supported by an Industrial Doctorate grant from the Generalitat of Catalonia to L.L.-P (DI 2016-051).

Author Contributions: C.A. and M.S.-M. conceived the study and designed the experiments. L.L.-P. carried out the experiments whereas A.N.-C., C.A. and M.S.-M. supervised them. All the authors analyzed and discussed the results as well as wrote the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

Table A1. Summary of classical rigid docking and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the two best models selected per meridianins A–G. All energy values are kcal/mol.

GSK3β		CK1δ		DYRK1A		CLK1					
Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)				
R0/R1		R0/R1		R0/R1		R0/R1					
A	-7.3/-7.3 -6.6/-6.1	-26.43 -24.95	A	-6.9/-6.9 -6.8/-6.8	-32.25 -30.10	A	-7.4/-7.3 -7.5/-7.4	-28.00 -31.43	A	-8.9/-8.9 -7.8/-7.8	-27.49 -30.70
B	-7.3/-7.2 -6.8/-6.7	-29.11 -29.25	B	-6.4/-6.4 -5.6/-5.5	-35.06 -32.30	B	-7.7/-6.9 -7.3/-7.9	-37.38 -34.03	B	-8.5/-8.5 -8.0/-8.0	-34.14 -30.38
C	-7.6/-7.6 -7.4/-7.5	-28.54 -31.44	C	-6.9/-6.9 -6.9/-6.7	-38.85 -35.84	C	-8.2/-8.2 -7.6/-7.6	-31.95 -35.90	C	-8.5/-8.5 -8.1/-8.1	-33.31 -34.92
D	-7.7/-7.7 -7.0/-6.9	-31.19 -30.01	D	-7.0/-7.0 -6.8/-6.6	-38.69 -38.06	D	-7.9/-7.9 -7.5/-7.6	-29.47 -34.59	D	-8.6/-8.6 -8.1/-8.1	-33.58 -35.90
E	-7.3/-7.3 -7.5/-7.5	-31.26 -28.43	E	-7.0/-7.0 -7.0/-7.0	-35.20 -34.97	E	-7.5/-7.4 -7.6/-7.4	-35.62 -32.55	E	-9.0/-8.8 -7.9/-7.9	-26.39 -31.63
F	-7.9/-7.9 -7.7/-7.9	-35.18 -34.73	F	-7.2/-7.3 -7.1/-7.1	-38.55 -38.93	F	-8.0/-7.8 -7.8/-7.7	-39.99 -39.91	F	-8.7/-8.7 -8.5/-8.5	-37.71 -37.61
G	-7.3/-7.3 -7.2/-7.2	-24.04 -27.03	G	-6.8/-6.8 -6.9/-6.9	-31.92 -32.94	G	-8.1/-8.1 -8.1/-8.1	-30.17 -30.52	G	-9.1/-9.1 -8.7/-8.7	-27.95 -29.88

To avoid false positives, each docking calculation was performed twice (R0 and R1).

Table A2. GSK3 β residues that interacted with meridianins (each represented by letters A–G) after molecular dynamics (MD) simulations. Those residues involved in all meridianin binding are bold and were considered important binding residues.

Residues	A	B	C	D	E	F	G
I62	X	X	X	X	X	X	X
F67		X		X	X		X
V70	X	X	X	X	X	X	X
A83	X	X	X	X	X	X	X
K85	X	X	X	X	X	X	X
E97	X			X	X	X	X
V110	X	X	X		X	X	X
L132	X	X	X	X	X	X	X
D133	X			X		X	X
Y134	X	X	X	X		X	
V135	X	X	X		X		X
T138		X	X		X		
Q185				X			
L188	X	X	X	X	X	X	
D200	X	X	X	X	X	X	X

Table A3. CK1 δ residues that interacted with meridianins (each represented by letters A–G) after MD simulations. Those residues involved in all meridianin binding are in bold and were considered important binding residues.

Residues	A	B	C	D	E	F	G
I15	X	X	X	X	X		
I23	X	X	X	X	X	X	X
Y24		X					
A36	X	X	X	X	X	X	X
K38	X	X	X	X	X	X	X
E52		X	X		X		
Y56	X			X		X	X
M80				X		X	X
M82	X	X	X	X	X	X	X
E83	X	X			X		
L84	X	X		X			
L85	X	X	X	X	X	X	X
L135	X	X	X	X	X	X	X
I148	X	X	X	X	X	X	X
D149	X	X	X	X	X		X

Table A4. DYRK1A residues that interacted with meridianins (each represented by letters A–G) after MD simulations. Those residues involved in all meridianin binding are in bold and were considered important binding residues.

Residues	A	B	C	D	E	F	G
I165	X	X		X	X	X	X
K167				X			
F170	X		X	X	X	X	
V173	X	X	X	X	X	X	X
A186	X	X	X		X	X	X
K188	X	X	X	X	X	X	X
E203							X
V222	X	X	X	X	X	X	X
F238	X	X	X	X	X	X	X
E239		X			X	X	X
L241	X	X	X	X	X	X	X
N244				X			
N292					X		
L294	X	X	X	X	X	X	X
V306	X	X	X	X	X	X	X
D307	X	X	X	X	X	X	X

Table A5. CLK1 residues that interacted with meridianins (each represented by letters A–G) after MD simulations. Those residues involved in all meridianin binding are in bold and were considered important binding residues.

Residues	A	B	C	D	E	F	G
L167	X	X	X	X	X	X	X
F172	X	X	X		X	X	
V175	X	X	X	X	X	X	X
A189	X	X	X	X	X	X	X
K191	X	X	X	X	X		X
E206			X				X
L210			X				
V225	X	X	X	X		X	X
F241	X	X	X	X		X	X
E242			X	X	X	X	
L243	X		X	X			
L244	X	X	X	X	X	X	X
G245					X		
S247					X		
N293							X
L295	X	X	X	X	X	X	X
V324	X	X	X	X	X		X
D325	X	X	X	X	X		X

Table A6. Binder columns represent those residues identified after a bibliographic and database research and that interacted with other inhibitors. In shared columns are those residues involved with all meridianins binding per target. Residue number corresponds to each Protein Data Bank (PDB) number.

GSK3 β		CK1 δ		DYRK1A		CLK1	
Binders	Shared	Binders	Shared	Binders	Shared	Binders	Shared
I62	I62					L167	L167
V70	V70	I23	I23	V173	V173	V175	V175
A83	A83	A36	A36			A189	A189
K85	K85	K38	K38	K188	K188	K191	
				V222	V222	V225	
L132	L132	M82	M82	F238	F238	F241	
V135		L85	L85	L241	L241	L244	L244
L188		L135	L135	L294	L294	L295	L295
		I148	I148	V306	V306	V324	
D200	D200	D149		D307	D307		

Table A7. Residues involved in all meridianins binding to GSK3 β , CK1 δ , DYRK1A and CLK1. Residue number corresponds to each PDB number.

GSK3β	CK1δ	DYRK1A	CLK1
I62	I15	I165 K167	L167
F67		F170	F172
V70	I23 Y24	V173	V175
A83	A36	A186	A189
K85	K38	K188	K191
E97	E52 Y56 M80	E203	E206 L210
V110		V222	V225
L132	M82	F238	F241
D133	E83	E239	E242
Y134	L84		L243
V135	L85	L241	L244
T138		N244	G245
Q185		N292	S247 N293
L188	L135 I148	L294 V306	L295 V324
D200	D149	D307	D325

Table A8. Summary of classical rigid docking of the best model selected per meridianin A–G and the derived compounds 1, 2 and 3, against others protein kinases and one non-kinase (IDH1).

	GSK3β	CK1δ	DYRK1A	CLK1	FGFR1	PRKACA
	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)
	R0/R1	R0/R1	R0/R1	R0/R1	R0/R1	R0/R1
A	-7.3/-7.3	A -6.9/-6.9	A -7.4/-7.3	A -8.9/-8.9	A -7.1/-7.1	A -8.5/-8.5
B	-7.3/-7.2	B -6.4/-6.4	B -7.7/-6.9	B -8.5/-8.5	B -6.7/-6.7	B -8.5/-8.5
C	-7.6/-7.6	C -6.9/-6.9	C -8.2/-8.2	C -8.5/-8.5	C -7.1/-7.1	C -9.0/-9.0
D	-7.7/-7.7	D -7.0/-7.0	D -7.9/-7.9	D -8.6/-8.6	D -7.1/-7.1	D -8.8/-8.8
E	-7.3/-7.3	E -7.0/-7.0	E -7.5/-7.4	E -9.0/-8.8	E -7.4/-7.4	E -7.6/-7.6
F	-7.9/-7.9	F -7.2/-7.3	F -8.0/-7.8	F -8.7/-8.7	F -7.3/-7.3	F -8.5/-8.5
G	-7.3/-7.3	G -6.8/-6.8	G -8.1/-8.1	G -9.1/-9.1	G -7.1/-7.1	G -8.6/-8.6
1	-7.6/-7.6	1 -7.0/-7.0	1 -8.1/-8.1	1 -9.2/-9.2	1 -7.1/-7.1	1 -8.2/-8.2
2	-7.7/-7.7	2 -7.1/-7.1	2 -8.2/-8.2	2 -7.4/-7.4	2 -6.2/-6.2	2 -8.3/-8.3
3	-8.0/-8.0	3 -7.3/-7.3	3 -7.9/-7.9	3 -7.8/-7.8	3 -6.4/-6.4	3 -8.0/-8.0
	HK2	MAP2K1	PIK3CG	PRKG1	IDH1	PKN1
	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)
	R0/R1	R0/R1	R0/R1	R0/R1	R0/R1	R0/R1
A	-7.1/-7.1	A -7.6/-7.6	A -6.8/-6.8	A -6.3/-6.3	A -5.8/-5.8	A -7.8/-7.8
B	-6.6/-6.6	B -7.4/-7.4	B -7.3/-7.3	B -6.5/-6.5	B -6.3/-6.3	B -7.7/-7.7
C	-7.0/-7.0	C -7.2/-7.2	C -7.0/-7.0	C -6.6/-6.6	C -5.8/-5.8	C -8.1/-8.1
D	-6.6/-6.6	D -7.4/-7.4	D -7.8/-7.8	D -6.9/-6.9	D -5.8/-5.8	D -7.3/-7.3
E	-6.8/-6.8	E -6.7/-6.7	E -7.1/-7.1	E -6.3/-6.3	E -5.6/-5.6	E -7.2/-7.2
F	-6.9/-6.9	F -7.5/-7.5	F -7.2/-7.2	F -6.8/-6.8	F -6.0/-6.0	F -7.8/-7.8
G	-6.9/-6.9	G -7.5/-7.5	G -7.3/-7.3	G -6.5/-6.5	G -6.3/-6.3	G -7.9/-7.9
1	-6.9/-6.9	1 -7.4/-7.4	1 -7.7/-7.7	1 -6.4/-6.4	1 -5.8/-5.8	1 -7.4/-7.4
2	-8.1/-8.1	2 -7.5/-7.5	2 -7.3/-7.3	2 -5.3/-5.3	2 -6.1/-6.1	2 -7.7/-7.7
3	-7.1/-7.1	3 -7.2/-7.2	3 -7.3/-7.3	3 -6.1/-6.1	3 -6.0/-6.0	3 -7.7/-7.7

All energy values are kcal/mol. To avoid false positives, each docking calculation was performed twice (R0 and R1).

Table A9. Summary of ADMET properties of meridianins (A to G) and the derived compounds extracted from the literature (1–3).

	LogS		Caco2 Permeability		Caco2 * Permeability		Intestinal Absorption		Skin Permeability		
	ABSORPTION	A	-4.18	A	H	A	0.99	A	93.38%	A	-2.76
B		-5.02	B	H	B	1.07	B	92.22%	B	-2.76	
C		-5.55	C	H	C	0.95	C	91.77%	C	-2.92	
D		-5.55	D	H	D	0.95	D	92.715	D	-2.91	
E		-5.04	E	H	E	0.98	E	90.98%	E	-2.74	
F		-6.16	F	H	F	0.98	F	91.49%	F	-2.92	
G		-4.51	G	H	G	0.86	G	93.44%	G	-2.90	
1		-4.18	1	OAD	1	0.93	1	91.41%	1	-2.90	
2		-5.02	2	OAD	2	0.8	2	89.89%	2	-2.884	
3		-5.55	3	H	3	0.819	3	91.04%	3	-2.895	
P-Glycoprotein substrate		P-Glycoprotein I/II Inhibitor		LogP		BBB		PPB			
A	Yes	A	No	DISTRIBUTION	A	1.53	A	No	A	>90%	
B	Yes	B	No		B	2.39	B	No	B	>90%	
C	Yes	C	No		C	3.10	C	No	C	>90%	
D	Yes	D	No		D	3.10	D	No	D	>90%	
E	Yes	E	No		E	2.40	E	No	E	>90%	
F	Yes	F	No		F	3.58	F	No	F	>90%	
G	Yes	G	No		G	2.44	G	No	G	<50%	
1	Yes	1	No		1	3.40	1	No	1	>90%	
2	Yes	2	No		2	3.40	2	No	2	>90%	
3	Yes	3	No		3	3.10	3	No	3	>90%	
VDss		CNS Permeability		CYP450 Metabolism *		Total Clearance		Renal OCT2 Substrate			
A	0.25	A	-2.92	METABOLISM	A	Yes	EXCRETION	A	0.57	A	No
B	0.24	B	-2.92		B	Yes		B	0.30	B	No
C	-0.06	C	-2.81		C	Yes		C	0.09	C	No
D	-0.01	D	-2.82		D	Yes		D	0.14	D	No
E	0.22	E	-2.93		E	Yes		E	0.15	E	No
F	0.07	F	-2.82		F	Yes		F	-0.19	F	No
G	-0.10	G	-2.12		G	Yes		G	0.71	G	No
1	-0.02	1	-2.83		1	Yes		1	-0.07	1	No
2	-0.09	2	-1.88		2	Yes		2	-0.092	2	No
3	-0.09	3	-1.99		3	Yes		3	0.132	3	No

Caco2 permeability is calculated using proprietary ML model and Caco2 * with the pkCSM webserver, as explained in the methods section. CYP450 metabolism * specific values of interaction with different CYP450 isoforms are listed in Table A11. BBB: blood brain Barrier, PPB: protein-protein binding, VDss: steady state volume of distribution, CNS: central nervous system, OCT2: organic cation transported 2.

Table A10. Summary of toxicity properties of meridianins A–G and the three derived compounds extracted from the literatures (1–3).

	MRTD	AMES Toxicity		hERG I/II Inhibition		Hepatotoxicity		Skin Sensitisation	
A	0.503	A	No	A	No	A	No	A	No
B	0.584	B	No	B	No	B	Yes	B	No
C	-0.107	C	Yes	C	No	C	No	C	No
D	-0.095	D	Yes	D	No	D	No	D	No
E	0.589	E	No	E	No	E	No	E	No
F	-0.088	F	Yes	F	No	F	Yes	F	No
G	-0.086	G	Yes	G	No	G	No	G	No
1	-0.068	1	Yes	1	No	1	No	1	No
2	-0.038	2	Yes	2	No	2	No	2	No
3	-0.058	3	Yes	3	No	3	No	3	No

MRTD: maximum recommended tolerated dose, hERG: human ether-a-go-go gene.

Table A11. Summary of specific values of interaction with different CYP450 isoforms properties of meridianins A–G and the three derived compounds (1–3).

CYP2D6		CYP3A4		CYP1A2		CYP2C19		CYP2C9		CYP2D6		CYP3A4	
Substrate		Substrate		Inhibitor		Inhibitor		Inhibitor		Inhibitor		Inhibitor	
A	No	A	No	A	Yes	A	No	A	No	A	No	A	No
B	No	B	No	B	Yes	B	No	B	No	B	No	B	No
C	No	C	No	C	Yes	C	No	C	No	C	No	C	No
D	No	D	No	D	Yes	D	No	D	No	D	No	D	No
E	No	E	No	E	Yes	E	No	E	No	E	No	E	No
F	No	F	No	F	Yes	F	Yes	F	No	F	No	F	No
G	No	G	No	G	Yes	G	No	G	No	G	No	G	No
1	No	1	No	1	Yes	1	Yes	1	No	1	No	1	Yes
2	No	2	Yes	2	Yes	2	Yes	2	No	2	No	2	No
3	No	3	No	3	Yes	3	No	3	No	3	No	3	Yes

CYP: Cytochrome

References

1. Paul, S.M.; Mytelka, D.S.; Dunwiddie, C.T.; Persinger, C.C.; Munos, B.H.; Lindborg, S.R.; Schacht, A.L. How to improve R&D productivity: The pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* **2010**, *9*, 203.
2. Leelananda, S.P.; Lindert, S. Computational methods in drug discovery. *Beilstein J. Org. Chem.* **2016**, *12*, 2694–2718.
3. Ou-Yang, S.-S.; Lu, J.-Y.; Kong, X.-Q.; Liang, Z.-J.; Luo, C.; Jiang, H. Computational drug discovery. *Acta Pharmacol. Sin.* **2012**, *33*, 1131–1140.
4. Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E.W. Computational Methods in Drug Discovery. *Pharmacol. Rev.* **2014**, *66*, 334–395.
5. Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular docking: A powerful approach for structure-based drug discovery. *Curr. Comput. Aided Drug Des.* **2011**, *7*, 146–157.
6. Durrant, J.D.; McCammon, J.A. Molecular dynamics simulations and drug discovery. *BMC Biol.* **2011**, *9*, 71.
7. Yu, W.; Mackerell, A.D. Computer-Aided Drug Design Methods. *Methods Mol. Biol.* **2017**, 1520, 93–94.
8. Acharya, C.; Coop, A.; Polli, J.E.; Mackerell, A.D., Jr. Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr. Comput. Aided Drug Des.* **2011**, *7*, 10–22.
9. Li, J.W.-H.; Vederas, J.C. Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science* **2009**, *325*, 161–165.
10. Harvey, A.L.; Edrada-Ebel, R.; Quinn, R.J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* **2015**, *14*, 111–129.
11. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **2016**, *79*, 629–661.
12. Prachayasittikul, V.; Worachartcheewan, A.; Shoombuatong, W.; Songtawe, N.; Simeon, S.; Prachayasittikul, V.; Nantasenamat, C. Computer-Aided Drug Design of Bioactive Natural Products. *Curr. Top. Med. Chem.* **2015**, *15*, 1780–1800.
13. Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on natural products for drug design. *Nat. Chem.* **2016**, *8*, 531–541.
14. Martin, L.; Latypova, X.; Wilson, C.M.; Magnaudeix, A.; Perrin, M.L.; Yardin, C.; Terro, F. Tau protein kinases: Involvement in Alzheimer's disease. *Ageing Res. Rev.* **2013**, *12*, 289–309.
15. Kolarova, M.; García-Sierra, F.; Bartos, A.; Ricny, J.; Ripova, D.; Ripova, D. Structure and Pathology of Tau Protein in Alzheimer Disease. *Int. J. Alzheimers Dis.* **2012**, *2012*, 1–13.
16. Citron, M. Alzheimer's disease: strategies for disease modification. *Nat. Rev. Drug Discov.* **2010**, *9*, 387–398.
17. Tell, V.; Hilgeroth, A. Recent developments of protein kinase inhibitors as potential AD

- therapeutics. *Front. Cell. Neurosci.* **2013**, *7*, 189.
18. Dolan, P.J.; Johnson, G.V.W. The role of tau kinases in Alzheimer's disease. *Curr. Opin. Drug Discov. Dev.* **2010**, *13*, 595–603.
 19. Lucke-Wold, B.P.; Turner, R.C.; Logsdon, A.F.; Simpkins, J.W.; Alkon, D.L.; Smith, K.E.; Chen, Y.-W.; Tan, Z.; Huber, J.D.; Rosen, C.L. Common mechanisms of Alzheimer's disease and ischemic stroke: The role of protein kinase C in the progression of age-related neurodegeneration. *J. Alzheimers Dis.* **2015**, *43*, 711–724.
 20. Montaser, R.; Luesch, H. Marine natural products: A new wave of drugs? *Future Med. Chem.* **2011**, *3*, 1475–1489.
 21. Blunt, J.W.; Copp, B.R.; Keyzers, R.A.; Munro, M.H.G.; Prinsep, M.R. Marine natural products. *Nat. Prod. Rep.* **2017**, *34*, 235–294.
 22. Kiuru, P.; D'Auria, M.; Muller, C.; Tammela, P.; Vuorela, H.; Yli-Kauhala, J. Exploring Marine Resources for Bioactive Compounds. *Planta Med.* **2014**, *80*, 1234–1246.
 23. Molinski, T.F.; Dalisay, D.S.; Lievens, S.L.; Saludes, J.P. Drug development from marine natural products. *Nat. Rev. Drug Discov.* **2009**, *8*, 69–85.
 24. The Nagoya Protocol on Access and Benefit-Sharing. Available online: <https://www.cbd.int/abs/about/default.shtml/> (accessed on 16 May 2017).
 25. Bharate, S.B.; Yadav, R.R.; Battula, S.; Vishwakarma, R.A. Meridianins: Marine-Derived Potent Kinase Inhibitors. *Mini-Rev. Med. Chem.* **2012**, *12*, 618–631.
 26. Núñez-Pons, L.; Carbone, M.; Vázquez, J.; Rodríguez, J.; Nieto, R.M.; Varela, M.M.; Gavagnin, M.; Avila, C. Natural Products from Antarctic Colonial Ascidians of the Genera *Aplidium* and *Synoicum*: Variability and Defensive Role. *Mar. Drugs* **2012**, *10*, 1741–1764.
 27. Traxler, P.; Furet, P. Strategies toward the Design of Novel and Selective Protein Tyrosine Kinase Inhibitors. *Pharmacol. Ther.* **1999**, *82*, 195–206.
 28. Huang, D.; Zhou, T.; Lafleur, K.; Nevado, C.; Caflisch, A. Kinase selectivity potential for inhibitors targeting the ATP binding site: A network analysis. *Bioinformatics* **2010**, *26*, 198–204.
 29. McGregor, M.J. A Pharmacophore Map of Small Molecule Protein Kinase Inhibitors. *J. Chem. Inf. Model.* **2007**, *47*, 2374–2382.
 30. Ebrahim, H.; El Sayed, K. Discovery of Novel Antiangiogenic Marine Natural Product Scaffolds. *Mar. Drugs* **2016**, *14*, 57.
 31. Feng, L.; Geisselbrecht, Y.; Blanck, S.; Wilbuer, A.; Atilla-Gokcumen, G.E.; Filippakopoulos, P.; Kräling, K.; Celik, M.A.; Harms, K.; Maksimoska, J.; et al. Structurally sophisticated octahedral metal complexes as highly selective protein kinase inhibitors. *J. Am. Chem. Soc.* **2011**, *133*, 5976–5986.
 32. Giraud, F.; Alves, G.; Debiton, E.; Nauton, L.; Théry, V.; Durieu, E.; Ferandin, Y.; Lozach, O.; Meijer, L.; Anizon, F.; et al. Synthesis, Protein Kinase Inhibitory Potencies, and in vitro

- Antiproliferative Activities of Meridianin Derivatives. *J. Med. Chem.* **2011**, *54*, 4474–4489.
33. Wang, G.; Zhu, W. Molecular docking for drug discovery and development: A widely used approach but far from perfect. *Future Med. Chem.* **2016**, *8*, 1707–1710.
 34. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. The Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* **2016**, *59*, 4035–4061.
 35. De Vivo, M.; Cavalli, A. Recent advances in dynamic docking for drug discovery. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1320.
 36. Aravindhan, G.; Coote, M.L.; Barakat, K. Molecular dynamics-driven drug discovery: Leaping forward with confidence. *Drug Discov. Today* **2017**, *22*, 249–269.
 37. Pagadala, N.S.; Syed, K.; Tuszynski, J. Software for molecular docking: A review. *Biophys. Rev.* **2017**, *9*, 91–102.
 38. Shin, W.-H.; Zhu, X.; Bures, M.; Kihara, D. Three-Dimensional Compound Comparison Methods and Their Application in Drug Discovery. *Molecules* **2015**, *20*, 12841–12862.
 39. Muegge, I.; Mukherjee, P. An overview of molecular fingerprint similarity search in virtual screening. *Expert Opin. Drug Discov.* **2016**, *11*, 137–148.
 40. Alonso, N.; Caamaño, O.; Romero-Duran, F.J.; Luan, F.; Cordeiro, M.N.D.S.; Yañez, M.; González-Díaz, H.; García-Mera, X. Model for High-Throughput Screening of Multitarget Drugs in Chemical Neurosciences: Synthesis, Assay, and Theoretic Study of Rasagiline Carbamates. *ACS Chem. Neurosci.* **2013**, *4*, 1393–1403.
 41. Simpraga, S.; Alvarez-Jimenez, R.; Mansvelter, H.D.; Van Gerven, J.M.A.; Groeneveld, G.J.; Poil, S.-S.; Linkenkaer-Hansen, K. EEG machine learning for accurate detection of cholinergic intervention and Alzheimer's disease. *Sci, Rep.* **2017**, *7*, 5775.
 42. Cummings, J.; Aisen, P.S.; DuBois, B.; Frölich, L.; Jack, C.R.; Jones, R.W.; Morris, J.C.; Raskin, J.; Dowsett, S.A.; Scheltens, P. Drug development in Alzheimer's disease: The path to 2025. *Alzheimers Res. Ther.* **2016**, *8*, 39.
 43. Solomon, K.A.; Sundararajan, S.; Abirami, V. QSAR Studies on N-aryl Derivative Activity Towards Alzheimer's Disease. *Molecules* **2009**, *14*, 1448–1455.
 44. Gopi Mohan, C.; Gupta, S. QSAR Models towards Cholinesterase Inhibitors for the Treatment of Alzheimer's Disease. In *Oncology: Breakthroughs in Research and Practice*; IGI Global: Hershey, PA, USA, 2016.
 45. PharmaSea. Available online: <http://www.pharma-sea.eu/> (accessed on 2 September 2017).
 46. MaCuMBA. Available online: <https://www.macumbaproject.eu/> (accessed on 2 September 2017).
 47. SeaBiotech. Available online: <http://spider.science.strath.ac.uk/seabiotech/index.php> (accessed on 2 September 2017).
 48. BlueGenics. Available online: <http://www.bluegenics.eu/cms/> (accessed on 2 September 2017).
 49. MicroB3. Available online: <https://www.microb3.eu/> (accessed on 2 September 2017).
 50. Jain, P.; Karthikeyan, C.; Moorthy, N.S.; Waiker, D.; Jain, A.; Trivedi, P. Human CDC2-Like Kinase 1

- (CLK1): A Novel Target for Alzheimer's Disease. *Curr. Drug Targets* **2014**, *15*, 539–550.
51. Yadav, R.R.; Sharma, S.; Joshi, P.; Wani, A.; Vishwakarma, R.A.; Kumar, A.; Bharate, S.B. Meridianin derivatives as potent Dyrk1A inhibitors and neuroprotective agents. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 2948–2952.
 52. Tahtouh, T.; Elkins, J.M.; Filippakopoulos, P.; Soundararajan, M.; Burgy, G.; Durieu, E.; Cochet, C.; Schmid, R.S.; Lo, D.C.; Delhommel, F.; et al. Selectivity, Cocrystal Structures, and Neuroprotective Properties of Leucettines, a Family of Protein Kinase Inhibitors Derived from the Marine Sponge Alkaloid Leucettamine B. *J. Med. Chem.* **2012**, *55*, 9312–9330.
 53. Wang, W.; Bodles-Brakhop, A.M.; Barger, S.W. A Role for P-Glycoprotein in Clearance of Alzheimer Amyloid β -Peptide from the Brain. *Curr. Alzheimer Res.* **2016**, *13*, 615–620.
 54. Cirrito, J.R.; Deane, R.; Fagan, A.M.; Spinner, M.L.; Parsadanian, M.; Finn, M.B.; Jiang, H.; Prior, J.L.; Sagare, A.; Bales, K.R.; et al. P-glycoprotein deficiency at the blood-brain barrier increases amyloid- β deposition in an Alzheimer disease mouse model. *J. Clin. Investig.* **2005**, *115*, 3285–3290.
 55. Miller, D.S.; Bauer, B.; Hartz, A.M.S. Modulation of P-glycoprotein at the blood-brain barrier: Opportunities to improve central nervous system pharmacotherapy. *Pharmacol. Rev.* **2008**, *60*, 196–209.
 56. Chang, K.L.; Pee, H.N.; Yang, S.; Ho, P.C. Influence of drug transporters and stereoselectivity on the brain penetration of pioglitazone as a potential medicine against Alzheimer's disease. *Sci. Rep.* **2015**, *5*, 9000.
 57. Lynch, T.; Price, A. The Effect of Cytochrome P450 Metabolism on Drug Response, Interactions, and Adverse Effects. *Am. Fam. Physician* **2007**, *76*, 391–396.
 58. Santamaria-Navarro, E.; Felix, E.; Nonell-Canals, A. Cabrakan. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
 59. Santamaria-Navarro, E.; Nonell-Canals, A. Hurakan. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
 60. Bento, A.P.; Gaulton, A.; Hersey, A.; Bellis, L.J.; Chambers, J.; Davies, M.; Krüger, F.A.; Light, Y.; Mak, L.; McGlinchey, S.; et al. The ChEMBL bioactivity database: An update. *Nucleic Acids Res.* **2014**, *42*, 1083–1090.
 61. Felix, E.; Santamaria-Navarro, E.; Sanchez-Martinez, M.; Nonell-Canals, A. Ixchel. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
 62. Berman, H.; Henrick, K.; Nakamura, H.; Markley, J.L. The worldwide Protein Data Bank (wwPDB): Ensuring a single, uniform archive of PDB data. *Nucleic Acids Res.* **2007**, *35*, 2006–2008.
 63. The UniProt Consortium Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res.* **2013**, *41*, 43–47.
 64. Mente, S.; Arnold, E.; Butler, T.; Chakrapani, S.; Chandrasekaran, R.; Cherry, K.; Dirico, K.; Doran, A.; Fisher, K.; Galatsis, P.; et al. Ligand-protein interactions of selective casein kinase 1 δ inhibitors. *J. Med. Chem.* **2013**, *56*, 6819–6828.

65. Lill, M.A. Efficient Incorporation of Protein Flexibility and Dynamics into Molecular Docking Simulations. *Biochemistry* **2011**, 50, 6157–6169.
66. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, 441, 1–13.
67. Alonso, H.; Bliznyuk, A.A.; Gready, J.E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* **2006**, 26, 531–568.
68. Felix, E.; Santamaría-Navarro, E.; Sanchez-Martinez, M.; Nonell-Canals, A. Itzamna. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
69. Felix, E.; Nonell-Canals, A. Kin. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
70. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, 26, 1781–802.
71. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, 78, 1950–1958.
72. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, 25, 1157–1174.
73. Martín-García, F.; Papaleo, E.; Gomez-Puertas, P.; Boomsma, W.; Lindorff-Larsen, K. Comparing Molecular Dynamics Force Fields in the Essential Subspace. *PLoS ONE* **2015**, 10, e0121114.
74. Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M.P.; Dror, R.O.; Shaw, D.E. Systematic validation of protein force fields against experimental data. *PLoS ONE* **2012**, 7, 1–6.
75. Wang, J.; Wang, W.; Kollman, P.A.; Case, D.A. Antechamber, An Accessory Software Package For Molecular Mechanical Calculations. *J. Comput. Chem.* **2005**, 25, 1157–1174.
76. Case, D.A.; Cheatham, T.E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, 26, 1668–1688.
77. Jorgensen, W.L.; Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *J. Comput. Chem.* **1998**, 19, 1179–1186.
78. Andersen, H.C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983**, 52, 24–34.
79. Rastelli, G.; Degliesposti, G.; Del Rio, A.; Sgobba, M. Binding Estimation after Refinement, a New Automated Procedure for the Refinement and Rescoring of Docked Ligands in Virtual Screening. *Chem. Biol. Drug Des.* **2009**, 73, 283–286.
80. Miller, B.R., III; McGee, T.D., Jr.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, 8, 3314–3321.

81. Mente, S.; Arnold, E.; Butler, T.; Chakrapani, S.; Chandrasekaran, R.; Cherry, K.; Dirico, K.; Doran, A.; Fisher, K.; Galatsis, P.; et al. Ligand-protein interactions of selective casein kinase 1 δ inhibitors. *J. Med. Chem.* **2013**, *56*, 6819–6828.
82. Halekotte, J.; Witt, L.; Ianes, C.; Krüger, M.; Bührmann, M.; Rauh, D.; Pichlo, C.; Brunstein, E.; Luxenburger, A.; Baumann, U.; et al. Optimized 4,5-Diarylimidazoles as Potent/Selective Inhibitors of Protein Kinase CK1 δ and Their Structural Relation to p38 α MAPK. *Molecules* **2017**, *22*, 522.
83. Fedorov, O.; Huber, K.; Eisenreich, A.; Filippakopoulos, P.; King, O.; Bullock, A.N.; Szklarczyk, D.; Jensen, L.J.; Fabbro, D.; Trappe, J.; et al. Specific CLK Inhibitors from a Novel Chemotype for Regulation of Alternative Splicing. *Chem. Biol.* **2011**, *18*, 67–76.
84. Vidal, D.; Nonell-Canals, A. ADMET Models. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
85. Pires, D.E.V.; Blundell, T.L.; Ascher, D.B. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J. Med. Chem.* **2015**, *58*, 4066–4072.
86. Muehlbacher, M.; Spitzer, G.M.; Liedl, K.R.; Kornhuber, J. Qualitative prediction of blood–brain barrier permeability on a large and refined dataset. *J. Comput. Aided Mol. Des.* **2011**, *25*, 1095–1106.
87. Zhu, X.-W.; Sedykh, A.; Zhu, H.; Liu, S.-S.; Tropsha, A. The Use of Pseudo-Equilibrium Constant Affords Improved QSAR Models of Human Plasma Protein Binding. *Pharm. Res.* **2013**, *30*, 1790–1798.
88. Huuskonen, J. Estimation of Aqueous Solubility for a Diverse Set of Organic Compounds Based on Molecular Topology. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 773–777.
89. Yuan, S.; Chan, H.C.S.; Hu, Z. Using PyMOL as a platform for computational drug design. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1298.
90. Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: Fully automated protein–ligand interaction profiler. *Nucleic Acids Res.* **2015**, *43*, W443–W447.
91. Prudhomme, M.; Rossignol, E.; Youssef, A.; Anizon, F.; Moreau, P.; Fabbro, D.; Cohen, P. Aminopyrimidylindoles structurally related to meridianins as kinase inhibitors. *Cancer Res.* **2014**, *68*.
92. Kingwell, K. Drug delivery: New targets for drug delivery across the BBB. *Nat. Rev. Drug Discov.* **2016**, *15*, 84–85.
93. Banks, W.A. From blood–brain barrier to blood–brain interface: New opportunities for CNS drug delivery. *Nat. Rev. Drug Discov.* **2016**, *15*, 275–292.
94. Saraiva, C.; Praça, C.; Ferreira, R.; Santos, T.; Ferreira, L.; Bernardino, L. Nanoparticle-mediated brain drug delivery: Overcoming blood–brain barrier to treat neurodegenerative diseases. *J. Control. Release* **2016**, *235*, 34–47.

Section II

Chapter 3

Kororamides, convolutamines, and indole derivatives as possible tau and dual-specificity kinase inhibitors for Alzheimer's disease: A computational study

Chapter 3

Kororamides, convolutamines, and indole derivatives as possible tau and dual specificity kinases inhibitors for Alzheimer's Disease.

Laura Llorach-Pares ^{1,2}, Alfons Nonell-Canals ², Conxita Avila^{1,*}, and Melchor Sánchez-Martínez^{2,*}

1. Department of Evolutionary Biology, Ecology and Environmental Sciences, Faculty of Biology and Biodiversity Research Institute (IRBio), Universitat de Barcelona, 08028 Barcelona, Catalonia.
2. Mind the Byte S.L., 08028 Barcelona, Catalonia, Spain.

Published in: *Marine Drugs* 16(10), 386 (2018) doi: 10.3390/md16100386

Abstract

Alzheimer's disease (AD) is becoming one of the most disturbing health and socioeconomic problems nowadays, as it is a neurodegenerative pathology with no treatment expected to grow further due to population ageing. Current treatments for AD produce only a modest amelioration of symptoms, although there is a constant ongoing research of new therapeutic strategies oriented to improve them, and even to completely cure the disease. A principal feature of AD is the presence of neurofibrillary tangles (NFT) induced by the aberrant phosphorylation of the microtubule-associated protein tau in the brains of affected individuals. Glycogen synthase kinase-3 beta (GSK3 β), casein kinase 1 delta (CK1 δ), dual specificity tyrosine phosphorylation regulated kinase 1A (DYRK1A), and dual specificity kinases cdc2-like kinase 1 (CLK1) have been identified as the principal proteins involved in this process. Because of that, the inhibition of these kinases has been proposed as a plausible therapeutic strategy to fight AD. In this study, we computationally tested the inhibitory activity of different marine natural compounds, as well as newly designed molecules from their scaffolds, over the mentioned protein kinases, finding some new possible inhibitors with potential therapeutical application.

Keywords: Meridianins; Kororamide A-B; Convolutamine I-J; Indole scaffold; Computer-Aided Drug Design; Alzheimer's disease; GSK3 β ; CK1 δ ; DYRK1A; CLK1.

Resum

La malaltia d'Alzheimer (MA) s'està convertint en una de les malalties més inquietants i en un problema socioeconòmic en l'actualitat, ja que és una patologia neurodegenerativa que no té tractament, i s'espera que la seva afectació augmenti encara més a causa de l'envelliment de la població. Els tractaments actuals per a la MA només produeixen una modesta millora dels símptomes, tot i que hi ha una constant i permanent investigació de noves estratègies terapèutiques orientades a millorar aquests símptomes, i fins i tot, per curar completament la malaltia. Una característica principal de la MA és la presència de cabdells neurofibril·lars, induïts per una aberrant fosforilació de la proteïna tau associada als microtúbuls, que es troben presents en el cervell dels individus afectats. La glycogen synthase kinase-3 beta (GSK3 β) i la casein kinase 1 delta (CK1 δ , CK1D o KC1D), així com les quinases de doble especificitat, com la dual specificity tyrosine phosphorylation regulated kinase 1 (DYRK1A) i les quinases cdc2-like (CLK1), s'han identificat com les principals proteïnes implicades en aquest procés d'hiperfosforilació. Per això, la inhibició d'aquestes quinases s'ha proposat com una estratègia terapèutica plausible per combatre la MA. En aquest estudi, hem estudiat computacionalment l'activitat inhibidora de diferents compostos naturals d'origen marí, així com molècules dissenyades a partir dels seus esquelets, sobre les esmentades proteïnes quinases, trobant alguns nous possibles inhibidors amb potencial aplicació terapèutica.

Paraules clau: Meridianines; kororamide A-B; convolutamine I-J; esquelets indòlics; disseny de fàrmacs assistit per ordinador; malaltia d'Alzheimer; GSK3 β ; CK1 δ ; DYRK1A; CLK1.

Introduction

Constituting about 2% of all human genes, protein kinases are an important family of enzymes with a critical role in signal transduction pathway by modification of substrate activity. They are also responsible to control different aspects of cell functions by its phosphorylation activity, which plays a critical role in intracellular communication during development, and in the function of the nervous and immune systems [1]. Due to that, kinases are related with many diseases such as Alzheimer's Disease (AD) or Amyotrophic Lateral Sclerosis (ALS), among others. AD, the neurodegenerative pathology that is considered to represent the most common type of dementia (60-80% of the total cases), is characterized by memory deterioration and modification of cognitive abilities. Alzheimer's pathologies are associated with the presence of senile plaques (SP), mainly composed by Beta-Amyloid ($A\beta$) peptides, and neurofibrillary tangles (NFT), that are intraneuronal aggregations principally composed of abnormal phosphorylated tau protein. Tau is a soluble microtubule-binding protein and is hyperphosphorylated in AD. Tau phosphorylation is regulated by a balance between tau kinase and phosphatase activities. Anti-phosphorylation strategies (kinase inhibitors) aim to inhibit these processes of aggregation and the formation of NFT [2–4]. The above mentioned evidences may suggest that one of the key strategies to prevent tau phosphorylation and thus, combat AD, could be the inhibition of the protein kinases involved in the tau phosphorylation pathway [4].

Despite the catalytic subunits of many protein kinases are highly conserved, there are several differences between them that allow to classify protein kinases into subfamilies: 1) Protein Kinases (EC 2.7.10); 2) Serine-Threonine protein kinases (EC 2.7.11); 3) Dual-specificity kinases (those acting on Ser/Thr and Tyr residues) (EC 2.7.12); 4) Protein-histidine kinases (EC 2.7.13); 5) Protein-Arginine kinases (EC 2.7.11.14) and 6) other protein kinases (EC 2.7.99), that can be also divided into sub-subfamilies, such as tau protein kinase (EC 2.7.11.26) and dual specificity kinase (EC 2.7.12.1). The main relevant protein kinases involved in tau phosphorylation belong to the sub-subfamilies tau protein kinase and dual specificity kinases. glycogen synthase kinase-3 beta (GSK3 β) and casein kinase 1 delta (CK1 δ) are tau protein kinases, while dual specificity tyrosine phosphorylation regulated kinase 1A (DYRK1A) and cdc2-like kinase 1 (CLK1) are dual specificity kinases. Each of them has different roles regarding AD pathology. For GSK3 β several authors suggest its link between $A\beta$ and tau pathology, and in AD patients it has been co-localized with NFT. GSK3 β is suggested to phosphorylate and hyper-phosphorylate tau, while increasing the production of $A\beta$ and mediating neuronal death. Phosphorylation of tau by GSK3 β occurs at 42 sites, where 29 of them are phosphorylated in AD brains. CK1 δ is part of the non-proline-directed protein kinase (non-PDPK) group inside the tau kinases and its levels are increased while is co-

localized with NFT. CK1 δ has an important role on protein aggregation and regulates the microtubule dynamics through tau phosphorylation at 46 sites, 25 of them phosphorylated in AD brains. DYRK1A phosphorylates the amyloid precursor protein (APP) and tau proteins, thus increasing neuronal death and the formation of aggregates. DYRK1A induces tau phosphorylation at serine 202, threonine 212, and serine 404, sites that were found phosphorylated in AD brains. Finally, cdc2-like kinase 1 (CLK1), one of the four isoforms conforming an evolutionary conserved group of dual specificity kinases, is related with AD by phosphorylating the serine residues in arginine-rich (SR) proteins [2,3,5–15].

The natural-product-inspired design plays an important role in chemical science, as historically natural products (NP) from diverse sources, such as plants or microbes, have been a rich source of compounds [16–18]. NP are optimized biologically active metabolites which can be used as a template to design drug-like compounds [16–18]. Evaluation of Food and Drug Administration (FDA) approved new molecular entities (NMEs) reveals that NP and their derived compounds represent over one-third of all NMEs [19], a percentage that is even higher regarding the active compounds in the central nervous system (CNS) domain [20]. AD is not an exception, and several drug candidates have been developed from natural sources against the different therapeutic targets identified to date [21–23]. In fact, few reasonable selective and potent GSK3 β , CK1 δ , DYRK1A and CLK1 inhibitors have been described so far, being most of them marine natural products or derived molecules from them [8,24–36].

Recently, it has been shown that meridianins, indole alkaloids from the marine tunicate *Aplidium* from the Southern Ocean, could act as inhibitors of these four kinases, with possible inhibitors being derived from them [24,29,34,37]. In addition to that, kororamide A-B, two brominated alkaloids from the bryozoan *Amathia tortuosa* from Australia, showed a phenotypic signature on Parkinson's disease [38]. Their structure resembles that of meridianins and because of that we decided to study whether these compounds could also act as inhibitors of the four mentioned kinases, although as far as we know this relation has never been established before. Following with this, and having into account that marine indole alkaloid conform a large group of compounds with diverse biological activities that make them attractive starting points for pharmaceutical development [39–41], we have designed here several compounds starting from this well-known scaffold as a core element. Further, we modified the structural features observed in meridianins and kororamides, as well as with the presence of halogen substituents (present also in both chemical species), which has been revealed as key player to increase activity over these four kinases [24,37,42,43].

To beefing up our initial assumption, we tested the indole scaffold and halogen substituents effect on the inhibition of GSK3 β , CK1 δ , DYRK1A and CLK1. To determine the importance of the indole

scaffold for the inhibition of the four studied kinases we also screened the MarinLit database [44] to find other possible marine compounds that were similar to meridianin F and kororamide A (which were the best theoretical inhibitors of the four kinases) or at least incorporate the indole scaffold. Thereafter, we analysed their binding behaviour against them. Moreover, and because of the importance of the halogen substituents, we decided to investigate whether the halogen substituents are important respect to the indole scaffold. To do that, we evaluated the inhibitory behaviour of convolutamine I-J, two halogenated heterocyclic compounds (that do not present an indole scaffold) extracted also from the bryozoan *Amathia tortuosa* and which are structurally and functionally related to kororamide A-B [38].

To sum up, with the general objective to contribute to the discovery of anti-AD drugs (protein inhibitor/s to reduce or alleviate AD symptoms), the concrete aim of this study is threefold 1) to validate if kororamide A-B and convolutamine I-J could act as novel inhibitors of the four studied kinases, 2) to test the indole scaffold importance on the kinases inhibition, and 3) to design new possible inhibitors of the four kinases starting from meridianin and kororamide indole scaffolds. To do so, a computational study targeting the adenosine triphosphate (ATP) binding site of the aforementioned kinases has been carried out by using computer-aided drug design (CADD) methods. CADD techniques are widely used in (marine natural products) drug discovery, as they constitute an appropriate tool to rationally design and develop new drug candidates, reducing the time and costs derived from their identification, characterization and structure-optimization [45].

Results and discussion

New possible GSK3 β , CK1 δ , DYRK1A and CLK1 ATP competitive inhibitors

It is generally accepted that the ATP binding site of protein kinases, despite the fact that their catalytic domains are highly conserved, still remains the most used cavity in (rational) drug design over this family of proteins [46]. Protein kinases have two different lobes, the N-lobe that is mainly formed by β -sheets, and the C-lobe formed by a helical structure. Between both lobes, the catalytic ATP cavity is found, and it can be divided into five regions: Glycine-rich region (GRR), hydrophobic pocket (HP), adenine region (AR), sugar pocket (SP), and the phosphate binding pocket (PBP) [46–48]. GRR and HP are located at the N-terminal lobe, while SP and PBP are placed at the C-terminal lobe. AR is in the middle of these regions, providing a link between them (Figure 1).

All five regions are quite evolutionary conserved between the kinases, but they are not identical [37]. GRG is a highly conserved region with a GxGxFG motif (Table 1). The same occurs with the HP, as all

the four kinases have a VAIK motif, except DYRK1A with a Valine (V) residue instead of an Isoleucine (I). On the contrary, the AR does not seem to have any conserved motif, while SP can be identified by the PxNxL pattern. For the PBP, only the last Aspartate (D) residue seems to be conserved along the four kinases.

Table 1. Summary of the ATP binding site regions of GSK3 β , CK1 δ , DYRK1A and CLK1. Five regions found inside the ATP cavity and their respective residues in a single letter code, as well as their sequence position that corresponds to each PDB file numbering.

	Glycine-rich Region	Hydrophobic Pocket	Adenine Region	Sugar Pocket	Phosphate Binding Pocket
GSK3β	GNGSFG	VAIK	LDYV	PQNLL	LKLCD
	63-68	82-85	132-135	184-188	196-200
CK1δ	GSGSFG	VAIK	MELL	PDNFL	VYIID
	16-21	35-38	82-85	131-135	145-149
DYRK1A	GKGSFG	VAIK	FEML	PENIL	IKIVD
	166-171	184-187	238-241	290-294	303-307
CLK1	GEGAFG	VAVK	FELL	PENIL	IKVVD
	168-173	188-191	241-244	291-295	312-325

As explained previously, the kinases ATP binding site is the most exploited cavity as far as inhibition is concerned. Several inhibitors have been reported in the past, being some of them marine natural products, such as meridianins [28,49]. Most of them can bind to all these regions, with a different binding strength depending on their chemical structure. Interestingly, a common feature seems to be shared between the majority of them: the presence of an indole scaffold [8,25,26,30,31,33–35].

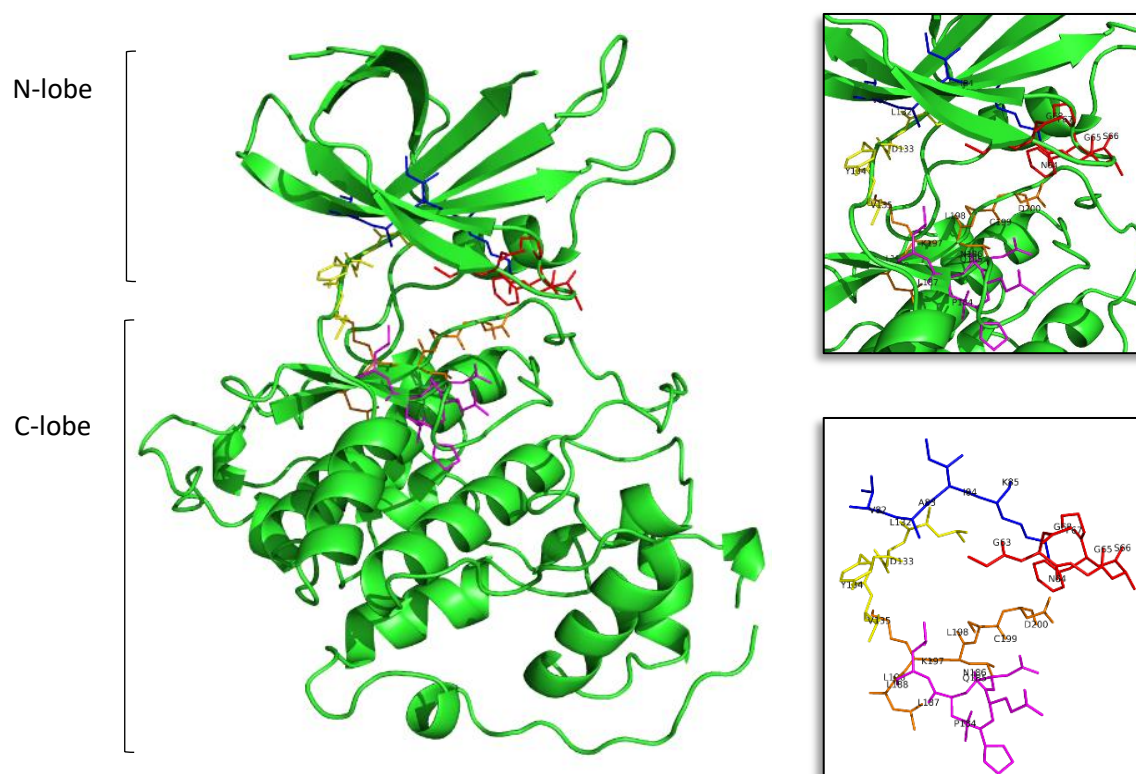


Figure 1. Structure of the tau protein kinase GSK3 β (Protein Data Bank ID (PDB) 3PUP). In the first, largest, image the two lobes can be seen in cartoon representation, and in sticks the residues that form the ATP cavity. In the top and bottom zoom images all the aminoacid residues involved on the ATP binding site are shown. Residues in red represent the glycine-rich region (GRR), in blue the hydrophobic pocket (HP), in yellow the adenine region (AR), in lilac the sugar pocket (SP), and finally, in orange the phosphate binding pocket (PBP). Letters and numbers correspond to their position in the aminoacid sequence and the PDB file numbering. Original from the author.

Kororamide A-B and Convolutamine I-J as possible kinase inhibitors

Indole alkaloids are marine natural products that show specific biological activities, such as anti-inflammatory and serotonin antagonism [41]. Moreover, the therapeutic importance of this kind of indole scaffolds is well known, as demonstrated by clinical and preclinical studies showing pharmacological activities over neurodegenerative diseases, such as AD [41,50]. Meridianins A-F belong to the group of compounds containing the indole moiety. These molecules constitute a group of indole alkaloids consisting of an indole framework linked to an aminopyrimidine ring with a reported inhibitory activity over GSK3 β , CK1 δ , DYRK1A and CLK1 [30,34,37]. Within the list of indole containing compounds, structurally similar to meridianins, different molecules can be found, among which are kororamides. Kororamide A and B are two tribrominated indole alkaloid compounds from the Australian bryozoan *Amanthia tortuosa*. These two marine molecules share a common halogenated indole scaffold with meridianins and, based on their chemical structural similarity, one could assume that kororamides could

have an inhibitory activity similar to meridianins. In the same study where kororamide B was identified, three other compounds were also isolated, kororamide A and convolutamine I and J. The last two compounds do not present an indole scaffold, but they are halogenated heterocyclic compounds as other known kinase inhibitors [51–53] (Figure 2). To test this hypothesis, docking calculations and Molecular Dynamics (MD) simulations were carried out to evaluate if kororamide A-B and convolutamine I-J could behave as meridianins regarding GSK3 β , CK1 δ , DYRK1A and CLK1 binding, thus indicating that they could be potential anti-AD therapeutic agents.

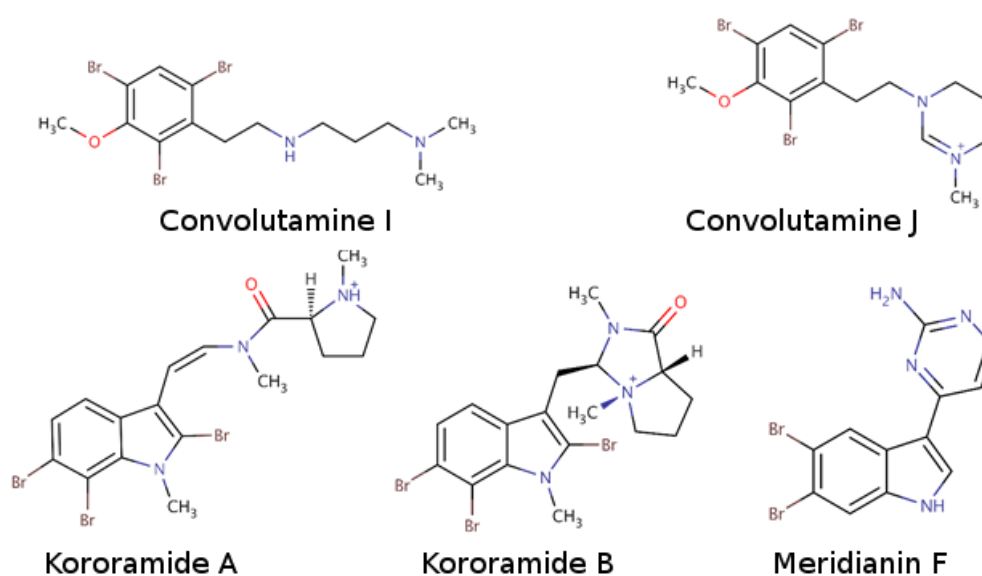


Figure 2. Chemical structures of convolutamine I, convolutamine J, kororamide A, kororamide B and meridianin F.

In previous studies the presence of halogen atoms was considered important to achieve a good inhibitory activity over the four studied kinases [24,37]. In order to test whether the presence of a halogenated indole scaffold, or just the presence of aromatic cycle substituted with halogen atoms enhances a higher binding affinity against GSK3 β , CK1 δ , DYRK1A and CLK1, we analyse it by means of docking and MD simulations. Thereafter, we compared the obtained results (Table 2) with the values from kororamide A-B and convolutamine I-J with meridianin F, the most promising compound of the chemical family (meridianin A-F) [37].

Table 2. Summary of classical rigid docking and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the two best models selected per meridianin F (F), convolutamine I (I) and J (J), kororamide A (A) and B (B).

GSK3 β		CK1 δ		DYRK1A		CLK1	
Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)
R0/R1		R0/R1		R0/R1		R0/R1	
F	-7.9/-7.9 -7.7/-7.9	F	-7.2/-7.3 -7.1/-7.1	F	-8.0/-7.8 -7.8/-7.7	F	-8.7/-8.7 -8.5/-8.5
I	-5.6/-5.6 -6.3/-6.3	I	-5.0/-5.0 -5.4/-5.4	I	-5.6/-5.6 -4.8/-4.8	I	-5.8/-5.5 -5.8/-5.8
J	-6.7/-6.7 -5.9/-5.9	J	-6.2/-6.2 -5.8/-5.8	J	-7.4/-7.4 -7.0/-7.0	J	-6.0/-6.0 -4.6/-4.6
A	-8.3/-8.3 -8.1/-8.1	A	-8.0/-8.0 -7.4/-7.4	A	-8.2/-8.2 -6.7/-6.7	A	-6.7/-6.7 -2.9/-2.9
B	-9.1/-9.1 -8.3/-8.3	B	-8.1/-8.1 -6.6/-6.6	B	-7.7/-7.7 -7.3/-7.3	B	-4.4/-4.4 -4.0/-4.0
	-35.18 -34.73		-38.55 -38.93		-39.99 -39.91		-37.71 -37.61
	-23.08 -18.38		-3.19 -11.26		-26.52 -11.02		-33.23 -31.93
	-31.58 -31.61		-37.76 -28.91		-31.35 -32.27		-21.47 -24.37
	-34.88 -31.02		-35.48 -33.94		-32.94 -14.61		-37.46 -38.93
	-31.80 -32.34		-28.68 -35.53		-23.83 -24.29		-28.71 -22.96

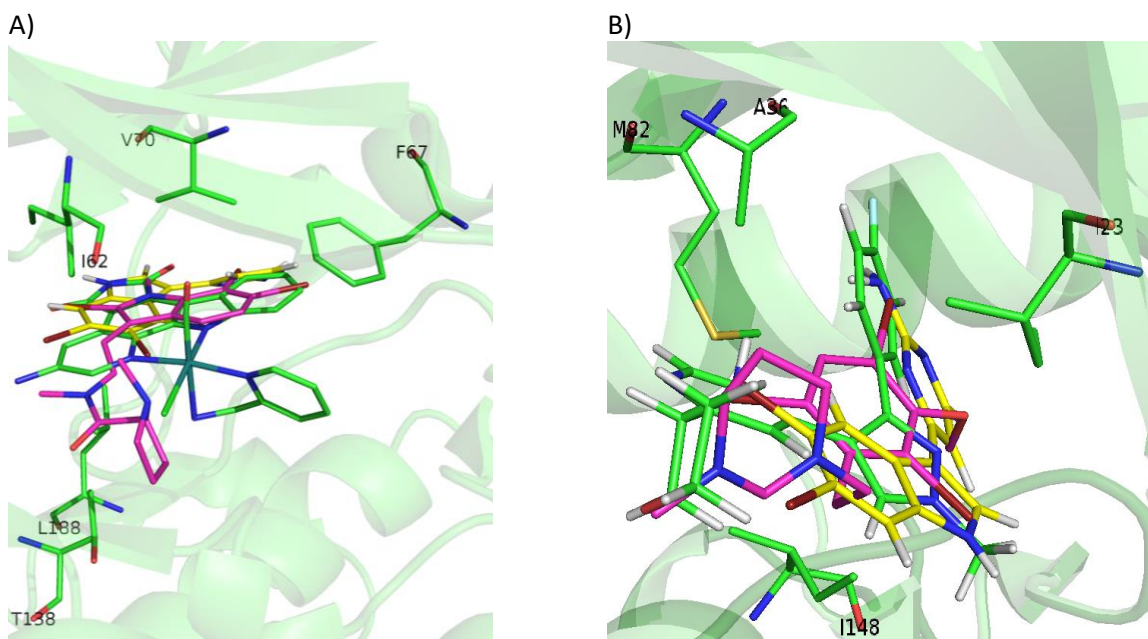
To avoid false positives, each docking calculation was performed twice (R0/R1). All energy values are kcal/mol. For each target the first (left) column refers to the results of docking calculations while the second (right) column indicate the binding energy results obtained after MD calculations.

Our results indicate that all the analysed compounds could bind to the ATP binding pocket of each of the mentioned kinases, thus theoretically acting as ATP competitive inhibitors (Figure 3). Binding energies obtained after docking and MD simulations (Table 2) show that convolutamine J and kororamide A tend to have higher energies than convolutamine I and kororamide B. To be more precise, kororamide A shows better energies when bound against GSK3 β , DYRK1A and CLK1, while convolutamine J shows better energies over CK1 δ . Comparing the energies obtained between the four tribrominated metabolites found on the bryozoa *Amathia tortuosa* and meridianin F, we observe that the last one has slightly better energies in all cases after MD. These energies do not allow us to discard any of the compounds as an ATP competitive inhibitor, although we can prioritize kororamide A and convolutamine J over kororamide B, and specially over convolutamine I. Besides, these results do not allow us to discriminate between which structural feature influences most the binding strength against the four studied kinases: the indole scaffold, the presence of halogen atoms, or the combination of both features.

With the aim of performing a deeper analysis of the inhibitory behaviour of these compounds, an interaction and binding mode analysis, of the best and prioritized compounds per target, was performed. On the ATP catalytic cavity of GSK3 β it is observed that key binders as I62, F67 and V70, conforming the GRR or placed nearby, and Y138 and L188 placed at C-terminal lobe near the AR and inside SP, respectively, are involved on the kororamide stabilization (Figure 3A). For CK1 δ it is observed that convolutamine J it is stabilized by interacting with several key binders as I23 that is placed near the

GRR and A36, M82 and I148 placed at HP, AR and PBP, respectively (Figure 3B). Looking at DYRK1A ATP cavity, it is observed that kororamide A, at the N-terminal region, is interacting with I165 and V173, as other known inhibitors like meridianin F or the co-crystal 3RA, both placed near the GRR. In the same way, kororamide A is also stabilized by A185, which is found at the HP, and at the C-terminal zone it is also interacting with E291 and L294, conforming PENIL motif, and the key binder V306 present at the PBP. Finally, kororamide A is also stabilized by L241 and D244, placed near the AR (Figure 3C). Looking at the ATP cavity of CLK1, it can be seen that on the N-terminal domine, L167, F172 and V175 at the GRR and K191 at the HP, some of them known key binders, are interacting with kororamide A. Moreover, on the C-lobe, kororamide A is interacting with F241 coming from the FELL adenin motif, E292 and L295 (a known key binders), placed at the SP, and V324 found at the PBP (Figure 3D).

The binding mode of the best derivatives, as well as of the four brominated compounds studied, per target pointed out that they are performing key contacts, according to our interaction analyses. This fact together with the obtained binding energies, reinforce their capacity to behave as inhibitors for the four analyzed kinases, in a similar way to meridianin F.



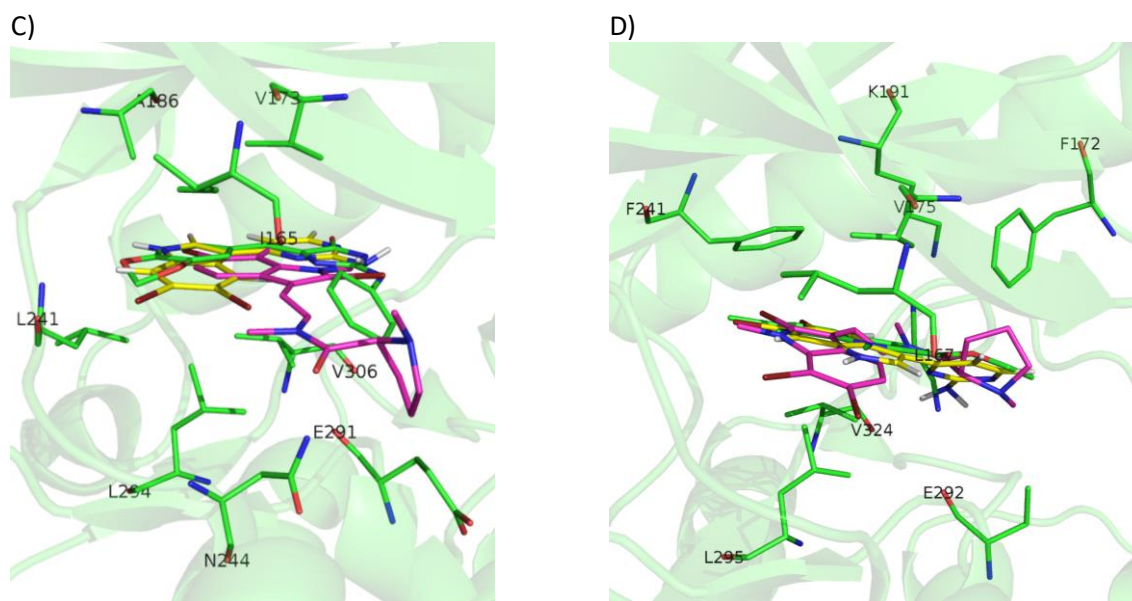


Figure 3. A) ATP cavity site of GSK3 β (Protein Data Bank ID (PDB) 3PUP) with meridianin F (yellow), the co-crystallized OS1 inhibitor (green), and the best pose of kororamide A (magenta). B) ATP cavity site of CK1 δ (Protein Data Bank ID (PDB) 4KBK) with meridianin F (yellow), the co-crystallized 1QG inhibitor (green), and the best pose of convolutamine J (magenta). C) ATP cavity site of DYRK1A (Protein Data Bank ID (PDB) 4AZE) with meridianin F (yellow), the co-crystallized 3RA inhibitor (green), and the best pose of kororamide A (magenta). D) ATP cavity site of CLK1 (Protein Data Bank ID (PDB) 2VAG) with meridianin F (yellow), the co-crystallized V25 inhibitor (green), and the best pose of kororamide A (magenta). Letters and numbers correspond to their position in the amino acid sequence and the PDB file numbering. Original from the author.

Marine natural products and indole scaffold validation

With the aim of testing the importance of the indole scaffold as structural key feature on the kinases ATP inhibitors and assuming the well-known Structure Activity Relationship (SAR) principle (i.e. structurally similar compounds will have similar biological activities) a substructure search was performed over the MarinLit database, a dataset that includes revised compounds from marine natural products [44]. In that sense, similar compounds to meridianin F and kororamide A and the indole scaffold were searched over this database. A list of 24 compounds was obtained, 18 compounds when the indole scaffold was used as a seed, and three using meridianin F and kororamide A, respectively. The list could contain more molecules if all the indole-containing compounds were selected. However, we decided that this number is adequate to test if the indole scaffold with several, mostly minor, additions is enough to have a theoretical inhibitory effect over the four kinases, or whether a complex structure like meridianin F or kororamide A is necessary. Docking calculations were performed to analyse the binding behaviour of all of them over GSK3 β , CK1 δ , DYRK1A, and CLK1 (Table 3).

Table 3. Summary of classical rigid docking calculations of the marine natural compounds found at MarinLit database after a substructure similarity search using an indole group, meridianin F and kororamide A as input molecules. To avoid false positives, each docking calculation was performed twice (R0/R1). All energy values are kcal/mol.

	GSK3β	CK1δ	DYRK1A	CLK1		GSK3β	CK1δ	DYRK1A	CLK1
	Binding Energy	Binding Energy	Binding Energy	Binding Energy		Binding Energy	Binding Energy	Binding Energy	Binding Energy
	R0/R1	R0/R1	R0/R1	R0/R1		R0/R1	R0/R1	R0/R1	R0/R1
L17640	-6.4 / -6.4	-7.3 / -7.3	-7 / -7	-6 / -6	L4950	-9.1 / -9.1	-9.1 / -9.1	-9.1 / -9.1	-9.1 / -9.1
L1189	-6.8 / -6.8	-7.6 / -7.6	-7.2 / -7.2	-5.9 / -5.9	L4949	-8.7 / -8.7	-8.7 / -8.7	-8.7 / -8.7	-8.7 / -8.7
L34	-7.2 / -7.2	-8.1 / -8.1	-8.2 / -8.2	-6.9 / -6.9	L4951	-9 / -9	-9 / -9	-9 / -9	-9 / -9
L4080	-6.1 / -6.1	-6.9 / -6.9	-6.8 / -6.8	-6 / -6					
L28238	-6.5 / -6.5	-7.8 / -7.8	-7.3 / -7.3	-5.8 / -5.8					
L7472	-6.3 / -6.3	-7.1 / -7.1	-6.8 / -6.8	-6.2 / -6.2					
L10723	-6.1 / -6.1	-6.7 / -6.7	-6.7 / -6.7	-5.2 / -5.2					
L17639	-6.4 / -6.4	-7.1 / -7.1	-6.8 / -6.8	-5.6 / -5.6					
L1192	-7.1 / -7.1	-7.6 / -7.6	-7.6 / -7.6	-5.9 / -5.9					
L17641	-6.8 / -6.8	-7 / -7	-7 / -7	-5.7 / -5.7					
	-6.2 / -6.2	-6.7 / -6.7	-6.6 / -6.6	-5.4 / -5.4					
L11375									
L35	-7.3 / -7.3	-8 / -8	-8.1 / -8.1	-7.1 / -7.1					
L28804	-7.1 / -7.1	-7.6 / -7.6	-7.2 / -7.2	-5.8 / -5.8					
L4081	-6.4 / -6.4	-7 / -7	-6.8 / -6.8	-6.4 / -6.4					
L29233	-8.5 / -8.5	-8.5 / -8.5	-9.3 / -9.3	-8.2 / -8.2					
L24201	-10.6 / -10.6	-8.8 / -8.8	-10.4 / -10.4	-9.3 / -9.3	L9830	-7.4 / -7.4	-7.4 / -7.4	-7.4 / -7.4	-7.4 / -7.4
L25368	-9.7 / -9.7	-8.8 / -8.8	-10.3 / -10.3	-8.9 / -8.9	L9831	-7.8 / -7.8	-7.8 / -7.8	-7.8 / -7.8	-7.8 / -7.8
L7473	-6.9 / -6.9	-7.3 / -7.3	-7.1 / -7.1	-6 / -6	L2330	-8.7 / -8.7	-8.7 / -8.7	-8.7 / -8.7	-8.7 / -8.7

The molecules names (Lxxxx) corresponds to the MarinLit entry code per each compound.

All those compounds with energies higher than -9.0 kcal/mol obtained in at least one of the studied targets were considered promising compounds. In fact, after analysing their scaffold, a trend can be seen because all of them have three or more aromatic rings and most of them have two indole scaffolds (Figure A 1). Some interesting kinase inhibitors described in recent years corroborate this finding, since they incorporate an indole moiety on their structures [24,25,28,54–56].

Moreover, looking at the top ranked compounds, it is easily observed that all of them have a bromine (Br) substituent. Actually, all the 24 compounds have at least one Br atom, a differential signature of marine compounds respect to terrestrial molecules. Unlike terrestrial species, many marine organism produce halogenated metabolites [43]. This corroborates the proposed importance of the indole scaffold on the kinases inhibition and seems to point out that the combination of an indole scaffold with halogen substituents could be a good starting point to design new possible inhibitors of the

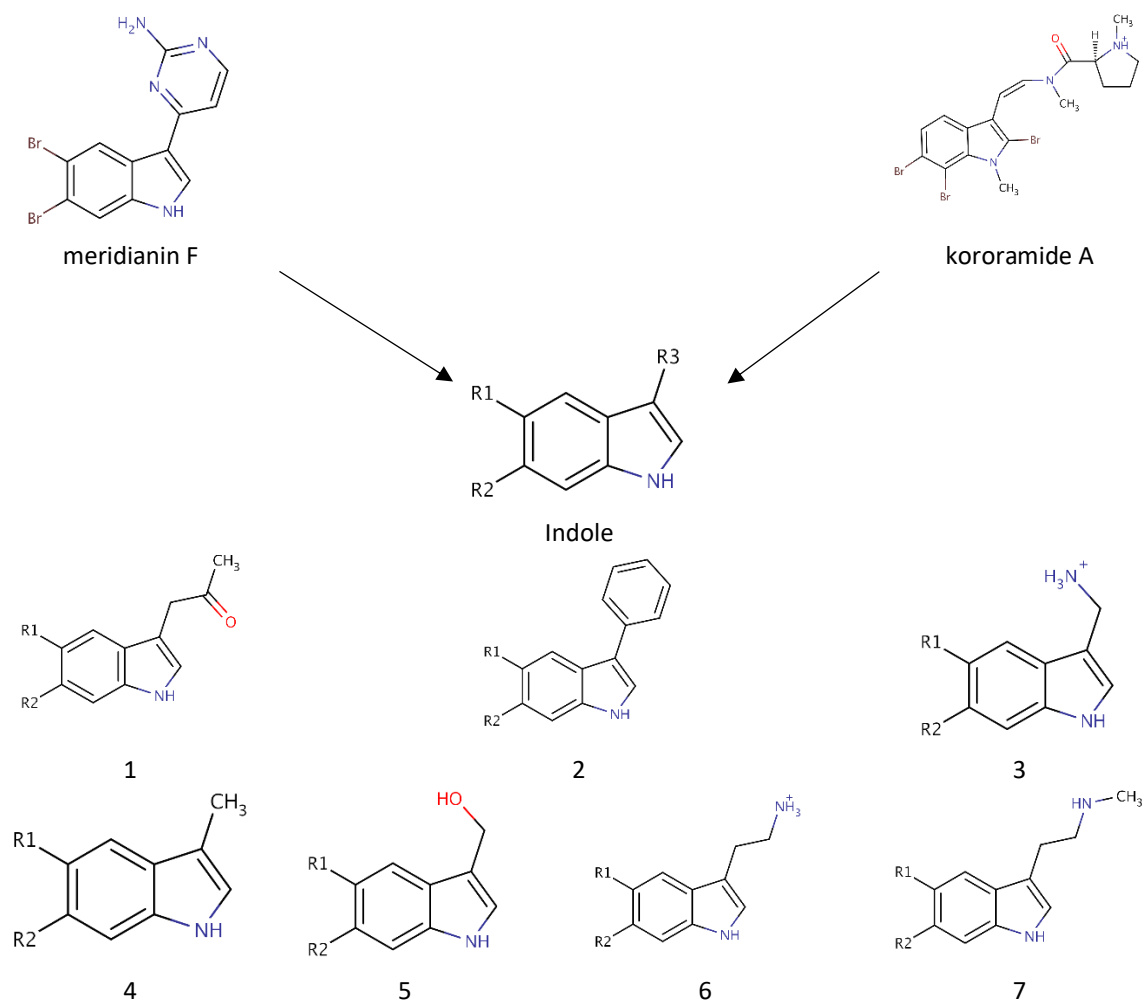
four kinases. This hypothesis is not an isolated fact as marine compounds with this moiety, different to meridianin F and kororamide A, have shown inhibitory effects against some of the studied kinases [24,28].

Indole derivatives

As mentioned above, the SAR hypothesis, widely used in drug discovery, has the premise that structurally similar molecules have similar biological activities and thus similar biological targets. Several known kinase inhibitors possess this moiety and some of them even present an halogenated version of it. In a previous work, we showed that meridianin F, which has an halogenated indole scaffold was the more active member of the family, highlighting the role of this moiety. Now, we have observed that kororamide A and B, given the similarity to meridianin following the SAR principle, could be possible inhibitors of these kinases. This fact is at least partially confirmed (further experiments are needed for a complete validation) because of the *in silico* obtained binding energies over GSK3 β , CK1 δ , DYRK1A and CLK1, reported above. All these facts together with the observed results at Table 3, made us hypothesize that starting from an halogenated indole moiety and following structural features extracted from meridianin F and kororamide A, we could design indole derivatives that could become kinase inhibitors. Concretely, the indole group was used as a template for the design of a series of seven analogue compounds with different fragments attached to the R3 position of the indole (compounds 1-7; Table 4) and substituted with different combinations of halogen atoms at positions R1 and R2 (a-g combinations; Table 4). Altogether, 49 compounds were designed.

Marine animals have demonstrated to be rich sources of halogenated metabolites and halogenated compounds have a wide range of biological activities [42]. Most halogenated drugs contain fluorine (F), followed by chlorine (Cl) and bromine (Br). Contrastingly, for marine-derived molecules, rather than chlorine, the most prevalent halogen found is Br [57]. Halogenated molecules are interesting therapeutic opportunities and it is estimated that one quarter of the total number of final compounds synthesized have an insertion that involves halogens [58]. Halogenated ligands lead to more stable complexes than non-halogenated ligands, and this is important to explain molecular recognition or to planning an screening study [59]. Moreover, the capability of halogen atoms to improve oral absorption, lipophilicity, blood brain barrier (BBB) permeability, metabolic and chemical stability, or even potency is well known [59,60]. Therefore, the three mentioned halogen groups at R1 and R2 positions were introduced and evaluated per compound (1-7 + a-g) with the aim of designing the best possible kinase inhibitors (Table 4).

Table 4. Scheme of the chemical design. The indole scaffold of meridianin F and kororamide A was selected to derive new compounds. More precisely, seven indole derivatives were designed (compound 1-7). The R3 position was fulfilled with diverse structural elements mainly inspired on meridianin and kororamide structures. Compound 1 with a ketone group, compound 2 with an aromatic ring, compound 3 with a methylamine, compound 4 with methyl group, compound 5 with methanol, compound 6 with an ethylamine and compound 7 with an ethylmethylamine. The R1 and R2 positions were completed with the permutation of Br, Cl and F halogen atoms (a-g) over both positions. At the end 49 indole analogue compounds were designed.

**Compounds**

a
b
c
d
e
f
g

R1

Br
F
Cl
Br
Br
F
Cl

R2

Br
F
Cl
F
Cl
Br
Br

***In silico* binding and binding mode analysis of indole derivatives**

To analyze the feasibility of the designed compounds as kinase inhibitors by an *in silico* binding analysis, their binding mode and binding strength against GSK3 β , CK1 δ , DYRK1A and CLK1 were analysed. To start with, docking experiments were performed. A total of 441 poses per target were obtained from the 49 compounds of the set. Thereafter the binding behaviour of all the poses was analysed, showing that the most populated binding region is, as expected, the ATP cavity. With all these results in hand, best poses per target in terms of binding mode and binding energy were selected to perform short MD simulations, for post-processing docking results. For some derivatives any pose was considered for further studies, as the selection of best compounds was carried out considering not only the binding energy but also the binding mode of each molecule, after an interaction analysis study. The poses that did not present good interactions were discarded. Finally, 166 simulations were carried out, corresponding to diverse poses belonging to 45 compounds for GSK3 β , 30 for CK1 δ , 46 in the case of DYRK1A and 45 for CLK1. After MD simulations, the binding energies of the target-ligand complexes were estimated by Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations. Table 5 summarizes the binding energies of the best indole derivatives, obtained after MD, per compound (1-7) and target. The rest of the binding energies obtained per derivative and target are reported at Table A 1, Table A 2, Table A 3 and Table A 4, respectively.

Table 5. Summary of Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the best derived analogues over the four targets studied. Lowercase letters indicate the halogen substituent group (a-g), as described in **Table**.

		GSK3 β		CK1 δ		DYRK1A		CLK1
		Binding Energy		Binding Energy		Binding Energy		Binding Energy
Compound 1	a	-30.3141	a	-35.4499	e	-32.8862	g	-30.3541
Compound 2	a	-31.2458	e	-37.8982	a	-37.8422	a	-34.1041
Compound 3	a	-13.8779	g	-28.7631	a	-15.2733	f	-20.4786
Compound 4	a	-27.6481	e	-28.6573	a	-30.7518	e	-28.3695
Compound 5	a	-27.6534	e	-28.5831	a	-31.2535	c	-29.4190
Compound 6	a	-18.5779	a	-26.4630	a	-18.9387	g	-30.7737
Compound 7	a	-18.8955	a	-18.4901	a	-20.8203	g	-25.4765

All energy values are kcal/mol.

As a general result, we observe that all the evaluated compounds present better binding interaction energies against CK1 δ , DYRK1A and CLK1 than GSK3 β , as observed for meridianins [37]. Also,

as a general trend, compound 1 and 2 always show better energies than the rest of derivatives, highlighting that the fragments (a ketone and an aromatic ring, respectively) introduced in the pyrrole ring (R3 position) of the indole scaffold could have beneficial effects to achieve better inhibitory activities over the ATP binding site of the four studied kinases. Finally, it must be remarked that the designed compounds that do not work against the kinases are different for each one of them, thus opening the door for exploiting these differences in the future to gain selectivity over the four analysed kinases.

GSK3 β

As said, the best docked complexes were selected to perform further analysis. For GSK3 β 75 poses were chosen and over them MD simulations were performed. From the total studied set, and with the aim of analysing the diverse derivatives, the best a-g combination for each of the 1-7 compounds per target was selected. Over the seven best compounds found after MD simulations in terms of binding energy, further analyses were performed, extracting some interesting features. Focusing on the halogen substituents, the best compounds are always those that contain two Br atoms at R1 and R2 position, reflecting the importance of Br substituents observed in previous studies [24,37].

A general pattern regarding the interactions performed by each of the seven best derived compounds at the catalytic ATP binding site was observed. In general, I62, V70, A83, V110, L132, D133, Y134, V135, Y138, and L188 are the most important aminoacids for their stabilization over the ATP catalytic pocket (Figure 4). The NH indole group is essential to establish hydrogen bond interactions with the carboxylic acid group (deprotonated under biological conditions) of D133 and/or V135. AR, described by LDYV motif, accommodates the seven best compounds, all of them showing the same binding mode/pose, stabilized by hydrophobic contacts. The indole group is wrapped by N-terminal I62 and V70 residues found near the GRR, together with A83 placed at the HP and C-terminal residues V110 and L188 present at the SP. As the binding mode analysis reveals, all the compounds have the same binding mode, thus binding energy results and MD simulations were used with the aim of identifying some differential features among them. MD analysis reveals that the indole scaffold is maintained wrapped in the same position during all the simulation while the fragments introduced at R3 are more flexible. A binding energy analysis showed that compound **2a** has a slightly better energy than compound **1a**, although both could be considered good plausible options, as the binding energy differences are around 1 kcal/mol, that seems to point out compound **2a** as the best possible inhibitor.

Looking at the literature, our results show that the binding mode displayed by most of the analysed compounds, specially by compound 2a, correlates with the binding mode of known inhibitors, and also that the residues involved on it are key binders [35,61].

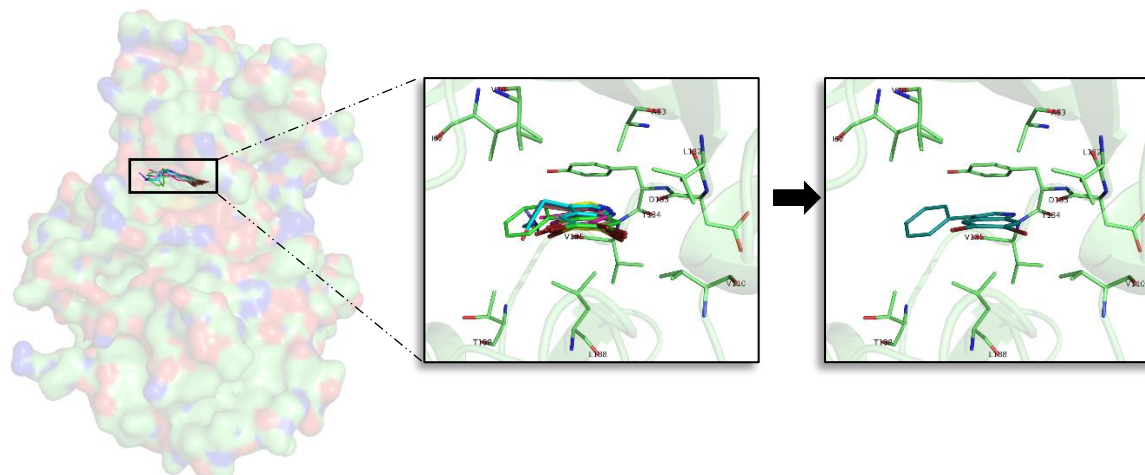


Figure 4. Superposition of the seven best compounds over GSK3 β (PDB code: 3PUP) ATP cavity. The active site aminoacid residues involved in the binding of the best compounds and the binding position of each of them are enlarged. In the first enlarged panel the seven top compounds are represented, whereas in the right panel only compound 2a is shown. Original from the author.

CK1 δ

For CK1 δ , 97 docking poses were subjected to MD simulations. Thereafter, the seven best compounds, in terms of binding energy were selected to be further analysed. Differently to GSK3 β , there is not a common binding mode shared by the seven derived analogues and there is not a specific location of the halogens in the ATP binding site, that can be inferred from the observed binding modes. Although a general pattern could not be observed, there are common features between the studied derivatives than can be highlighted. There is a common behaviour between compound 1, 6, and 7, and compounds 2, 4, and 5 (Figure 5). For the first group the best halogen composition is Br-Br (compound 1a, compound 6a and compound 7a), whereas for the second group, the best halogen composition is e (Br-Cl), while for compound 3, that behaves differently to the rest of the compounds, is g (Cl-Br). In all compounds a Br atom is present, which seems to indicate that this presence could be important to increase the binding strength. In general, with few exceptions, the worst binding energies are obtained when there is no Br atom present. This trend is also observed on the rest of kinases (Table A 1, Table A 2, Table A 3 and Table A 4). In addition, an accurate analysis of the most important residues involved on the seven compounds binding mode, was performed. This analysis reveals that despite each compound

has a different binding pose, there are conserved interactions at the ATP catalytic cavity. According to that, the most important residues on the binding of the seven compounds to CLK1 δ are I23, A36, Y56, L84, I148, and D149. All the seven derivatives are placed between the HP defined by A36 and the residue I23 that is placed near the GRR, both zones located at the N-terminal region and L84, I149 and D149 placed at the AR and PBP at the C-terminal domain. All the interactions observed between the analogues and the residues are mainly hydrophobic contacts. Binding energies reveal that compound 2e (Br-Cl) seems to be a slightly better inhibitor than compound 1a, although both can be considered good options as the energy differences are around 2 kcal/mol.

Different studies have been addressed to find novel and potent CK1 δ inhibitors in the last years. Looking at them, it is easy to observe that the interactions made by of all these molecules are aligned, validating it, with the binding mode of our proposed derivatives [27,32,33,36].

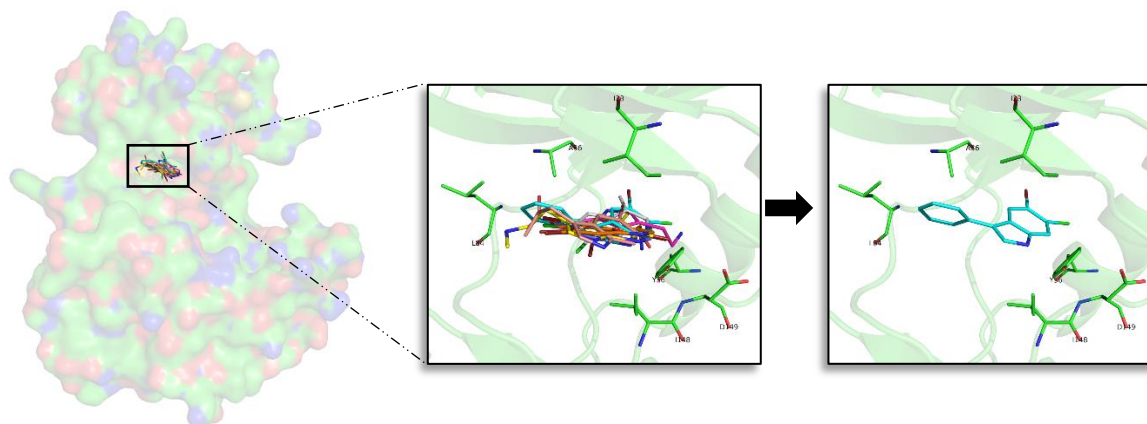


Figure 5. Superposition of the seven best compounds over CK1 δ (PDB code: 4KBK) ATP cavity. The active site aminoacid residues involved in the binding of the best compounds and the binding position of each of them are enlarged. In the first enlarged panel the seven top compounds are represented, whereas in the right panel only compound **2e** is shown. Original from the author.

DYRK1A

For DYRK1A, 72 docking poses were selected for further analysis. MD simulations were performed over all of them, and thereafter the best compound per target, as for the rest of kinases, was selected. Despite the indole derivatives tested do not shown a shared binding mode as GSK3 β , it is more conserved than for CK1 δ . All compounds, except compound 3 that is oriented right upside down and moreover shows the worst binding energy, shared the same placement at the ATP catalytic pocket (Figure 6). Analysing the halogen composition of the best compounds it is observed that Br-Br, at R1-R2

positions, is the most common substituent; only compound 1 has a different combination (Br-Cl). As a general conclusion, as with the other three kinases, the presence of at least one Br atom is important to have a good binding affinity.

Looking forward to extract common patterns from the binding modes of the top seven derivatives, it is clear that all these compounds placed at the catalytic ATP cavity are interacting with I165 and V173, both residues delimitate the GRR, and A186 that is found at the HP, all of them located at the N-terminal region. The important AR formed by a FEML motif also participates on each of the seven bindings, being F238 and L241 the most important residues to stabilize the analysed derivatives. At the C-terminal region, V306 and D307, present at the PBP, are also key binders. Interaction analysis reveals that most of the interactions performed by the derived analogues were hydrophobic contacts. For DYRK1A, after analysing the MD obtained results, it is observed again that compound 2a, is the best derivative in terms of binding energy and binding mode. Interestingly, the observed binding patterns are shared by most of the known inhibitors of this target, that could be found in the literature. Even more, all of them are proposed as ATP competitive inhibitors like the derivatives we described here [25,26,34].

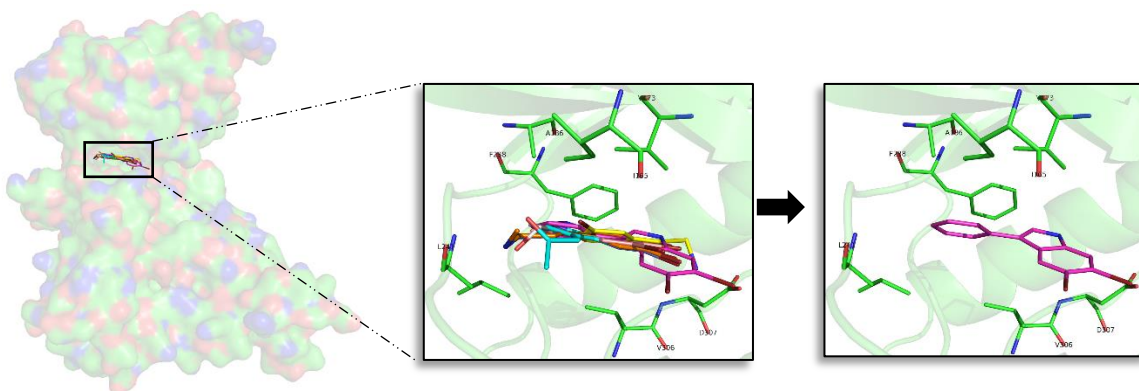


Figure 6. Superposition of the seven best compounds over DYRK1A (PDB code: 4AZE) ATP cavity. The active site aminoacid residues involved in the binding of the best compounds and the binding position of each of them are enlarged. In the first enlarged panel the seven top compounds are represented, whereas in the right panel only compound **2a** is shown. Original from the author.

CLK1

For CLK1, 87 docking poses were selected for further analysis. All of them were subjected to MD simulations selecting then the best one per target. A first binding mode observation reveals that a common binding mode was found for compounds 1, 2, and 6 (Figure 7). These three compounds have the best binding energy, and this could point out the importance of R1, R2 and R3 substituents to gain inhibitory capacity. Compound 3, despite having a similar binding pose, does not show good energies.

The other compounds (4, 5, and 7) show slightly lower binding energies and a different binding mode, even between them. Focused on the halogen groups, in this target there is not a clear trend, as the seven best compounds show five different halogen substituents (a, c, e, f and g). Despite this fact, not observed in the rest of studied kinases, a similar trend can be seen. Most of the seven top compounds have a Br atom, except compound 5c. Moreover, in agreement with the rest of compounds the seven top derivatives are mainly Br or Cl substituents on R1 or R2 position, with the exception of compound 3f. This seems to suggest that, as for the other targets, all of the analysed halogen substituents combinations could give good inhibitory results, but the presence of a Br is a key factor. In fact, for this target, as seen for the other kinases, the compound 2a is the best one in terms of binding energy.

A detailed analysis of the displayed binding modes by each compound at the ATP cavity site, reveals interesting shared patterns. On the N-terminal domain L167, F172, and V175 can be found at the GRR, and A189 at the HP acting as key binders. Adenine motif FELL was also revealed important, in particular F241 and L244. On the C-terminal region, residues E292 and L295 at the SP and V324 placed at PBP are the most important aminoacids to stabilize the derived compounds over CLK1. As compound 2a, the other best compounds tend to point their halogen groups between the AR and the HP, fact that facilitate residues as F175 placed at the GRR and E292, or L295 placed at the opposite SP, surround and fixed the indole scaffolds. Interestingly, the binding mode exposed here for the derived analogues in general, and also for the best compound, 2a, in particular, is validated by other inhibitors reported in the literature [30,31].

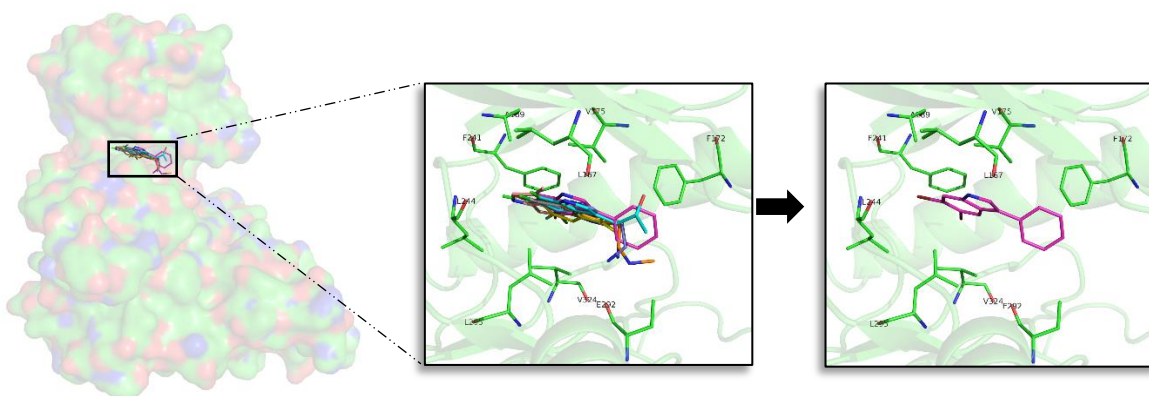


Figure 7. Superposition of the seven best compounds over CLK1 (PDB code: 2VAG) ATP cavity. The active site aminoacid residues involved in the binding of the seven best compounds and the binding position of each of them are enlarged. In the first enlarged panel the seven top compounds are represented, whereas in the right panel only compound **2a** is shown. Original from the author.

The *in silico* binding studies performed over the four kinases indicate that the derivatives coming from compound 2, 2a and 2e, located at the middle of the ATP binding cavity, seem to be the most plausible ATP competitive inhibitors. However, other derivatives, especially for compound 1 should not be discarded. In general, the presence of the benzene ring at position R3 could have a more positive influence on compounds stabilization at the catalytic site than other substituents. Looking at the literature, several inhibitors described for GSK3 β , CK1 δ , DYRK1A, and CLK1, as well as other members of the protein kinase family, have aromatic rings in the terminal positions. Moreover, the analysis of the effect of the halogen groups used as substituents at R1 and R2 positions pointed out that its presence can influence the binding strength of the complex (ligand-target). In general, if at least one of the substituents is a Br atom the binding energy is better. An interesting trend found here is that Br seems to be the “best” halogen, followed by Cl and F, which in general give worst binding energies. This finding is in line with what is observed in nature, since marine natural halogenated indole alkaloids contain mostly bromine and chlorine, being the iodinated and fluorinated compounds less abundant [43].

Selectivity

One of the most important challenges on the design of novel kinase inhibitors is the lack of selectivity over the ATP binding site, which is critical in clinical effectiveness of most drugs (Traxler & Furet, 1999; Huang et al., 2010). Most kinase small-molecule inhibitors bind to the ATP catalytic cavity near the AR and wrapped by GRR and HP on the T-lobe and SP and PBP at the C-lobe. The herein performed study does not reveal any significant selectivity over the four kinases for any of the analysed compounds, which could be easily observed looking at the obtained binding modes and energies. However, analysing the residues involved on the binding and the regions occupied by the analogues, some interesting trends that could be exploited in the future can be observed. Interestingly, regarding the binding modes, the best binding energies were obtained on those compounds that are (partially) placed at the PBP. This region, that is very exposed to the solvent and is not usually exploited to gain binding affinity, can be useful to improve the inhibitors selectivity since it contains non-conserved aminoacids [46].

Regarding the binding energy results per se, without having into account the binding mode, remarkable significant differences are not observed. The best compound for each target (2a and 2e respectively) comes from the same scaffold, being compound 2a the best theoretical inhibitor for three of the four targets. If we analyse the binding energies of these top compounds, compounds 2a and 2e over DYRK1A and CK1 δ , respectively, show a better interaction energy, around 6 kcal/mol of difference,

respect to the binding energy of compound 2a over GSK3 β and 3 kcal/mol over CLK1. However, although a slight preference could be inferred from this, the binding of these four compounds to all the four targets is possible with a reasonably good strength. In general, the main differences are observed between the derived compounds 2 (mainly) and 1, which seem to have better energies than those molecules coming from analogues 3 to 6, and specially respect to the molecules coming from analogue compound 7 (Table 5). For GSK3 β the best compounds coming from derivatives from 2 and 1 (1 kcal/mol of difference between them) are displaying the best interaction energies, followed by those from analogues 4 and 5 (around 4 kcal/mol of difference to compound 2a), and finally the worst compounds come from analogues 3, 6 and 7 with differences around 13 to 17 kcal/mol respect compound 2a. In the case of CK1 δ , as for GSK3 β , the top ranked compounds from analogues 2 and 1 (1.5 kcal/mol of difference between them) have the best binding energies, followed by those from analogues 3, 4, 5 and 6 with differences around 9 to 11 kcal/mol respect compound 2e, and finally compound 7a with a difference of around 19 kcal/mol respect to 2e. For DYRK1A, the best from compound 2 is the top molecule in terms of interaction energy. Compound 1a shows a difference of around 5 kcal/mol, whereas compounds 4 and 5 present differences between 6.5 and 7 kcal/mol, respectively, and compounds coming from scaffolds 3, 6 and 7 between 17 and 22.5 kcal/mol. In the case of CLK1, compound 2a has the better binding energies, followed by those from analogues 1 and 6 (differences around 3 kcal/mol), molecules derived from compounds 4 and 5 (differences around 5 kcal/mol), and finally those from analogues 3 and 7, that show differences around 8.5 to 13 kcal/mol respect to the binding energy.

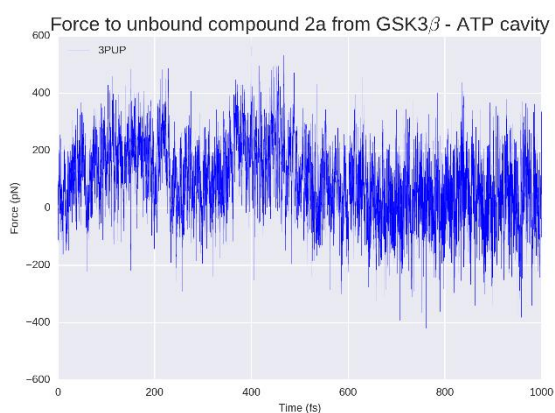
Looking to compounds 1-7 per target, it can be observed that for compounds 4 and 5 the binding energy differences between the top a-g derivatives range between 2 and 4 kcal/mol between the four kinases. For compounds 1, 2 and 7 the differences are higher, ranging between 2.5 and 7 kcal/mol, depending on the compound and target. Finally, for compounds derived from analogues 3 and 6 the differences are even higher, ranging between 8 and 15, and 4 and 12 kcal/mol, respectively. In general, there is not any noticeable selectivity trend derived from the binding energy, although there are some features that could be further exploited. For instance, for DYRK1A and CLK1 an aromatic ring at R3 position is the best choice to gain activity over them, whereas for GSK3 β and CK1 δ a ketone group at this position could also work, enhancing a way to design selective compounds at least for some of the four kinases.

Exploring the effect of the halogen atoms over the binding strength, as said above, some general trends could be observed but again, its presence does not give any clearly marked or significant selectivity trend between targets. The presence of Br atoms seems to increase the binding strength more than the presence of Cl or F, being in general Cl “better” than F to get good energetic results. However, a possible selectivity feature could be observed due to compound 2e. Docking energy results are similar for the four kinases, but it only performs good interactions for CK1 δ . This is the reason why MD simulation over this compound was only performed in complex with it, while for the other three kinases it was not selected. Compound 2a gave good docking energies for all four targets but performed good interactions only with GSK3 β , DYRK1A and CLK1, so the fact of having a Br-Cl combination at R1-R2 plus an aromatic ring at R3 could be a sign of selectivity over CK1 δ , although this should be further explored, as other Cl combinations give good results for the other kinases (Table A 1, Table A 2, Table A 3 and Table A 4).

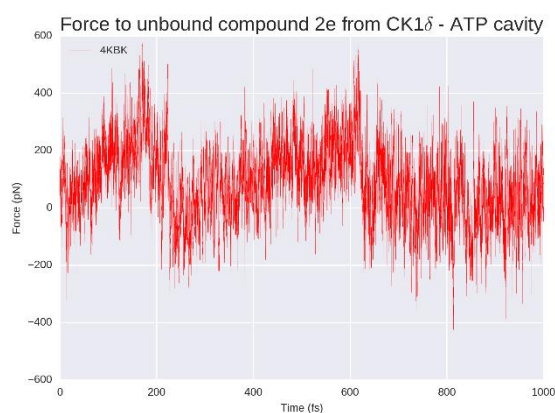
2a and 2e unbinding

To reinforce and validate the observed binding trends, as well as to find a differential feature that could help to enhance the selectivity of future derived compounds over the four kinases, steered molecular dynamics (SMD) simulations were performed. Since at the energy and binding mode level there are no significant differences, we intended to see if there was some type of selectivity derived from the protein structure that influences the facility/difficulty of unbinding of the most promising inhibitory compounds 2a and 2e per target (Figure 8).

A)



B)



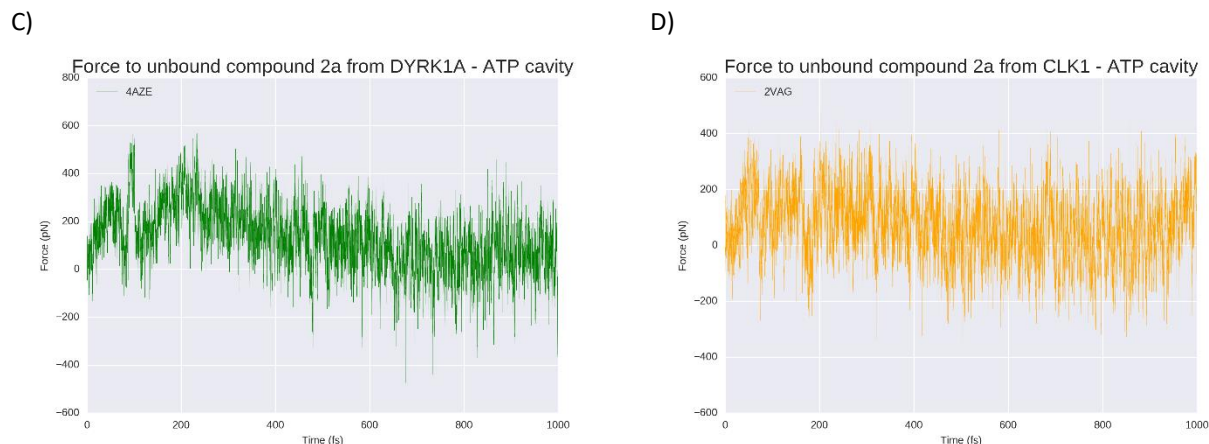


Figure 8. Exerted force in piconewtons (pN) needed to A) remove compound **2a** (blue) from the GSK3 β ATP catalytic cavity, B) remove compound **2e** (red) from the CK1 δ ATP catalytic cavity, C) remove compound **2a** (green) from the DYRK1A ATP catalytic cavity and remove compound **2a** (orange) from CLK1 ATP catalytic cavity. The x axis represents the computational residence time in femtoseconds (fs).

At the beginning of each simulation ($t = 0$), the compound is in the bound state, placed inside the ATP cavity interacting with the residues previously described. After 1 nanosecond (ns), at the four kinases, the ligand is completely out of the cavity. In the case of GSK3 β (Figure 6A) a force of around 400 pN (piconewtons) is needed to extract compound **2a** from its catalytic cavity. The compound dissociation from the target takes place at 200 femtoseconds (fs), moment when the force decrease approach zero pN which means that the compound is out of the cavity. For CK1 δ (Figure 7B) that hosts the best compound in terms of binding energy, **2e**, the necessary force to break the ligand-target complex is higher than for the GSK3 β complex, with forces that reach up to 600 pN. The ligand unbinding takes place at a similar time than for GSK3 β complex, around 200 ps although it takes slightly more time. The dissociation of compound **2a** from DYRK1A (Figure 7C) is done in two phases. A primary rupture force seems to occur before 100 fs, and immediately afterwards the highest energy point can be observed (around 500 pN), corresponding to the second break. A visual inspection of the SMD confirmed that at this moment, the compound is still inside the ATP pocket. Over 200 fs, the force, after a progressive decrease, arrives to zero pN. This progressive decline correlates with the progressive loose of interactions during the way out of the compound from the catalytic cavity. For the CLK1-compound **2a** complex (Figure 7C) a similar situation is observed. A primary rupture around 100 fs, moment when, as for DYRK1A, the compound is still placed at the ATP binding site and it is not until later on 200 fs, when a sudden drop in the energy can be observed, indicating the complete loss of interactions and therefore, the leaving of the cavity. As a general trend, around 200ps **2a** and **2e** compounds leave the catalytic pocket of the different kinases, requiring a different amount of force that is in line with the

observed binding energy. Usually, the better the binding energy, the higher the force needed to break the complex and the longer the residence time. In that sense, the SMD results corroborate what has been seen so far: CK1 δ -2e and DYRK1A-compound 2a complexes that have the higher binding energies also seem to have (slightly) longer residence times and require a higher force to take out their respective ligands from their catalytic pockets. Thus, there is no feature that suggests a selectivity trend derived from the unbinding process that could not be extracted from the binding energy results. Compound 2a is more selective (it binds stronger and requires a higher effort to remove it) for DYRK1A than for GSK3 β and CLK1 but could bind to all of them. Compound 2e seems to be more difficult to unbind than compound 2a, but this correlates with the higher binding energy it shows after MD.

Pharmacokinetic properties of kororamide A-B, convolutamine I-J and the designed derivatives

Due to the importance of pharmacokinetics (PK) and its impact on drug discovery process, convolutamine I-J, kororamide A-B and the whole set of 49 analogues compounds were analysed, studying their ADME/Tox features. The PK properties of the two best derived compounds 2a and 2e are summarized on Table 6 and Table 7. The full set of derivatives were also analysed and results can be found at Table A 5 (absorption and distribution) and Table A 6 (metabolism, excretion and toxicity).

The first PK property analysed was molecular weight, and all compounds show values under 500 Dalton (Da). The higher molecular mass was found for compound 2a with 351 Da, which is in good agreement with the sizes that a small therapeutic molecule that should cross the BBB should have.

Absorption properties

Absorption describes the process by which drug candidates move from the point of administration to the blood. LogS descriptor confirmed good solubility in water and good bioavailability for each compound. The derivatives coming from compounds 2 show values that are between -5.1 and -6.1, while for the rest of the derivatives, values go between -3 and -4. Caco-2 permeability revealed medium to high values for all the compounds, except for kororamide A that was low. The compounds that have a benzene at R3 position as well as the derivatives with F at R1 and R2 positions show moderate permeability and should be optimized in the future. Regarding P-glycoprotein (Pgp) binding, none of the compound was predicted to act over it. The interaction with Pgp has many pharmacological implications that could result in pharmaceutical advantages or contraindications. For instance, Pgp modulation has been suggested as a mechanism to improve CNS pharmacotherapy [62–65], but none of the derivatives here seem to have this ability. On the other hand, intestinal absorption values higher

than 30% are considered well-absorbed compounds, and for all our set the obtained values were higher than 89%. All these absorption results suggest good absorption properties for the 49 designed derivatives, plus kororamides and convolutamines.

Distribution properties

Distribution describes the migration of a compound from the circulation to the extravascular system. LogP values lower than 5 indicate that the compounds have an appropriate hydrophobicity and permeability. In that sense, the derivatives coming from compound 2, as well as convolutamine J and kororamide A have the highest values (≈ 4) while the rest of compounds are between 2 and 3. Opposite to LogP behaviour, plasma-protein binding (PPB) and steady state volume of distribution (VDss) are not showing as good tendencies for the best derivatives compounds coming from scaffold 1 and 2, and Convolutamine J. Most of the analysed molecules showed medium to high PBP values (except kororamide A and compounds 3b, 5b, 6b and 7b with low PPB values) indicating that a high percentage of the administrated compounds will be found attached to proteins, affecting its diffusion and its efficiency. As less bound a drug is to plasma proteins, the more efficient it is, as it can traverse cell membranes or diffuse. Regarding VDss, derivatives from scaffolds 3, 6 and 7 and convolutamine I-J plus kororamide A, have high VDss values (>0.45), while for the rest of compounds, distribution is low to medium, in a close agreement with PBP results. BBB descriptors with results higher than > 0.3 reveal good distribution to the brain, as they could pass the blood brain barrier. The highest values are found for convolutamine I, kororamide A and the derivatives coming from scaffolds 2 and 4, as well as for compounds 5b and 7b. However, it should be considered that most of the compounds not predicted to cross BBB, have values near the threshold. In addition to BBB, Central nervous system (CNS) permeability was measured. This seems to be a more precise measure than BBB, as it is a more direct measurement [66]. Kororamides and convolutamines do not show good permeability values, whereas all the derived compounds showed good results (>-2) allowing us to consider that most of the designed compounds could penetrate the CNS, specially the compounds coming from scaffolds 2 and 4, among which there are the two best candidates 2a and 2e.

Metabolism properties

Cytochrome (CYP) P450 is an important enzyme used to predict drug metabolism. Many drugs could be deactivated or activated by CYP450, as cytochrome P450 enzymes that can be inhibited or induced by drugs, resulting in clinically significant drug-drug interactions that can cause unanticipated adverse reactions or therapeutic failures. Our results revealed that all the analysed compounds, except

kororamide B, are likely to be metabolised by CYP450, so their properties should be carefully analysed to design lead compounds from the herein studied molecules [67,68].

Excretion properties

Regarding excretion properties, describing the transport of drugs into the urine or bile, good results were obtained. It was found that only kororamide A and B tend to act as a substrate of the organic cation transporter 2 (OCT2 or Solute carrier family 22 member 2, SLC22A2), which means that in general, and for the two best derivatives 2a and 2e, non-clearance problems and adverse interactions with co-administrated OCT2 inhibitors are expected.

Toxicity properties

During drug development, safety is always the most important issue, including a variety of toxicities and adverse drug effects that should be evaluated in preclinical and clinical trial phases [69]. Between the measured properties, the inhibition of the potassium channels encoded by the human ether-a-go-go gene (hERG) is basic. Our results indicate that none of the compounds seem to be toxic due to hERG. In the same way, none of the designed derivatives is susceptible to be hepatotoxic. However, convolutamine I and J as well as kororamide A tend to be hepatotoxic. Looking at AMES toxicity, which predicts mutagenic and carcinogenic properties, our results revealed that the derivatives from compound 2, as the top derivatives 2a and 2e, and kororamide B are predicted to be toxic, while the rest of the set does not. Regarding the maximum recommended tolerated dose (MRTD), the four brominated alkaloids as well as compounds coming from scaffolds 3, 6 and 7 showed low values/doses, which is not the best scenario, whereas the rest of the compounds present good MRTD values.

Table 6. Summary of absorption and distribution properties of the two best compounds **2a** and **2e** found on the four studied kinases. BBB: blood brain barrier, PPB: protein-protein binding, VDss: steady state volume of distribution, CNS: central nervous system.

Compound	Absorption					Distribution				
	Mol weight	LogS	P-Glycoprotein	Caco-2 permeability	Intestinal absorption	LogP	BBB	PPB	VDss	CNS permeability
Compound 2a	351	-6.1	inactive	Moderate	90.067	4.1	0.477	High	0.234	-0.894
Compound 2e	290.1	-5.7	inactive	Moderate	91.036	3.8	0.508	High	0.076	-0.92

Table 7. Summary of metabolism, excretion and toxicity properties of the two best compounds **2a** and **2e** found on the four studied kinases. CYP: cytochrome, OCT2: organic cation transporter 2, hERG: human ether-a-go-go gene, MRTD: maximum recommended tolerated dose.

Compound	Metabolism	Excretion	Toxicity			
	CYP450	OCT2 substrate	hERG	MRTD	AMES toxicity	Hepatotoxicity
Compound 2a	Yes	No	<4.0	0.673	Yes	No
Compound 2e	Yes	No	<4.0	0.641	Yes	No

The well-known Lipinski's rule of five, formulated in 1997 and that remains in force [70] was also used in combination with the different ADME/Tox properties described above with the aim of evaluate/determine druglikeness of the analysed compounds. To assess how druglike is a substance based on Lipinski's rules it is accepted that it should have 1) not more than five hydrogen bond donors and 2) ten hydrogen bond acceptors, 3) a molecular mass less than 500 Da and 4) a LogP not greater than 5. Focusing on the two best compounds (2a and 2e), both have one hydrogen donor and no acceptors. Also, as seen in Table 6 and Table 7, the other Lipinski requirements are also met. Thus, taking into consideration all the ADMET results described previously, these two compounds can be proposed as good hit candidates, having into account that some properties, such as the possible carcinogenesis and mutagenesis problems should be carefully addressed. In fact, absorption, distribution metabolism, excretion, and toxicity properties should be more or less improved for all the designed compounds, in a further Hit to lead (H2L) optimization process. Toxicity should be removed, and compounds interaction with cytochrome P450 carefully analysed and, given the case, eliminated or modulated. Moreover, Caco-2 permeability could be increased as well as their distribution properties should be improved, lowering the PPB and VDss, to be able to diffuse and penetrate into cells easily.

Materials and Methods

Computational virtual screening

It is well known that there is a correlation between (chemical) structure and (biological) activity, Structure Activity Relationship (SAR). This SAR is widely exploited in many aspects of the drug discovery pipeline, ranging from compound screening to lead optimization processes, at experimental and computational level. Herein, we have performed a 2D virtual screening search over MarinLit database using its substructure search functionality. Using as an input meridianin F and kororamide A (the two indole compounds that have shown a better binding strength against the four analysed kinases), as well as the indole scaffold alone, a similarity search was performed over MarinLit obtaining a list of compounds having an indole scaffold in their structure and/or being structurally similar to meridianin F and/or kororamide A. The name and structure of the similar compounds could be found at Figure A 1.

Structure Modelling

Meridianin F structure was modelled from the 2-dimensions (2D) chemical structure published by Núñez-Pons and co-workers [71]. Convolutamine J, I, and kororamide A and B, were modelled from

Dashti et al. [38]. Ligands were prepared to generate energetically minimized 3 dimensions (3D) coordinates.

To do computational work, obtaining good structures to start with is crucial, so prior to any calculation, the best structures of the four analysed targets were modelled from 3D crystal structures extracted from the Protein Data Bank (RCSB PDB) [72]. Those structures represented human targets and are the structures and chains that cover the maximum amino acid region sequence, in general, and the binding region of each of the selected targets in particular. Since all the four kinases biological assembly is in monomeric forms, GSK3 β and CK1 δ chain B and DYRK1A and CLK1 chain A were respectively selected to perform further studies. To do so, due to the fact that the four studied targets have 3D crystallographic structures, the ATP competitive inhibitors OS1 co-crystallized with GSK3 β (PDB: 3PUP) [35]; 1QG co-crystallized with CK1 δ (PDB: 4KBK) [36], the crystal structures of DYRK1A in complex with 3RA (PDB: 4AZE) [26] and finally, V25 co-crystallized with CLK1 (PDB: 2VAG) [31] were used as a template to perform rigid docking calculations using Itzamna (Mind the Byte.SL, Barcelona, Spain) [73].

Docking calculations

Docking calculations can identify small molecules (ligands) that fit well into the putative binding pocket of a given protein (target) in an optimal way. Without any other specification, generally speaking, docking refers to classical (rigid) docking where only the movement of the ligand is allowed [74]. This kind of calculations allows to elucidate the candidate binding mode (pose) that is predicted when ligand and receptor bind to each other, and scoring functions allows to classify and rank compounds based on the binding energies obtained. Proteins are flexible entities, they move, and this dynamic is necessary to carry out any protein function. Taking into account this flexibility regarding the binding mode prediction, is key to obtain results capable of being experimentally correlated [75,76]. A good option to add flexibility at the complex is the post-processing of docking results, that consist on docking validation and/or refinement by MD simulations [77]. Rigid docking calculations can predict optimal ligand placement at the binding site of a target, but not all the interactions between the ligand and the receptor are usually depicted accurately. MD simulations can optimize the predicted binding mode, allowing to observe the so-called induced fit events arising from the conformational adaptation of the target to the ligand, and also check the stability of the docked complex, as a bad docking pose will generate an unstable MD trajectory, during which the ligand could even leave the binding site [75,78]. In this study, we have employed a pipeline aimed to simulate a flexible docking protocol in a similar way to other studies reported in the literature, in that we post-processed the obtained docking poses [79].

Concretely, the energetically minimized ligand conformations were docked into the active site of the four kinases studied; possible binding poses were determined and subsequently ranked based on their calculated binding affinities and then post-processed using MD simulations. This two-step protocol constitutes a (probably the most) practical and convenient approach to address the docking problem [77]. It is in general less computationally expensive and provides the results that we need in an accurate way, comparable to “real” flexible docking methodologies (such as ensemble-based or flexible induced-fit docking). In general, using MD as a post-processing tool, a smaller fraction of the conformational space is usually covered, but without the several limitations that affect sampling and scoring algorithms for docking.

All docking calculations were performed using Itzamna software tool (Mind the Byte.SL, Barcelona, Spain) [73]. Itzamna needs the structure of a molecule, or a set of compounds, to dock, as well as the cavity where it should be placed as an input. When the used 3D crystal structures were co-crystallized with a ligand, this cavity was employed. If it were not the case, the cavity was defined by the residues that describe the cavity. Docking studies were performed between kororamide A-B, convolutamine I-J and the set of 49 derived compounds against GSK3 β , CK1 δ , DYRK1A and CLK1. Two runs were carried out for each calculation to avoid false positives.

Molecular dynamics simulations

MD simulations are one of the principal tools in the computational study of biomolecules as the dynamic nature of proteins is a well-established phenomenon that these simulations can capture. It can be described as a theoretical method for studying the physical movements of atoms and molecules with the aim of exploring as much as possible the conformational states that proteins can adopt. MD calculates the time dependent behaviour of a ligand-receptor complex; in other words, atoms and molecules are allowed to interact for a fixed period of time, which gives a view of the dynamic evolution of the system.

Short (1 ns) MD simulations were performed using NAMD program version 2.11 over the best-docked complexes, which were selected based on Gibbs free energy (ΔG_{bind}) [80]. The Amber ff99SB-ILDN and the General Amber Force Field (GAFF) set of parameters were employed for modelling receptors and ligands, respectively [81,82]. The election of these force-fields was based on the fact that both have been extensively tested, being two of the most used for protein and protein-ligand simulations [81–84]. It has been shown that ff99SB-ILDN correlates consistently well with experimental data, and the GAFF force-field can conveniently and quickly produce reasonable ligand (especially

organic molecules) parameters. Moreover, as amber force-fields, both are compatible, giving combined satisfactory results in several studies. Ligand GAFF parameters were obtained using Antechamber, whereas the receptor structures were modelled using the leap module of Amber Tools [85,86]. Simulations were carried out in explicit solvent using the TIP3P water model with the imposition of periodic boundary conditions via a cubic box [87]. Electrostatic interactions were calculated by the particle-mesh Ewald method using constant pressure and temperature conditions. Each complex was solvated with a minimum distance of 10 Å from the surface of the complex to the edge of the box. Temperature was kept at 300 Kelvin (K) using a Langevin Piston barostat. The time step employed was 2 fs. Bond lengths to hydrogens were constrained with the SHAKE algorithm [88]. Before production runs, the system was energy minimized. Next, the solvent surrounding the protein was equilibrated at the target temperature using harmonic position restraints on the heavy atoms. Finally, the system was submitted to a slow heating-up phase, from 0 to 300 K. For the production run, all position restraints were removed.

Molecular Mechanics/Generalized Born Surface Area

The molecular mechanics energies combined with generalized Born and surface area continuum solvation (MM/GBSA) are popular computational approaches to estimate ΔG binding of small molecules to proteins [79]. These methods are used to predict ligand-binding affinities based on docking or MD simulations to get a more realistic view of the interaction of docked complexes. The obtained energies are more realistic than those obtained after docking calculations, as it is generally accepted that they outperform docking results, allowing a better ranking of the analysed compounds. It should be noted that these results although improve docking binding energy values, are far to be biological comparable. In our case and following similar approaches, we applied reweighting techniques, specifically MM/GBSA, over the generated MD trajectories for post-processing docking results [78,79,89].

MM/GBSA rescoring was performing using the MMPBSA python algorithm contained within the Amber Tools suite [90]. The snapshots generated at the end of MD simulations (1ns) were used as input into the post-simulation MM/GBSA calculation of binding free energy.

Steered molecular dynamics

Steered molecular dynamics (SMD) is a simulation tool used to explore processes, which cannot usually be achieved by standard MD simulation, such as ligand-protein unbinding or certain protein conformational changes. Here, we have employed it to study ligand unbinding processes. In that sense, in SMD simulations, a time-dependent external force is applied to the ligand, from an internal atom of

the protein, to facilitate its unbinding. For a given ligand bound to a target, it allows to establish a theoretical correlation between unbinding forces and residence time, and in turn its inhibitory capacity; the larger is the force and time needed to unbind a ligand from a receptor, the higher its binding affinity [91–94].

SMD simulations were performed using NAMD version 2.11 [80]. Simulations of compounds 2a and 2e over GSK3 β , DYRK1A, CLK1 and CK1 δ , respectively, were performed. The last frame obtained from the post-processing MD simulations was used as an input. A harmonic constraint force constant of 4kcal/mol/Å with a constant velocity of 0.00002 Å/ns was applied. The time length for each simulation was 1 ns, using a timestep of 2 fs, which was enough to observe the entire ligand unbinding process. The rest of the parameters of the simulations were the same employed for MD simulations. The generated trajectory was finally analysed using visual molecular dynamics (VMD) to extract the exerted force (pN) per simulation frame [95].

Interaction analysis

To analyse the key residues of the active site involved in the inhibitor binding, we examined the obtained binding modes after *in silico* binding studies, docking and/or MD simulations, with already known binders of each one of the targets. The known binders are key residues that have been revealed as necessary for the binding of known substrates/inhibitors and were identified through an evidence-base interaction analysis carried out by a bibliographical search plus a database analysis. The bibliographical search was conducted using several studies in which inhibitors against the selected kinases were identified describing each compound binding mode [25–27,30–36,61]. The database search was done using an in-house, recently constructed database. It was built by crossing ChEMBL and the RCSB PDB [62], and it contains all PDB structures per UniProtKB ID with active compounds (by now there are only PDBs with compounds not competing against cofactors). Moreover, the database also contains the residues to which each active compound (per PDB) is bound [96,97]. This allows to easily check whether the proposed derived compounds behave as a binder or not.

ADME/Tox properties prediction

ADME/Tox properties prediction were carried out using proprietary machine-learning (ML) models and the pkCSM webserver [66,98]. The proprietary ML models covered logS (molecular aqueous coefficient), logP (octanol/water partition coefficient), Pgp, caco-2 permeability, BBB penetration and PPB. The first two models were generated by super vector regression (SVR) techniques, Pgp by Random Forest, and the other three employing support vector machines (SVM). For training and testing the logS,

logP, caco-2 and Pgp models, ChEMBL (logS, logP, caco-2 and Pgp) and Huuskonen (logS), Poongavanam and co-workers (Pgp) and Sedykh et al. (Pgp) datasets were employed, and for BBB and PPB, the datasets described by Muehlbacher et al. and Zhu and co-workers [99–103]. The pkCSM webserver allows the prediction of PK properties based on (I) compound general properties (including molecular properties, toxicophores and pharmacophores) and (II) distance-based graph signatures. Given an input molecule, both sources of information are used to train and test machine learning-based predictors. The webserver is composed of 28 (not all employed in this work) regression and classification ML models that have been generated and trained against 30 datasets (described at Pires et al.) [66].

Graphical representations

Graphical representations of protein-ligand complexes were prepared using PyMOL version 1.7 and PLIP version 1.3.3 [104,105]. 2D ligand images were prepared using RDKit [106] python library and SMD plots using matplotlib [107] and seaborn [108] python libraries.

Conclusions

Kororamide A-B and convolutamine I-J can act as tau (GSK3 β and CK1 δ) and dual specificity (DYRK1A and CLK1) protein kinases inhibitors. Kororamide A-B are brominated indole alkaloids structurally very similar to meridianins. Only having this fact into account and following the SAR principle, a kororamides kinase inhibitory effect could be hypothesized, therefore the *in silico* binding results we obtained were expected. These results corroborate the idea of that kororamides could be kinase inhibitors with a therapeutic role in AD. Convolutamine I-J, which are not structurally similar to meridianins or kororamides, but are brominated heterocyclic compounds like other known kinase inhibitors, have also shown a plausible inhibitory capacity over GSK3 β , CK1 δ , DYRK1A and CLK1. Altogether, the results highlight the role of the indole scaffold and the halogen substituents on these kinases inhibition, being common features among all the compounds.

However, as happened with several other compounds acting over kinases, their main problem is the selectivity. These compounds seem to be somehow selective for one of the kinases, and it is clear which kinase is the preferred one to bind and which one is the belittled, but in general the obtained energy differences are not enough to consider that these compounds are selective. Moreover, the four brominated alkaloids should be optimized according to their ADMET properties. They have moderated good absorption properties, but caco-2 permeability could be increased, especially for kororamide A, as well as the distribution properties. Besides, the four compounds show a tendency to have toxicity

problems that should be carefully revised, in the same way as compounds interaction with cytochrome P450, although kororamide B does not show this cytochrome interaction.

Through the inclusion of convolutamine into the analysis (as they are brominated but not indole compounds), as well as the exploration of some indole-containing compounds from the MarinLit database, we intended to disentangle whether the indole or the halogen substituents presence is the most important feature to gain activity over the four kinases studied. However, the main conclusion extracted is that individually both are equally important, and probably the best way to profit both features is combining them into halogenated indole scaffolds.

Natural products possess a huge therapeutic potential, as reported here and in the mentioned literature. Within natural products, those of unexplored marine origin are of great interest in the discovery of novel chemical structures, since they harbour most of the biodiversity of the world (Montaser & Luesch, 2011; Blunt et al., 2018b). Life started in the oceans and many organisms live only there. Because of that, they should be successfully exploited in the future using sustainability criteria and respecting biodiversity. All this makes computational CADD contributions very relevant, since no biological sample is needed to perform an *in silico* analysis (Molinski et al., 2009b; Grosso et al., 2014; Kiuru et al., 2014; Martins et al., 2014). Having all these facts into account and taking profit from the scaffolds showed by meridianins and kororamides (examples of the importance of halogenated indole scaffolds to gain kinase inhibitory activity) we designed 49 marine natural products derivatives. Concretely, we performed a detailed computational study for the development of specific tau (GSK3 β and CK1 δ) and dual specificity (DYRK1A and CLK1) protein kinases inhibitors, starting from marine natural products, meridianin F and kororamide A, till the rational design of indole scaffolds derivatives as possible ATP competitive kinase inhibitors for the treatment of AD. We illustrated how the indole derivative compounds derived from scaffold 2 (an indole with an aromatic ring at R3 position and halogen substituents at R1 and R2) in general and compounds 2a and 2e in particular, could be proposed as good hit compounds to start a H2L optimization process. Altogether, it could be concluded that kororamides, specially A, convolutamines, specially J, and compounds 2a and 2e could be possible ATP competitive inhibitors against GSK3 β , CK1 δ , DYRK1A and CLK1. These results come from *in silico* experiments and further *in vitro* and *in vivo* studies are required. Our results constitute a promising starting point for the development of novel anti-AD drugs.

Acknowledgments: This research was partially supported by an Industrial Doctorate grant from the Generalitat of Catalonia to L.L.-P (DI 2016-051). We want to thank John W. Blunt and Murray H. G. Munro for their initial idea of building the MarinLit Database.

Author Contributions: M.S.-M., L.L.-P and C.A. conceived the study and designed the experiments. L.L.-P. carried out the experiments whereas A.N.-C., C.A. and M.S.-M. supervised them. All the authors analysed and discussed the results as well as wrote the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

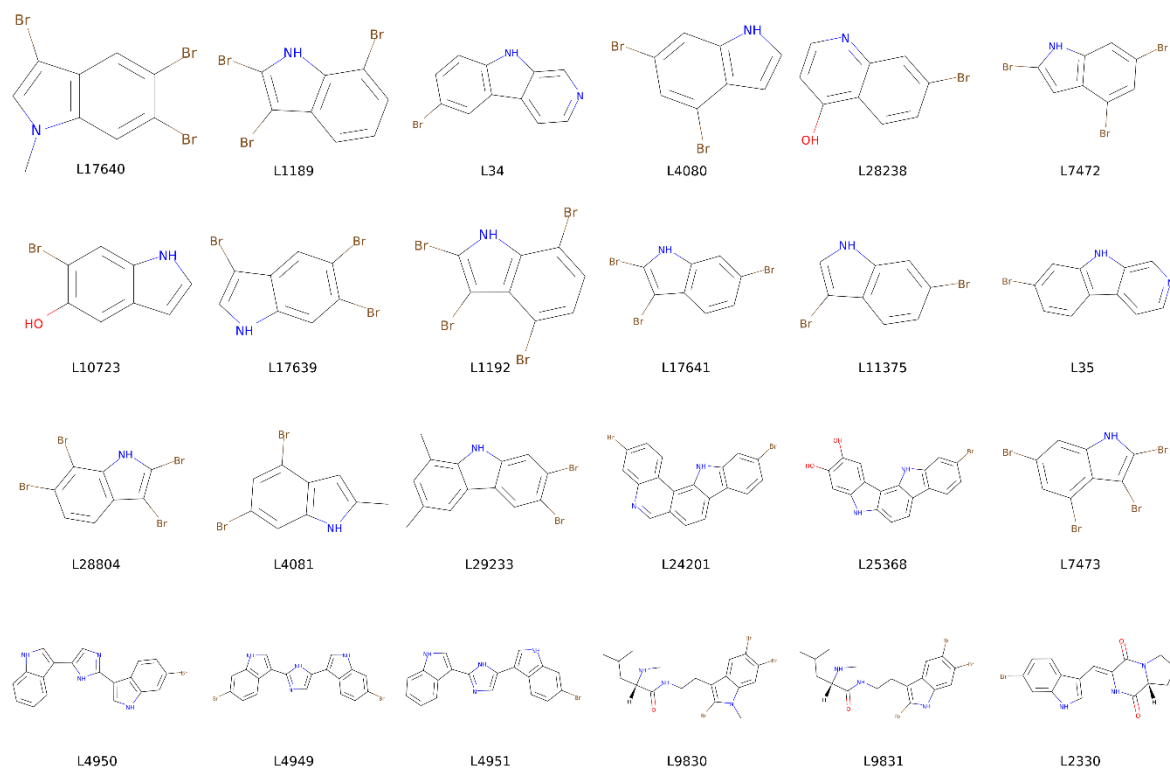


Figure A1. Structures of the 24 compounds found at MarinLit database after a similarity based substructure search using meridianin F, kororamide A, and the indole scaffold as a seed.

Table A1. Summary of classical rigid docking calculations of the derived analogues compound set over the GSK3 β and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the best poses analogues compounds. Lowercase letters represent the employed halogen group (a-g).

		GSK3 β							
		1	2	3	4	5	6	7	
Binding Energy	R0/R1	a	-6.9/-6.9	-8.1/-8.1	-5.9/-5.9	-6.5/-6.5	-6.3/-6.3	-6.2/-6.2	-6.6/-6.6
			MM/GBSA	-30.3141	-31.2458	-13.8779	-27.6481	-27.6534	-18.5779
Binding Energy	R0/R1	b	-6.8/-6.8	-7.9/-7.9	-5.1/-5.1	-6.6/-6.6	-5.6/-5.6	-5.8/-5.8	-5.9/-5.9
			MM/GBSA	-22.0902	-23.9910	-7.0321	-17.6371	-19.3959	-9.2248
Binding Energy	R0/R1	c	-6.8/-6.8	-8.1/-8.1	-5.8/-5.8	-6.3/-6.3	-6.3/-6.3	-6.2/-6.2	-6.7/-6.7
			MM/GBSA	-26.1345	-28.4927	-10.3167	-24.8857	-23.8207	-14.7674
Binding Energy	R0/R1	d	-7/-7	-8.1/-8.1	-5.2/-5.2	-6.8/-6.8	-6.5/-6.5	-6.1/-6.1	-6.7/-6.7
			MM/GBSA	-26.2805	-29.6158	-12.1225	-22.7717	-23.6754	-14.4347
Binding Energy	R0/R1	e	-7/-7	-8.1/-8.1	-6/-6	-6.7/-6.7	-5.8/-5.8	-6.1/-6.1	-6.7/-6.7
			MM/GBSA	-27.2898	-19.5564	-25.4497	-26.4593	-17.3314	-18.9577
Binding Energy	R0/R1	f	-6/-6	-8.2/-8.2	-5.8/-5.8	-6.3/-6.3	-6.2/-6.2	-6/-6	-6.4/-6.4
			MM/GBSA	-26.4517	-6.2475	-23.7315	-19.8909	-13.3104	
Binding Energy	R0/R1	g	-6.2/-6.2	-8.2/-8.2	-5.8/-5.8	-6.5/-6.5	-6.2/-6.2	-6.3/-6.3	-5.9/-5.9
			MM/GBSA	-28.2864	-14.2501	-24.4026	-24.7124	-16.8986	-20.7272

To avoid false positives, each docking calculation was performed twice (R0/R1). All energy values are kcal/mol.

Table A2. Summary of classical rigid docking calculations of the derived analogues compound set over the CK1 δ and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the best poses analogues compounds. Lowercase letters represent the employed halogen group (a-g).

			CK1 δ						
			1	2	3	4	5	6	7
Binding Energy	R0/R1	a	-5.1/-5.1	-7.7/-7.7	-5.2/-5.2	-5.3/-5.3	-5.6/-5.6	-5.6/-5.6	-5.3/-5.3
MM/GBSA			-35.4499		-3.7149	-26.5327		-26.4630	-18.4901
Binding Energy	R0/R1	b	-5.8/-5.8	-7.7/-7.7	-5.5/-5.5	-5.3/-5.3	-5.8/-5.8	-5.6/-5.6	-5.2/-5.2
MM/GBSA			-24.0479	-30.2266		-21.2435		-6.4142	-11.6429
Binding Energy	R0/R1	c	-5.6/-5.6	-7.6/-7.6	-4.8/-4.8	-5.7/-5.7	-5.2/-5.2	-5.1/-5.1	-5.6/-5.6
MM/GBSA			-29.9803		-19.4546	-22.8644	-26.4159	-12.0871	
Binding Energy	R0/R1	d	-5.9/-5.9	-7.5/-7.5	-4.7/-4.7	-5.5/-5.5	-5.9/-5.9	-5.1/-5.1	-5.5/-5.5
MM/GBSA					-19.8892	-25.5694			
Binding Energy	R0/R1	e	-6.1/-6.1	-7.5/-7.5	-5.4/-5.4	-5.3/-5.3	-5.1/-5.1	-5.3/-5.3	-5.3/-5.3
MM/GBSA				-37.8982		-28.6573	-28.5831	-16.2323	
Binding Energy	R0/R1	f	-6.2/-6.2	-7.5/-7.5	-5.4/-5.4	-5.2/-5.2	-5.1/-5.1	-5.5/-5.5	-5.4/-5.4
MM/GBSA				-34.6944		-22.6616	-26.3915	-13.6562	-15.4050
Binding Energy	R0/R1	g	-6.1/-6.1	-7.3/-7.3	-4.8/-4.8	-5.2/-5.2	-5.1/-5.1	-5.6/-5.6	-5.5/-5.5
MM/GBSA				-33.2393	-28.7631	-26.0731	-28.0238		

To avoid false positives, each docking calculation was performed twice (R0/R1). All energy values are kcal/mol.

Table A3. Summary of classical rigid docking calculations of the derived analogs compound set over the DYRK1A and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the best poses analogs compounds. Lowercase letters represent the employed halogen group (a-g).

			DYRK1A						
			1	2	3	4	5	6	7
Binding Energy	R0/R1	a	-6.2/-6.2	-8.6/-8.6	-5.9/-5.9	-6.8/-6.8	-6.7/-6.7	-5.7/-5.7	-6/-6
MM/GBSA				-37.8422	-15.2733	-30.7518	-31.2535	-18.9387	-20.8203
Binding Energy	R0/R1	b	-6.5/-6.5	-8.5/-8.5	-7/-7	-7.2/-7.2	-6.6/-6.6	-7/-7	-7.3/-7.3
MM/GBSA			-23.7829	-28.0642	-8.4887	-19.4730	-21.8802	-10.2503	-11.8981
Binding Energy	R0/R1	c	-6.3/-6.3	-8.6/-8.6	-5.9/-5.9	-7.2/-7.2	-6.5/-6.5	-6.9/-6.9	-7.3/-7.3
MM/GBSA			-30.2004	-34.1231	-12.1852	-26.5473	-28.3000	-18.7332	-13.2158
Binding Energy	R0/R1	d	-5.9/-5.9	-8.6/-8.6	-6.7/-6.7	-6.6/-6.6	-6.3/-6.3	-6.5/-6.5	-6.6/-6.6
MM/GBSA			-30.8597	-36.1125	-10.3667	-26.3748	-26.9602	-17.1653	-16.4510
Binding Energy	R0/R1	e	-6.4/-6.4	-8.6/-8.6	-6.5/-6.5	-6.6/-6.6	-6.4/-6.4	-6.6/-6.6	-6.4/-6.4
MM/GBSA			-32.8862		-13.5197	-28.9038	-29.6780		-20.0635
Binding Energy	R0/R1	f	-6.2/-6.2	-8.6/-8.6	-5.8/-5.8	-6.8/-6.8	-6.3/-6.3	-6.4/-6.4	-6.9/-6.9
MM/GBSA			-28.9419	-33.0823	-14.5027	-25.4579	-28.4081	-15.8583	-17.0209
Binding Energy	R0/R1	g	-6.7/-6.7	-8.7/-8.7	-5.8/-5.8	-6.8/-6.8	-6.7/-6.7	-6.1/-6.1	-6.9/-6.9
MM/GBSA			-30.3186	-35.5805	-13.4928	-27.1510	-29.6010	-18.0146	-17.3973

To avoid false positives, each docking calculation was performed twice (R0/R1). All energy values are kcal/mol.

Table A4. Summary of classical rigid docking calculations of the derived analogs compound set over the CLK1 and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the best poses analogs compounds. Lowercase letters represent the employed halogen group (a-g).

			CLK1						
			1	2	3	4	5	6	7
Binding Energy	R0/R1	a	-7.5/-7.5	-8.8/-8.8	-6.4/-6.4	-6.8/-6.8	-6.2/-6.2	-6.4/-6.4	-6.9/-6.9
MM/GBSA			-29.8546	-34.1041	-15.2943	-26.7584	-25.8120	-27.7489	-24.0832
Binding Energy	R0/R1	b	-7.1/-7.1	-8.4/-8.4	-7/-7	-5.6/-5.6	-7.5/-7.5	-6.6/-6.6	-7.2/-7.2
MM/GBSA			-25.9089	-27.3711	-13.7919	-26.2326	-22.1539	-16.6984	-20.9149
Binding Energy	R0/R1	c	-7.6/-7.6	-9.1/-9.1	-7/-7	-7.3/-7.3	-6.4/-6.4	-6.6/-6.6	-6.4/-6.4
MM/GBSA				-34.0221	-16.2165	-24.6708	-29.4190	-25.8368	-20.9436
Binding Energy	R0/R1	d	-7.6/-7.6	-8.4/-8.4	-6.8/-6.8	-7.8/-7.8	-6.3/-6.3	-6.9/-6.9	-7/-7
MM/GBSA			-26.9398	-30.7361		-25.6581	-25.1797	-19.3712	-17.5727
Binding Energy	R0/R1	e	-6.8/-6.8	-8.9/-8.9	-6.9/-6.9	-7.5/-7.5	-5.9/-5.9	-7.1/-7.1	-7/-7
MM/GBSA			-30.0891		-16.2097	-28.3695	-28.0697	-27.0478	-17.4985
Binding Energy	R0/R1	f	-7.5/-7.5	-8.6/-8.6	-6.7/-6.7	-7.5/-7.5	-6.2/-6.2	-6.4/-6.4	-6.7/-6.7
MM/GBSA				-28.1471	-20.4786	-23.9274	-27.3596	-15.4231	-21.2829
Binding Energy	R0/R1	g	-7/-7	-8.9/-8.9	-6/-6	-7.4/-7.4	-6.8/-6.8	-6.4/-6.4	-6.8/-6.8
MM/GBSA			-30.3541	-33.9082	-16.3122	-25.0002		-30.7737	-25.4765

To avoid false positives, each docking calculation was performed twice (R0/R1). All energy values are kcal/mol.

Table A5. Summary of absorption and distribution properties of the all set of 49 derived compounds and the four brominated alkaloids convolutamine I (I) and J (J), kororamide A (A) and B (B).

Compound	Absorption					Distribution				
	Mol weight	LogS	P-Glycoprotein	Caco-2 permeability	Intestinal absorption	LogP	BBB	PPB	VDss	CNS permeability
Compound 1a	331	-4.8	inactive	High	92.328	2.8	0.298	High	0.231	-1.832
Compound 1b	209.2	-3.4	inactive	Moderate	94.522	2.6	0.186	Medium	0.131	-1.888
Compound 1c	242.1	-4.3	inactive	High	92.462	2.9	0.279	High	0.37	-1.832
Compound 1d	270.1	-4	inactive	High	93.463	2.6	0.152	High	0.249	-1.866
Compound 1e	286.6	-4.5	inactive	High	92.395	2.8	0.278	High	0.385	-1.832
Compound 1f	270.1	-4	inactive	High	93.49	2.6	0.152	High	0.256	-1.87
Compound 1g	286.6	-4.5	inactive	High	92.395	2.8	0.278	High	0.385	-1.832
Compound 2a	351	-6.1	inactive	Moderate	90.067	4.1	0.477	High	0.234	-0.894
Compound 2b	229.2	-5.1	inactive	Moderate	92.006	3.4	0.539	High	-0.081	-0.946
Compound 2c	262.1	-5.9	inactive	Moderate	90.201	4.6	0.482	High	0.197	-0.894
Compound 2d	306.6	-6.1	inactive	Moderate	90.134	4.4	0.48	High	0.215	-0.894
Compound 2e	290.1	-5.7	inactive	Moderate	91.036	3.8	0.508	High	0.076	-0.92
Compound 2f	290.1	-5.7	inactive	Moderate	91.721	3.8	0.708	High	0.051	-1.339
Compound 2g	306.6	-6.1	inactive	Moderate	90.819	4.4	0.68	High	0.196	-1.313
Compound 3a	304	-4.7	inactive	High	89.848	2.8	0.227	High	0.95	-1.961
Compound 3b	182.2	-3.4	inactive	Moderate	91.82	2.7	0.375	Low	0.764	-2.017
Compound 3c	215.1	-4.2	inactive	High	89.982	2.9	0.23	High	0.919	-1.961
Compound 3d	243.1	-4	inactive	High	90.899	2.7	0.222	Medium	0.868	-1.999
Compound 3e	259.5	-4.4	inactive	High	89.915	2.8	0.228	High	0.934	-1.961
Compound 3f	243.1	-4	inactive	High	90.872	2.7	0.222	Medium	0.845	-1.995
Compound 3g	259.5	-4.4	inactive	High	89.915	2.8	0.228	High	0.934	-1.961
Compound 4a	289	-4.9	inactive	High	91.487	3.3	0.351	High	0.432	-1.66
Compound 4b	167.2	-3.4	inactive	Moderate	93.459	3.2	0.437	Medium	0.248	-1.715
Compound 4c	200.1	-4.5	inactive	High	91.621	3.6	0.357	High	0.401	-1.66
Compound 4d	228.1	-4.1	inactive	High	92.538	3.2	0.382	Medium	0.344	-1.697
Compound 4e	244.5	-4.7	inactive	High	91.554	3.5	0.354	High	0.416	-1.66
Compound 4f	228.1	-4.1	inactive	High	92.511	3.2	0.382	Medium	0.32	-1.693
Compound 4g	244.5	-4.7	inactive	High	91.554	3.5	0.354	High	0.416	-1.66
Compound 5a	305	-4.4	inactive	High	89.763	3	0.284	High	0.253	-1.98
Compound 5b	183.2	-2.9	inactive	Moderate	91.734	2.6	0.432	Low	0.086	-2.036
Compound 5c	216.1	-3.6	inactive	High	89.897	2.7	0.287	High	0.223	-1.98
Compound 5d	244.1	-3.5	inactive	High	90.814	2.7	0.279	Medium	0.169	-2.017
Compound 5e	260.5	-3.9	inactive	High	89.83	2.9	0.286	High	0.238	-1.98
Compound 5f	244.1	-3.5	inactive	High	90.786	2.7	0.279	Medium	0.143	-2.014
Compound 5g	260.5	-3.9	inactive	High	89.83	2.9	0.286	High	0.238	-1.98
Compound 6a	318	-4.7	inactive	High	90.757	2.7	0.146	High	1.061	-1.917
Compound 6b	196.2	-3.2	inactive	Moderate	92.728	2.7	0.293	Low	0.867	-1.973
Compound 6c	229.1	-4.3	inactive	High	90.891	3	0.148	Medium	1.031	-1.917
Compound 6d	257.1	-3.9	inactive	High	91.78	2.6	0.14	Medium	0.956	-1.951
Compound 6e	273.6	-4.4	inactive	High	90.824	2.8	0.147	Medium	1.046	-1.917
Compound 6f	257.1	-3.9	inactive	High	91.808	2.6	0.14	Medium	0.978	-1.954
Compound 6g	273.6	-4.4	inactive	High	90.824	2.8	0.147	Medium	1.046	-1.917
Compound 7a	332	-4.9	inactive	High	92.246	3	0.191	High	1.141	-1.487
Compound 7b	210.2	-3.4	inactive	Moderate	94.218	2.9	0.349	Low	0.994	-1.543
Compound 7c	243.1	-4.6	inactive	High	92.38	3.3	0.225	Medium	1.11	-1.487
Compound 7d	271.1	-4.1	inactive	High	93.297	2.9	0.258	Medium	1.084	-1.525
Compound 7e	287.6	-4.7	inactive	High	92.313	3.2	0.208	High	1.125	-1.487
Compound 7f	271.1	-4.1	inactive	High	93.27	2.9	0.258	Medium	1.055	-1.521
Compound 7g	287.6	-4.7	inactive	High	92.313	3.2	0.208	High	1.125	-1.487
J	470	-4.4	inactive	Moderate	90,483	4.4	0,386	High	0,868	-2,215
I	473	-4.3	inactive	Moderate	91,515	3.9	0,193	High	1,474	-2,024
A	534.1	-4.3	inactive	Low	90,979	4.6	0,316	Low	1,112	-2,449
B	535.1	-3.9	inactive	Moderate	100	3.4	0,184	High	0,002	-2,93

BBB: blood brain barrier, PPB: protein-protein binding, VDss: steady state volume of distribution, CNS: central nervous system.

Table A6. Summary of metabolism, excretion and toxicity properties of the all set of 49 derived compounds and the four brominated alkaloids convolutamine I (I) and J (J), kororamide A (A) and B (B).

Compound	Metabolism	Excretion	Toxicity			
	CYP450	OCT2 substrate	hERG	MRTD	AMES toxicity	Hepatotoxicity
Compound 1a	Yes	No	<4.0	0.482	No	No
Compound 1b	Yes	No	<4.0	0.666	No	No
Compound 1c	Yes	No	<4.0	0.503	No	No
Compound 1d	Yes	No	<4.0	0.45	No	No
Compound 1e	Yes	No	<4.0	0.492	No	No
Compound 1f	Yes	No	<4.0	0.574	No	No
Compound 1g	Yes	No	<4.0	0.492	No	No
Compound 2a	Yes	No	<4.0	0.673	Yes	No
Compound 2b	Yes	No	<4.0	0.608	Yes	No
Compound 2c	Yes	No	<4.0	0.671	Yes	No
Compound 2d	Yes	No	<4.0	0.672	Yes	No
Compound 2e	Yes	No	<4.0	0.641	Yes	No
Compound 2f	Yes	No	<4.0	0.585	Yes	No
Compound 2g	Yes	No	<4.0	0.616	Yes	No
Compound 3a	Yes	No	<4.0	0.381	No	No
Compound 3b	Yes	No	<4.0	0.512	No	No
Compound 3c	Yes	No	<4.0	0.402	No	No
Compound 3d	Yes	No	<4.0	0.455	No	No
Compound 3e	Yes	No	<4.0	0.391	No	No
Compound 3f	Yes	No	<4.0	0.303	No	No
Compound 3g	Yes	No	<4.0	0.391	No	No
Compound 4a	Yes	No	<4.0	0.525	No	No
Compound 4b	Yes	No	<4.0	0.716	No	No
Compound 4c	Yes	No	<4.0	0.544	No	No
Compound 4d	Yes	No	<4.0	0.625	No	No
Compound 4e	Yes	No	<4.0	0.534	No	No
Compound 4f	Yes	No	<4.0	0.471	No	No
Compound 4g	Yes	No	<4.0	0.534	No	No
Compound 5a	Yes	No	<4.0	0.55	No	No
Compound 5b	Yes	No	<4.0	0.678	No	No
Compound 5c	Yes	No	<4.0	0.572	No	No
Compound 5d	Yes	No	<4.0	0.627	No	No
Compound 5e	Yes	No	<4.0	0.561	No	No
Compound 5f	Yes	No	<4.0	0.47	No	No
Compound 5g	Yes	No	<4.0	0.561	No	No
Compound 6a	Yes	No	<4.0	0.376	No	No
Compound 6b	Yes	No	<4.0	0.502	No	No
Compound 6c	Yes	No	<4.0	0.397	No	No
Compound 6d	Yes	No	<4.0	0.293	No	No
Compound 6e	Yes	No	<4.0	0.387	No	No
Compound 6f	Yes	No	<4.0	0.441	No	No
Compound 6g	Yes	No	<4.0	0.387	No	No
Compound 7a	Yes	No	<4.0	0.2	No	No
Compound 7b	Yes	No	<4.0	0.311	No	No
Compound 7c	Yes	No	<4.0	0.219	No	No
Compound 7d	Yes	No	<4.0	0.259	No	No
Compound 7e	Yes	No	<4.0	0.209	No	No
Compound 7f	Yes	No	<4.0	0.117	No	No
Compound 7g	Yes	No	<4.0	0.209	No	No
I	Yes	Yes	<4.0	0,029	No	Yes
J	Yes	Yes	<4.0	-0,814	No	Yes
A	Yes	No	<4.0	-0,599	No	Yes
B	No	No	<4.0	0,405	Yes	No

CYP: cytochrome, OCT2: organic cation transporter 2, hERG: human ether-a-go-go gene, MRTD: maximum recommended tolerated dose.

References

1. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* 2002, 298, 1912–34, doi:10.1126/science.1075762.
2. Kolarova, M.; García-Sierra, F.; Bartos, A.; Ricny, J.; Ripova, D.; Ripova, D. Structure and Pathology of Tau Protein in Alzheimer Disease. *Int. J. Alzheimers. Dis.* 2012, 2012, 1–13, doi:10.1155/2012/731526.
3. Martin, L.; Latypova, X.; Wilson, C. M.; Magnaudeix, A.; Perrin, M.-L.; Yardin, C.; Terro, F. Tau protein kinases: Involvement in Alzheimer's disease. *Ageing Res. Rev.* 2013, 12, 289–309, doi:10.1016/j.arr.2012.06.003.
4. Citron, M. Alzheimer's disease: strategies for disease modification. *Nat. Rev. Drug Discov.* 2010, 9, 387–398, doi:10.1038/nrd2896.
5. Hernandez, F.; Lucas, J. J.; Avila, J. GSK3 and Tau: Two Convergence Points in Alzheimer's Disease. *J. Alzheimer's Dis.* 2012, 33, S141–S144, doi:10.3233/JAD-2012-129025.
6. Ishizawa, T.; Sahara, Narahiko Ishizawa, T.; Sahara, N.; Ishiguro, K.; Kersh, J.; MCGOWAN, E.; Lewis, J.; Hutton, M.; Dickson, D. W.; Yen, S.-H. C.-L. of G. S. K.-3 with N. T. and G. D. in T. M. 2003; V. 163.; Ishiguro, K.; Kersh, J.; MCGOWAN, E.; Lewis, J.; Hutton, M.; Dickson, D. W.; Yen, S.-H. Co-Localization of Glycogen Synthase Kinase-3 with Neurofibrillary Tangles and Granulovacuolar Degeneration in Transgenic Mice; 2003; Vol. 163;.
7. Ryoo, S.-R.; Kyeong Jeong, H.; Radnaabazar, C.; Yoo, J.-J.; Cho, H.-J.; Lee, H.-W.; Kim, I.-S.; Cheon, Y.-H.; Ahn, Y. S.; Chung, S.-H.; Song, W.-J. DYRK1A-mediated Hyperphosphorylation of Tau A functional link between down syndrome and alzheimer disease *. 2007, doi:10.1074/jbc.M707358200.
8. Tell, V.; Hilgeroth, A. Recent developments of protein kinase inhibitors as potential AD therapeutics. *Front Cell Neurosci* 2013, 1118, 5–1, doi:10.3389/fncel.2013.00189.
9. Dolan, P. J.; Johnson, G. V. W. The role of tau kinases in Alzheimer's disease. *Curr. Opin. Drug Discov. Devel.* 2010, 13, 595–603.
10. Stotani, S.; Giordanetto, F.; Medda, F. DYRK1A inhibition as potential treatment for Alzheimer's disease. *Future Med. Chem.* 2016, 8, 681–696, doi:10.4155/fmc-2016-0013.
11. Branca, C.; Shaw, D. M.; Belfiore, R.; Gokhale, V.; Shaw, A. Y.; Foley, C.; Smith, B.; Hulme, C.; Dunckley, T.; Meechoo, B.; Caccamo, A.; Oddo, S. Dyrk1 inhibition improves Alzheimer's disease-like pathology. *Aging Cell* 2017, 16, 1146–1154, doi:10.1111/acer.12648.
12. Li, G.; Yin, H.; Kuret, J. Casein kinase 1 delta phosphorylates tau and disrupts its binding to microtubules. *J. Biol. Chem.* 2004, 279, 15938–45, doi:10.1074/jbc.M314116200.
13. Hooper, C.; Killick, R.; Lovestone, S. The GSK3 hypothesis of Alzheimer's disease. *J. Neurochem.* 2008, 104, 1433–9, doi:10.1111/j.1471-4159.2007.05194.x.
14. Llorens-Martín, M.; Jurado, J.; Hernández, F.; Avila, J. GSK-3 β , a pivotal kinase in Alzheimer disease. *Front. Mol. Neurosci.* 2014, 7, 46, doi:10.3389/fnmol.2014.00046.
15. Hernández, F.; Gómez de Barreda, E.; Fuster-Matanzo, A.; Lucas, J. J.; Avila, J. GSK3: A possible link between beta amyloid peptide and tau protein. *Exp. Neurol.* 2010, 223, 322–325, doi:10.1016/j.expneurol.2009.09.011.
16. Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on natural products for drug design. *Nat. Chem.* 2016, 8, 531–541, doi:10.1038/nchem.2479.
17. W-H Li, J.; Vederas, J. C. Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science* (80-.). 2009, 325, 161–165.
18. Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* 2015, 14, 111–129, doi:10.1038/nrd4510.

19. Patridge, E.; Gareiss, P.; Kinch, M. S.; Hoyer, D. An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov. Today* 2016, 21, 204–207, doi:10.1016/J.DRUDIS.2015.01.009.
20. Bharate, S. S.; Mignani, S.; Vishwakarma, R. A. Why Are the Majority of Active Compounds in the CNS Domain Natural Products? A Critical Analysis. *J. Med. Chem.* 2018, acs.jmedchem.7b01922, doi:10.1021/acs.jmedchem.7b01922.
21. Russo, P.; Kisialiou, A.; Lamonaca, P.; Moroni, R.; Prinzi, G.; Fini, M. New Drugs from Marine Organisms in Alzheimer's Disease. *Mar. Drugs* 2015, 14, 5, doi:10.3390/md14010005.
22. Ansari, N.; Khodagholi, F. Natural Products as Promising Drug Candidates for the Treatment of Alzheimer's Disease: Molecular Mechanism Aspect. *Curr. Neuropharmacol.* 2013, 11, 414–429, doi:10.2174/1570159X11311040005.
23. Dey, A.; Bhattacharya, R.; Mukherjee, A.; Pandey, D. K. Natural products against Alzheimer's disease: Pharmaco-therapeutics and biotechnological interventions. *Biotechnol. Adv.* 2017, 35, 178–216, doi:10.1016/J.BIOTECHADV.2016.12.005.
24. B. Bharate, S.; R. Yadav, R.; Battula, S.; A. Vishwakarma, R. Meridianins: Marine-Derived Potent Kinase Inhibitors. *Mini-Reviews Med. Chem.* 2012, 12, 618–631, doi:10.2174/138955712800626728.
25. Meine, R.; Becker, W.; Falke, H.; Preu, L.; Loaëc, N.; Meijer, L.; Kunick, C. Indole-3-Carbonitriles as DYRK1A Inhibitors by Fragment-Based Drug Design. *Molecules* 2018, 23, 64, doi:10.3390/molecules23020064.
26. Tahtouh, T.; Elkins, J. M.; Filippakopoulos, P.; Soundararajan, M.; Burgy, G.; Durieu, E.; Cochet, C.; Schmid, R. S.; Lo, D. C.; Delhommel, F.; Oberholzer, A. E.; Pearl, L. H.; Carreaux, F.; Bazureau, J.-P.; Knapp, S.; Meijer, L. Selectivity, Cocrystal Structures, and Neuroprotective Properties of Leucettines, a Family of Protein Kinase Inhibitors Derived from the Marine Sponge Alkaloid Leucettamine B. *J. Med. Chem.* 2012, 55, 9312–9330, doi:10.1021/jm301034u.
27. Salado, I. G.; Redondo, M.; Bello, M. L.; Perez, C.; Liachko, N. F.; Kraemer, B. C.; Miguel, L.; Lecourtois, M.; Gil, C.; Martinez, A.; Perez, D. I. Protein Kinase CK-1 Inhibitors As New Potential Drugs for Amyotrophic Lateral Sclerosis. *J. Med. Chem.* 2014, 57, 2755–2772, doi:10.1021/jm500065f.
28. Eldar-Finkelman, H.; Martinez, A. GSK-3 Inhibitors: Preclinical and Clinical Focus on CNS. *Front. Mol. Neurosci.* 2011, 4, 32, doi:10.3389/fnmol.2011.00032.
29. Giraud, F.; Alves, G.; Debiton, E.; Nauton, L.; Th Ery, V.; Durieu, E.; Ferandin, Y.; Lozach, O.; Meijer, L.; Anizon, F.; Pereira, E.; Moreau, P. Synthesis, Protein Kinase Inhibitory Potencies, and *in vitro* Antiproliferative Activities of Meridianin Derivatives. *J. Med. Chem.* 2011, 54, 4474–4489, doi:10.1021/jm200464w.
30. Jain, P.; Karthikeyan, C.; Moorthy, N. S.; Waiker, D.; Jain, A.; Trivedi, P. Human CDC2-Like Kinase 1 (CLK1): A Novel Target for Alzheimer's Disease. *Curr. Drug Targets* 2014, 15, 539–550, doi:10.2174/1389450115666140226112321.
31. Fedorov, O.; Huber, K.; Eisenreich, A.; Filippakopoulos, P.; King, O.; Bullock, A. N.; Szklarczyk, D.; Jensen, L. J.; Fabbro, D.; Trappe, J.; Rauch, U.; Bracher, F.; Knapp, S. Specific CLK Inhibitors from a Novel Chemotype for Regulation of Alternative Splicing. *Chem. Biol.* 2011, 18, 67–76, doi:10.1016/j.chembiol.2010.11.009.
32. Halekotte, J.; Witt, L.; Ianes, C.; Krüger, M.; Bührmann, M.; Rauh, D.; Pichlo, C.; Brunstein, E.; Luxenburger, A.; Baumann, U.; Knippschild, U.; Bischof, J.; Peifer, C. Optimized 4,5-Diarylimidazoles as Potent/Selective Inhibitors of Protein Kinase CK1 δ and Their Structural Relation to p38 α MAPK. *Molecules* 2017, 22, 522, doi:10.3390/molecules22040522.
33. Cozza, G.; Gianoncelli, A.; Montopoli, M.; Caparrotta, L.; Venerando, A.; Meggio, F.; Pinna, L. A.; Zagotto, G.; Moro, S. Identification of novel protein kinase CK1 delta (CK1 δ) inhibitors through

- structure-based virtual screening. *Bioorg. Med. Chem. Lett.* 2008, 18, 5672–5675, doi:10.1016/J.BMCL.2008.08.072.
34. Yadav, R. R.; Sharma, S.; Joshi, P.; Wani, A.; Vishwakarma, R. A.; Kumar, A.; Bharate, S. B. Meridianin derivatives as potent Dyrk1A inhibitors and neuroprotective agents. *Bioorg. Med. Chem. Lett.* 2015, 25, 2948–2952, doi:10.1016/j.bmcl.2015.05.034.
 35. Feng, L.; Geisselbrecht, Y.; Blanck, S.; Wilbuer, A.; Atilla-Gokcumen, G. E.; Filippakopoulos, P.; Kräling, K.; Celik, M. A.; Harms, K.; Maksimoska, J.; Marmorstein, R.; Frenking, G.; Knapp, S.; Essen, L.-O.; Meggers, E. Structurally sophisticated octahedral metal complexes as highly selective protein kinase inhibitors. *J. Am. Chem. Soc.* 2011, 133, 5976–86, doi:10.1021/ja1112996.
 36. Mente, S.; Arnold, E.; Butler, T.; Chakrapani, S.; Chandrasekaran, R.; Cherry, K.; Dirico, K.; Doran, A.; Fisher, K.; Galatsis, P.; Green, M.; Hayward, M.; Humphrey, J.; Knafels, J.; Li, J.; Liu, S.; Marconi, M.; McDonald, S.; Ohren, J.; Paradis, V.; Sneed, B.; Walton, K.; Wager, T. Ligand-protein interactions of selective casein kinase 1 delta inhibitors. *J. Med. Chem.* 2013, doi:10.1021/jm4006324.
 37. Llorach-Pares, L.; Nonell-Canals, A.; Sanchez-Martinez, M.; Avila, C. Computer-Aided Drug Design Applied to Marine Drug Discovery: Meridianins as Alzheimer's Disease Therapeutic Agents. *Mar. Drugs* 2017, 15, 366, doi:10.3390/md15120366.
 38. Dashti, Y.; Vial, M.-L.; Wood, S. A.; Mellick, G. D.; Roullier, C.; Quinn, R. J. Kororamide B, a brominated alkaloid from the bryozoan *Amathia tortuosa* and its effects on Parkinson's disease cells. *Tetrahedron* 2015, 71, 7879–7884, doi:10.1016/J.TET.2015.08.017.
 39. Netz, N.; Opatz, T. Marine Indole Alkaloids. *Mar. Drugs* 2015, 13, 4814–4914, doi:10.3390/md13084814.
 40. Blunt, J. W.; Carroll, A. R.; Copp, B. R.; Davis, R. A.; Keyzers, R. A.; Prinsep, M. R. Marine natural products. *Nat. Prod. Rep.* 2018, 35, 8–53, doi:10.1039/c7np00052a.
 41. Gul, W.; Hamann, M. T. Indole alkaloid marine natural products: An established source of cancer drug leads with considerable promise for the control of parasitic, neurological and other diseases. *Life Sci.* 2005, 78, 442–453, doi:10.1016/J.LFS.2005.09.007.
 42. Gribble, G. W. Biological Activity of Recently Discovered Halogenated Marine Natural Products. *Mar. Drugs* 2015, 13, 4044–136, doi:10.3390/md13074044.
 43. Pauletti, P. M.; Cintra, L. S.; Braguine, C. G.; da Silva Filho, A. A.; Silva, M. L. A. E.; Cunha, W. R.; Januário, A. H. Halogenated indole alkaloids from marine invertebrates. *Mar. Drugs* 2010, 8, 1526–49, doi:10.3390/md8051526.
 44. Munro, M. H. G.; Blunt, J. W. *MarinLit*.
 45. Acharya, C.; Coop, A.; Polli, J. E.; Mackerell, A. D.; Jr. Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr. Comput. Aided. Drug Des.* 2011, 7, 10–22, doi:10.2174/157340911793743547.
 46. Traxler, P.; Furet, P. Strategies toward the Design of Novel and Selective Protein Tyrosine Kinase Inhibitors. *Pharmacol. Ther.* 1999, 82, 195–206, doi:10.1016/S0163-7258(98)00044-8.
 47. McGregor, M. J. A Pharmacophore Map of Small Molecule Protein Kinase Inhibitors. *J. Chem. Inf. Model.* 2007, 47, 2374–2382, doi:10.1021/ci700244t.
 48. Huang, D.; Zhou, T.; Lafleur, K.; Nevado, C.; Caflich, A. Kinase selectivity potential for inhibitors targeting the ATP binding site: a network analysis. *Bioinformatics* 2010, 26, 198–204, doi:10.1093/bioinformatics/btp650.
 49. Jarhad, D. B.; Mashelkar, K. K.; Kim, H.-R.; Noh, M.; Jeong, L. S. Dual-Specificity Tyrosine Phosphorylation-Regulated Kinase 1A (DYRK1A) Inhibitors as Potential Therapeutics. *J. Med. Chem.* 2018, acs.jmedchem.8b00185, doi:10.1021/acs.jmedchem.8b00185.

50. Klein-Junior, L.; Santos Passos, C.; Moraes, A.; Wakui, V.; Konrath, E.; Nurisso, A.; Carrupt, P.-A.; Alves de Oliveira, C.; Kato, L.; Henriques, A. Indole Alkaloids and Semisynthetic Indole Derivatives as Multifunctional Scaffolds Aiming the Inhibition of Enzymes Related to Neurodegenerative Diseases – A Focus on Psychotria L. Genus. *Curr. Top. Med. Chem.* 2014, 14, 1056–1075, doi:10.2174/1568026614666140324142409.
51. Ren, P.; Liu, Y.; Li, L.; Chan, K.; Wilson, T. E. Heterocyclic kinase inhibitors 2008.
52. Almstetter, M. T.; Trembl, A.; Koestler, R.; Yehia, N. Heterocyclic compounds as kinase inhibitors 2011.
53. Tremblay, J. L.; Kutok, D. G.; Winkler, V. J.; Palombella, Alfredo C. Castro, C. A.; Evans, Somarajannair Janardanannair, A. L. T.; Liu, M. R. Heterocyclic compounds and uses thereof 2014.
54. An, W. F.; Germain, A. R.; Bishop, J. A.; Nag, P. P.; Metkar, S.; Ketterman, J.; Walk, M.; Weiwer, M.; Liu, X.; Patnaik, D.; Zhang, Y.-L.; Gale, J.; Zhao, W.; Kaya, T.; Barker, D.; Wagner, F. F.; Holson, E. B.; Dandapani, S.; Perez, J.; Munoz, B.; Palmer, M.; Pan, J. Q.; Haggarty, S. J.; Schreiber, S. L. Discovery of Potent and Highly Selective Inhibitors of GSK3 β ; National Center for Biotechnology Information (US), 2010;
55. Benek, O.; Hroch, L.; Aitken, L.; Gunn-Moore, F.; Vinklarova, L.; Kuca, K.; Perez, D. I.; Perez, C.; Martinez, A.; Fisar, Z.; Musilek, K. 1-(Benzo[d]thiazol-2-yl)-3-phenylureas as dual inhibitors of casein kinase 1 and ABAD enzymes for treatment of neurodegenerative disorders. *J. Enzyme Inhib. Med. Chem.* 2018, 33, 665–670, doi:10.1080/14756366.2018.1445736.
56. Coombs, T. C.; Tanega, C.; Shen, M.; Wang, J. L.; Auld, D. S.; Gerritz, S. W.; Schoenen, F. J.; Thomas, C. J.; Aubé, J. Small-molecule pyrimidine inhibitors of the cdc2-like (Clk) and dual specificity tyrosine phosphorylation-regulated (Dyrk) kinases: development of chemical probe ML315. *Bioorg. Med. Chem. Lett.* 2013, 23, 3654–61, doi:10.1016/j.bmcl.2013.02.096.
57. Fenical, W. Chapter 12 Natural Halogenated Organics. Elsevier Oceanogr. Ser. 1981, 31, 375–393, doi:10.1016/S0422-9894(08)70334-9.
58. Hernandez, M.; Cavalcanti, S. M.; Moreira, D. R.; de Azevedo Junior, W.; Leite, A. C. Halogen Atoms in the Modern Medicinal Chemistry: Hints for the Drug Design. *Curr. Drug Targets* 2010, 11, 303–314, doi:10.2174/138945010790711996.
59. Hernandez, M.; Cavalcanti, S. M.; Moreira, D. R.; de Azevedo Junior, W.; Leite, A. C. Halogen Atoms in the Modern Medicinal Chemistry: Hints for the Drug Design. *Curr. Drug Targets* 2010, 11, 303–314, doi:10.2174/138945010790711996.
60. Metrangolo, P.; Resnati, G. Chemistry. Halogen versus hydrogen. *Science* 2008, 321, 918–9, doi:10.1126/science.1162215.
61. Kramer, T.; Schmidt, B.; Lo Monte, F. Small-Molecule Inhibitors of GSK-3: Structural Insights and Their Application to Alzheimer’s Disease Models. *Int. J. Alzheimers. Dis.* 2012, 2012, 381029, doi:10.1155/2012/381029.
62. Wei, W.; Bodles-Brakhop, A. M.; Barger, S. W. A Role for P-Glycoprotein in Clearance of Alzheimer Amyloid β - Peptide from the Brain HHS Public Access. *Curr Alzheimer Res* 2016, 13, 615–620.
63. Miller, D. S.; Bauer, B.; Hartz, A. M. S. Modulation of P-Glycoprotein at the Blood-Brain Barrier: Opportunities to Improve Central Nervous System Pharmacotherapy. *Pharmacol. Rev.* 2008, 60, 196–209, doi:10.1124/pr.107.07109.
64. Chang, K. L.; Pee, H. N.; Yang, S.; Ho, P. C. Influence of drug transporters and stereoselectivity on the brain penetration of pioglitazone as a potential medicine against Alzheimer’s disease. *Sci. Rep.* 2015, 5, 9000, doi:10.1038/srep09000.
65. Cirrito, J. R.; Deane, R.; Fagan, A. M.; Spinner, M. L.; Parsadanian, M.; Finn, M. B.; Jiang, H.; Prior, J. L.; Sagare, A.; Bales, K. R.; Paul, S. M.; Zlokovic, B. V.; Piwnicka-Worms, D.; Holtzman, D. M. P-

- glycoprotein deficiency at the blood-brain barrier increases amyloid- deposition in an Alzheimer disease mouse model. *J. Clin. Invest.* 2005, 115, 3285–3290, doi:10.1172/JCI25247.
66. Pires, D. E. V.; Blundell, T. L.; Ascher, D. B. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *J. Med. Chem.* 2015, 58, 4066–4072, doi:10.1021/acs.jmedchem.5b00104.
 67. Lynch, T.; Price, A. The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. *Am. Fam. Physician* 2007, 76, 391–396, doi:10.1046/j.1365-2125.1999.00073.x.
 68. Bibi, Z. Role of cytochrome P450 in drug interactions. *Nutr. Metab. (Lond)*. 2008, 5, 27, doi:10.1186/1743-7075-5-27.
 69. Yang, H.; Sun, L.; Li, W.; Liu, G.; Tang, Y. *In silico* Prediction of Chemical Toxicity for Drug Design Using Machine Learning Methods and Structural Alerts. *Front. Chem.* 2018, 6, 1–12, doi:10.3389/fchem.2018.00030.
 70. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 2012, 64, 4–17, doi:10.1016/j.addr.2012.09.019.
 71. Núñez-Pons, L.; Carbone, M.; Vázquez, J.; Rodríguez, J.; Nieto, R. M.; Varela, M. M.; Gavagnin, M.; Avila, C. Natural Products from Antarctic Colonial Ascidians of the Genera *Aplidium* and *Synoicum*: Variability and Defensive Role. *Mar. Drugs* 2012, 10, 1741–1764, doi:10.3390/md10081741.
 72. Berman, H.; Henrick, K.; Nakamura, H.; Markley, J. L. The worldwide Protein Data Bank (wwPDB): Ensuring a single, uniform archive of PDB data. *Nucleic Acids Res.* 2007, 35, 2006–2008, doi:10.1093/nar/gkl971.
 73. Felix, E.; Santamaría-Navarro, E.; Sanchez-Martinez, M.; Nonell-Canals, A. Itzamna.
 74. Yu, W.; Mackerell, A. D. Computer-Aided Drug Design Methods. *Methods Mol. Biol.* 2017, 1520, 93–94, doi:10.1007/978-1-4939-6634-9_5.
 75. Ganesan, A.; Coote, M. L.; Barakat, K. Molecular dynamics-driven drug discovery: leaping forward with confidence. *Drug Discov. Today* 2017, 22, doi:10.1016/j.drudis.2016.11.001.
 76. Lill, M. A. Efficient incorporation of protein flexibility and dynamics into molecular docking simulations. *Biochemistry* 2011, 50, 6157–6169, doi:10.1021/bi2004558.Efficient.
 77. Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* 2006, 26, 531–568, doi:10.1002/med.20067.
 78. de Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. The Role of Molecular Dynamics and Related Methods in Drug Discovery The Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* 2016, doi:10.1021/acs.jmedchem.5b01684.
 79. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 2015, 0441, 1–13, doi:10.1517/17460441.2015.1032936.
 80. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 2005, 26, 1781–1802, doi:10.1002/jcc.20289.
 81. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 2010, 78, 1950–8, doi:10.1002/prot.22711.
 82. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* 2004, 25, 1157–1174, doi:10.1002/jcc.20035.
 83. Martín-García, F.; Papaleo, E.; Gomez-Puertas, P.; Boomsma, W.; Lindorff-Larsen, K. Comparing Molecular Dynamics Force Fields in the Essential Subspace. *PLoS One* 2015, 10, e0121114, doi:10.1371/journal.pone.0121114.

84. Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Systematic validation of protein force fields against experimental data. *PLoS One* 2012, 7, 1–6, doi:10.1371/journal.pone.0032131.
85. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Antechamber, An Accessory Software Package For Molecular Mechanical Calculations. *Natl. Institutes Heal.*
86. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* 2005, 26, 1668–88, doi:10.1002/jcc.20290.
87. Jorgensen, W. L.; Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *J. Comput. Chem.* 1998, 19, 1179–1186, doi:10.1002/(SICI)1096-987X(19980730)19:10<1179::AID-JCC6>3.0.CO;2-J.
88. Andersen, H. C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* 1983, 52, 24–34.
89. Rastelli, G.; Degliesposti, G.; Del Rio, A.; Sgobba, M. Binding estimation after refinement, a new automated procedure for the refinement and rescoring of docked ligands in virtual screening. *Chem. Biol. Drug Des.* 2009, 73, 283–286, doi:10.1111/j.1747-0285.2009.00780.x.
90. Miller III, B. R.; McGee Jr., T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* 2012, 3314–3321.
91. Suan Li, M.; Khanh Mai, B. Steered Molecular Dynamics-A Promising Tool for Drug Design. *Curr. Bioinform.* 2012, 7, 342–351, doi:10.2174/157489312803901009.
92. Patel, J. S.; Berteotti, A.; Ronsisvalle, S.; Rocchia, W.; Cavalli, A. Steered Molecular Dynamics Simulations for Studying Protein–Ligand Interaction in Cyclin-Dependent Kinase 5. *J. Chem. Inf. Model.* 2014, 54, 470–480, doi:10.1021/ci4003574.
93. Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wrighers, W.; Schulten, K. Steered Molecular Dynamics. In: Springer, Berlin, Heidelberg, 1999; pp. 39–65.
94. Isralewitz, B.; Gao, M.; Schulten, K. Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.* 2001, 11, 224–230, doi:10.1016/S0959-440X(00)00194-9.
95. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Molec. Graph.* 1996, 14, 33–38.
96. Bento, a. P.; Gaulton, A.; Hersey, A.; Bellis, L. J.; Chambers, J.; Davies, M.; Krüger, F. a.; Light, Y.; Mak, L.; McGlinchey, S.; Nowotka, M.; Papadatos, G.; Santos, R.; Overington, J. P. The ChEMBL bioactivity database: An update. *Nucleic Acids Res.* 2014, 42, 1083–1090, doi:10.1093/nar/gkt1031.
97. The UniProt Consortium Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res.* 2013, 41, D43–7, doi:10.1093/nar/gks1068.
98. Vidal, D.; Nonell-Canals, A. ADMET models.
99. Huuskonen, J. Estimation of Aqueous Solubility for a Diverse Set of Organic Compounds Based on Molecular Topology. 2000, doi:10.1021/CI9901338.
100. Muehlbacher, M.; Spitzer, G. M.; Liedl, K. R.; Kornhuber, J. Qualitative prediction of blood–brain barrier permeability on a large and refined dataset. *J. Comput. Aided. Mol. Des.* 2011, 25, 1095–1106, doi:10.1007/s10822-011-9478-1.
101. Zhu, X.-W.; Sedykh, A.; Zhu, H.; Liu, S.-S.; Tropsha, A. The Use of Pseudo-Equilibrium Constant Affords Improved QSAR Models of Human Plasma Protein Binding. *Pharm. Res.* 2013, 30, 1790–1798, doi:10.1007/s11095-013-1023-6.
102. Poongavanam, V.; Haider, N.; Ecker, G. F. Fingerprint-based *in silico* models for the prediction of P-glycoprotein substrates and inhibitors. *Bioorg. Med. Chem.* 2012, 20, 5388–5395, doi:10.1016/J.BMC.2012.03.045.

103. Sedykh, A.; Fourches, D.; Duan, J.; Hucke, O.; Garneau, M.; Zhu, H.; Bonneau, P.; Tropsha, A. Human intestinal transporter database: QSAR modeling and virtual profiling of drug uptake, efflux and interactions., doi:10.1007/s11095-012-0935-x.
104. Yuan, S.; Chan, H. C. S.; Hu, Z. Using PyMOL as a platform for computational drug design. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2017, 7, e1298, doi:10.1002/wcms.1298.
105. Salentin, S.; Schreiber, S.; Haupt, V. J.; Adasme, M. F.; Schroeder, M. PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Res.* 2015, 43, W443–W447, doi:10.1093/nar/gkv315.
106. RDKit: Open-source cheminformatics.
107. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 2007, 9, 90–95.
108. Waskom, M.; Botvinnik, O.; O’Kane, D.; Hobson, P.; Ostblom, J.; Lukauskas, S.; Gemperline, D. C.; Augspurger, T.; Halchenko, Y.; Cole, J. B.; Warmenhoven, J.; Ruitter, J. de; Pye, C.; Hoyer, S.; Vanderplas, J.; Villalba, S.; Kunter, G.; Quintero, E.; Bachant, P.; Martin, M.; Meyer, K.; Miles, A.; Ram, Y.; Brunner, T.; Yarkoni, T.; Williams, M. L.; Evans, C.; Fitzgerald, C.; Brian; Qalieh, A. *mwaskom/seaborn: v0.9.0* (July 2018). 2018, doi:10.5281/ZENODO.1313201.
109. Montaser, R.; Luesch, H. Marine natural products: a new wave of drugs? *Future Med. Chem.* 2011, 3, 1475–89, doi:10.4155/fmc.11.118.
110. Kiuru, P.; D’Auria, M.; Muller, C.; Tammela, P.; Vuorela, H.; Yli-Kauhaluoma, J. Exploring Marine Resources for Bioactive Compounds. *Planta Med.* 2014, 80, 1234–1246, doi:10.1055/s-0034-1383001.
111. Grosso, C.; Valentão, P.; Ferreres, F.; Andrade, P. B. Review: Bioactive marine drugs and marine biomaterials for brain diseases. *Mar. Drugs* 2014, 12, 2539–2589, doi:10.3390/md12052539.
112. Martins, A.; Vieira, H.; Gaspar, H.; Santos, S. Marketed marine natural products in the pharmaceutical and cosmeceutical industries: Tips for success. *Mar. Drugs* 2014, 12, 1066–1101, doi:10.3390/md12021066.
113. Molinski, T. F.; Dalisay, D. S.; Lievens, S. L.; Saludes, J. P. Drug development from marine natural products. *Nat. Rev. Drug Discov.* 2009, 8, 69–85, doi:10.1038/nrd2487.

Section III

A combined computational and experimental study

Section III

Chapter 4

Meridianins and lignarenones as potential GSK3 β inhibitors and inductors of structural synaptic plasticity

Chapter 4

Meridianins and lignarenones as potential GSK3 β inhibitors and inducers of structural synaptic plasticity

Laura Llorach-Pares ^{1,2}, Ened Rodriguez-Urgelles ^{3,4,5}, Alfons Nonell-Canals ², Melchor Sanchez-Martinez ^{2*}, Jordi Alberch ^{3,4,5} Albert Giralt ^{3,4,5*} and Conxita Avila ¹,

1. Department of Evolutionary Biology, Ecology and Environmental Sciences, Faculty of Biology and Biodiversity Research Institute (IRBio), Universitat de Barcelona, 08028 Barcelona, Catalonia.
2. Mind the Byte S.L., 08028 Barcelona, Catalonia, Spain
3. Departament de Biomedicina, Facultat de Medicina, Institut de Neurociències, Universitat de Barcelona, Barcelona, Spain.
4. Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain.
5. Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, Spain.

Abstract

Glycogen Synthase Kinase 3 (GSK3) is an essential protein, with a relevant role in many pathogenesis such as diabetes, cancer, and neurodegenerative diseases. Particularly, the isoform GSK3 β , is related to pathologies such as Alzheimer's disease (AD). This enzyme constitutes a very interesting target for the discovery and/or design of new therapeutic agents against AD due to its relation to the hyperphosphorylation of the microtubule-associated protein tau (MAPT), and therefore, its contribution to neurofibrillary tangles (NFT) formation. An *in silico* study identified two marine molecular families, the indole alkaloids meridianins from the tunicate *Aplidium*, and lignarenones, the secondary metabolites of the shelled cephalaspidean mollusc *Scaphander lignarius*, as possible GSK3 β inhibitors. The analysis of the surface of GSK3 β revealed that both marine molecules can act over the ATP and/or substrate binding regions. Here, the predicted inhibitory potential of these two marine molecules was experimentally validated *in vitro* by the comparison Ser9 phosphorylation levels to total GSK3 β levels and also, we determined that both molecules potentiate structural synaptic plasticity. These allowed us to suggest that meridianins and lignarenone B could be used as possible therapeutic candidates for the treatment of GSK3 β involved pathologies, such as AD.

Keywords: Computer-aided drug design; Alzheimer disease; marine natural products, *Aplidium* tunicates; *Scaphander* molluscs.

Resum

La glycogen synthase kinase-3 (GSK3) és una proteïna essencial, amb un paper rellevant en moltes patologies com ara la diabetis, el càncer i les malalties neurodegeneratives. En particular, la isoforma GSK3 β està relacionada amb patologies com la malaltia d'Alzheimer (MA). Aquest enzim constitueix un objectiu molt interessant per al descobriment i/o disseny de nous agents terapèutics contra la MA a causa de la seva relació amb la hiperfosforilació de la proteïna tau associada als microtúbuls, i per tant, la seva contribució en la formació de cabdells neurofibril·lars. Els estudis *in silico* van identificar dues famílies de molècules marines, els alcaloides indòlics meridianines, del tunicat *Aplidium*, i les lignarenones, metabòlits secundaris del mol·lusc cefalaspidi amb closca *Scaphander lignarius*, com a possibles inhibidors GSK3 β . L'anàlisi de la superfície de GSK3 β revelà que les dues molècules marines poden actuar sobre les regions d'unió del trifosfat d'adenosina (ATP) i/o del substrat. El predit potencial inhibitori d'aquestes dues molècules marines va ser validat experimentalment *in vitro* mitjançant la comparació dels nivells de fosforilació de Ser9 i els nivells totals de GSK3 β , a la vegada que es va poder observar que les dues molècules potencien la plasticitat sinàptica estructural. Aquests fets ens permeten suggerir que les meridianines i la lignarenone B podrien ser utilitzats com a possibles candidats terapèutics per al tractament de GSK3 β implicada en patologies com ara la MA.

Paraules clau: Disseny de fàrmacs assistit per ordinador, Malaltia d'Alzheimer, Productes naturals marins, Tunicats *Aplidium*, Mol·luscs *Scaphander*.

Introduction

Glycogen Synthase Kinase 3 (GSK3) is one of the best known kinases and it has been widely studied since in 1992, when its key contribution to the abnormal phosphorylation of the microtubule-binding protein tau (MAPT) in the process thought to cause neurofibrillary tangles (NFT) formation in Alzheimer's disease (AD), was first discovered [1,2]. The most consistent AD disease manifestations are extracellular senile plaques composed by amyloid- β (A β) proteins and NFT, mainly formed by hyperphosphorylated tau protein. In the last years, GSK3 has emerged as a potential drug target. GSK3 is a ubiquitous serine (Ser)/threonine (Thr) protein kinase widely expressed in many types of cells and tissues, and particularly abundant in the brain. GSK3 is involved in a diversity of processes and pathways, catalysing the transfer of a phosphate group from adenosine triphosphate (ATP) to Ser and Thr aminoacid residues of target substrates. GSK3 was considered interesting because its wide involvement in diseases and cellular processes, but also for its unconventional characteristics for being a kinase. GSK3 is constitutively active, its substrates usually need to be pre-phosphorylated by another kinase, and it is inhibited, rather than activated, in response to stimulation of the insulin and Wnt pathways [3–5]. There are two highly conserved isoforms of GSK3, that is GSK3 α and GSK3 β , sharing an overall identity of 84%, and a 98% of sequence identity on their catalytic domains. Due to the historical relation of the isoform β with several neurological diseases as well as aging processes or diabetes, it has received more attention than GSK3 α . GSK3 β , widely present in the brain, is associated with several neurodegenerative diseases, including Parkinson's disease (PD), AD, and Huntington's disease (HD) [6–8].

The predominant hypothesis in AD suggest that the activity of phosphatases and kinases, in particular GSK3 β , is affected by the amyloid peptides. All these changes result in an increase of protein Tau phosphorylation. Changes in kinase activity are an intrinsic aspect of the pathological problem in AD, as they negatively affect, even interrupting, synaptic signals essential for learning and memory [9]. GSK3 activity can be regulated by serine 9/21 phosphorylation and by Thr phosphorylation at residues 216 and 279, although this last phosphorylation is less common. GSK3 is usually phosphorylated at different sites, but the regulatory outcomes of this remain unclear [3].

In AD, GSK3 β is commonly regulated by inhibitory phosphorylation on Ser9, located at the N-terminal tail. The deregulation of this process results in a GSK3 β permanent abnormal activation which then, hyperphosphorylate tau leading to its aggregation [7,10–12]. From a drug development perspective, the strategies aimed to target GSK3 β are oriented towards reducing tau hyperphosphorylation by its inhibition. Thus, this is considered a promising therapeutic avenue for AD, even more now that the amyloid hypothesis, that was never universally accepted, has not given too

promising results [13,14]. Although it has been the predominant target and therapeutic hypothesis in the last years, as corroborated by its predominant presence in the 112 molecules tested in clinical trials from Phase I to III in 2018 [15,16]. Inhibition of tau aggregation is conceptually more tempting, because there seems to be a wider general consensus about its damaging effects [17]. In that sense, GSK3 β is related to the amyloid hypothesis as the main therapeutic target, where A β fibrillar forms, are a possible therapeutic alternative. Significant efforts have been made in the past years to design new potent and selective GSK3 inhibitors, acting over the ATP catalytic pocket or over other allosteric cavities [18]. However, most of the obtained compounds have been considered as hits or starting points, but have not advanced to the clinic because of administration, distribution, metabolism, excretion and toxicity (ADMET) problems [19]. In fact, some of the early GSK3 β inhibitors that entered into clinical trials failed for toxicity problems or because off-target interactions, among other reasons [20,21]. Thus, there is still an opportunity to develop better and safer GSK3 β inhibitors, and marine natural products could play a key role on this [22–27].

Bioactive natural products, from animals, plants, fungi, or microorganisms, are a source of inspiration and play an important role in the discovery and design of new drugs [28–30]. Also, these biological active metabolites can be used as a template to design drug-like compounds. One of the most promising lines in the drug discovery pipeline is the identification of natural products capable of modulating protein functions in pathogenesis-related pathways [31]. Within natural products, those coming from the still under explored marine world are of great interest since they shelter most of the biodiversity of the world [32,33]. An assessment of all Food and Drug Administration (FDA) approved new molecular entities (NMEs) reveals that natural products and their derivatives represent over 38% of all NMEs [34]. In other words, by the end of 2013 the FDA approved 547 natural products [34]. AD is a pathology where several natural products have been proposed as drug candidates, covering different therapeutic activities [22–24]. Indeed, in the last few years, an increasing number of possible GSK3 β inhibitors have been reported from marine invertebrates, as hymenialdisine isolated from a marine sponge *Axinella* [35], the natural phenylmethylene hydantoin (PMH) isolated from the Red sea sponge *Hemimycale arabica* [36], or even meridianins and its derivatives from tunicates, revealing the potential of marine natural products as GSK3 β potential therapeutic agents [25–27,37]. These compounds can and should be a starting point to develop new bioinspired drugs against GSK3 β . World oceans and its coasts are exceptionally rich in species diversity and its exploration offers unique life forms full of structurally diverse organisms and biological compounds which can also be tested as therapeutic agents

[38]. The study of some of these marine natural products revealed their potential inhibitory function, that can be further exploited and also developed as marketable drugs [39,40].

In a previous study of our group, aimed to unraveling the potential biological activities of a set of selected small marine molecules, two of them were found particularly interesting as potential therapeutic agents against GSK3 β : meridianin A and lignarenone B [41]. Meridianins are a family of marine indole alkaloids isolated from the cold waters of Antarctica. The natural products of this ascidian consist of an indole framework linked to an aminopyrimidine ring, among other compounds. In contrast, lignarenones, isolated from a mollusc in the temperate waters of the Mediterranean Sea, are two phenyl conjugated trienones, also classified as polyketides (Figure 1). These compounds have been isolated from specimens of the tunicate genus *Aplidium* and the shelled cephalaspidean mollusc *Scaphander lignarius*, respectively [42,43].

Supporting this finding, a large number of heterocyclic inhibitors of GSK3 have been identified in the past few years and they can be classified as ATP-competitive and non-ATP-competitive. The largest group is that of compounds acting in the ATP catalytic site, *i.e.* competing against ATP, including pyrazolopyrimidines, benzimidazoles, pyridinones, pyrimidines, indolylmaleimides, imidazopyridines, oxadiazoles, or pyrazines among others [31]. Inhibitors that do not compete with ATP, acting over allosteric cavities, include for instance thiadiazolidinones (TDZDs), halomethylketones (HMK), or 5-imino-1,2,4-thiadiazole (ITZDs) derivatives [18,44,45]. In addition to the aforementioned groups of inhibitors, in the last few years, a growing number of GSK3 inhibitors have been reported from marine invertebrates, such as hymenialdisine, indirubines, manzamines, isoflavones or the linear furanosesquiterpenes palinurin and tricantin [25,27,46]. Interestingly, the chemical structures of most of these marine inhibitors have an heterocyclic scaffold, similar to the indole scaffold of meridianins (Figure 1). However, not all of them are heterocyclic compounds; for example, palinurin and tricantin, are sesquiterpenes which present a linear structure, together with aromatic cycles in the terminal part, similar to lignarenones (Figure 1). Also, it has to be mentioned that these compounds have been synthesized in the past and this is relevant for the development of new derivatives from both meridianins and lignarenones [47–50].

Meridianin A	R1 = OH, R2 = H, R3 = H, R4 = H
Meridianin B	R1 = OH, R2 = H, R3 = Br, R4 = H
Meridianin C	R1 = H, R2 = Br, R3 = H, R4 = H
Meridianin D	R1 = H, R2 = H, R3 = Br, R4 = H
Meridianin E	R1 = OH, R2 = H, R3 = H, R4 = Br
Meridianin F	R1 = H, R2 = Br, R3 = Br, R4 = H
Meridianin G	R1 = H, R2 = H, R3 = H, R4 = H

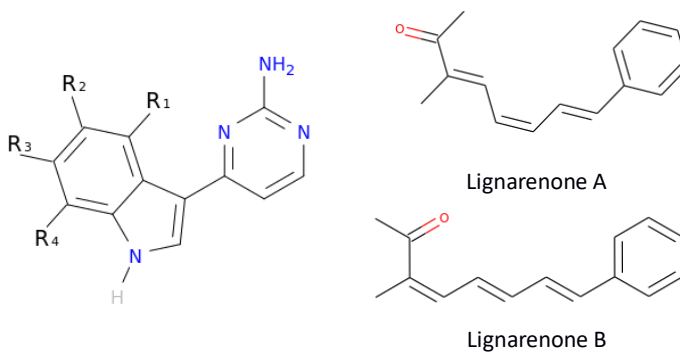


Figure 1. Structure of meridianins A-G and lignarenones A-B, natural compounds isolated from marine organisms.

In this study, we analyze and report the possible inhibitory activity of meridianins and lignarenones. In the case of the lignarenones, due to the high structural similarity between both (A and B) species, almost identical from a computational point of view, computational studies were performed only over lignarenone B. We aimed to show whether these compounds were able to inhibit GSK3 β activity. To do so, a first computational approach using docking calculations and molecular dynamics (MD) simulations was performed with the aim of elucidating the ability of meridianins and lignarenones to act as ATP competitive or non-ATP-competitive inhibitors, and finding possible allosteric binding cavities if it was the case. Moreover, an experimental validation of the inhibitory activity of both molecular families was performed by Dr. Alberch group, interested in identifying new therapeutic targets to develop new treatments for neurodegenerative diseases.

Results and discussion

Exploring druggable binding sites on GSK3 β .

In this study, with the objective of identifying novel allosteric binding sites, using fpocket [51] to analyse the GSK3 β surface, we obtained 15 plausible cavities (Figure 2). Previous search of crystallographic structures of GSK3 β from the Protein Data Bank (PDB) [52] resulted in more than thirteen different GSK3 β complexes available with a good resolution (lower than 2.5 Å). After a careful check of all of them, we selected the 6B8J crystal structure, representing human GSK3 β together with CHIR99021, a selective inhibitor, released in 2017 [53]. This new PDB was not used in previously published works where similar studies were performed [18]. GSK3 β ATP catalytic pocket has been widely explored for decades, often showing the specificity problems characteristic of protein kinases. Allosteric

cavities open the possibility of designing inhibitors without these inconveniences. The search of druggable cavities over the protein surface, that could be an allosteric pocket or a place where a molecule can be retained, for instance, affecting the efficacy of a given drug, is becoming more common nowadays. There are several software tools designed to do that, based on different principles like geometrical, probe and/or energy-based algorithms. Detection, comparison, and analyses of ligand binding pockets is a key step in structure-based drug design [54]. One of the main issues associated with pocket finding is druggability prediction, as it is important to avoid failing during the drug discovery process and to focus the efforts of the process on the discovery of cavity sites that can offer a better prospect. The concept of druggability adds a new dimension to pocket finding, as it evaluates the likelihood that small drug-like molecules can bind a given cavity, and thus a concrete target; in some cases, even evaluating if the binding could have sufficient potency to alter the protein activity [55].

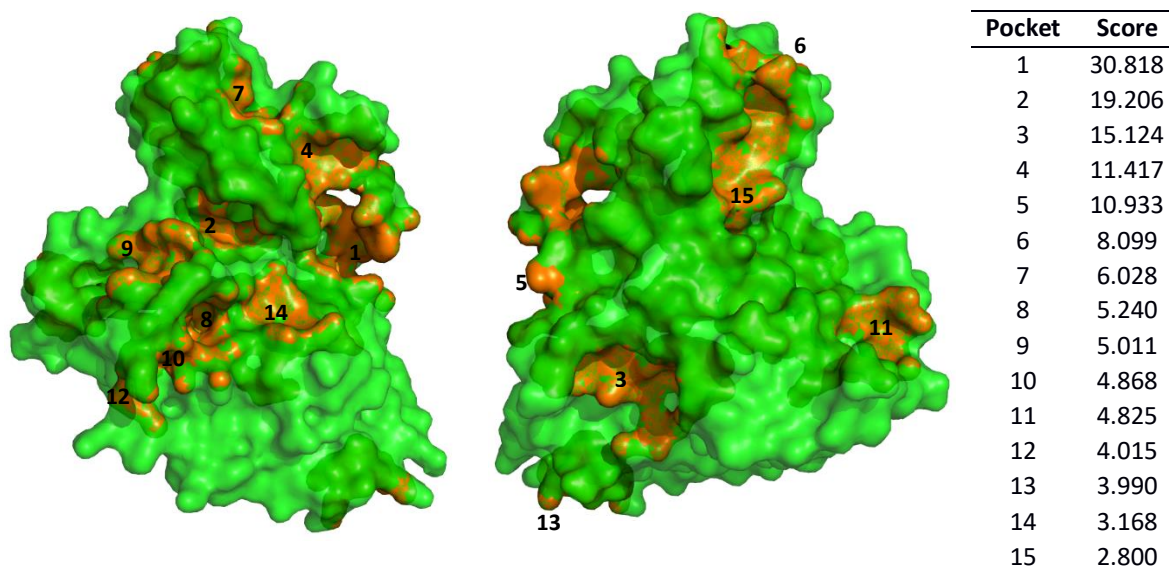


Figure 2. Cavities found (orange) by fpocket and plot on surface of GSK3 β (green). Numbers correspond to the ranking given by fpocket (1 is the best, and 15 the worst). In the list the score values of each pocket are shown. Noworse details regarding the scoring function of fpocket can be found at Le Guilloux, et al. (2009).

Most of the pockets found here were previously described by Palomo et al. (2011) over a different crystal structure, thus supporting the output of the performed cavity detection. Interestingly, looking at the obtained ranked pocket list (Figure 2), and the corresponding images, the best cavity, number 1, is not the well-known ATP-binding site, which is number 2. Instead, the best cavity corresponds to the substrate binding site. The finding of the substrate pocket as a plausible drug-binding

cavity is supported by literature data. For instance, Manzamine A, a β -carboline alkaloid isolated from the marine sponges *Haliclona* and *Acanthostrongylophora*, or its derivatives, can bind to GSK3 β near the activation pocket formed by the key residues ARG96, ARG180 and LYS205 [56]. These residues are located at the predicted cavity 1, the substrate pocket. Moreover, Zefirov et al. published in 2010 a study confirming the interaction of manzamine A with this pocket [57]. Therefore, despite the druggability score of the substrate pocket being lower than that obtained for the ATP cavity, our results are in agreement with the literature, reinforcing the election of pocket 1 as a good binding site for small molecules. It is worth mentioning that the fpocket ranking is based in a general score calculated from different properties, including druggability (see methods section and Le Guilloux et al., 2009 for further details) [51]. Thus, the pocket with the highest druggability may not be the “best” pocket, as seen here, because the ATP-pocket is more druggable than the substrate pocket, but the rest of the properties placed it in the second place. Anyway, both pockets are good enough to host drug-like compounds.

Due to the interest in developing allosteric inhibitors and the high number of cavities found on the surface of GSK3 β , it is relevant to study the capacity of meridianins and lignarenones to act as allosteric inhibitors as well as ATP-competitive inhibitors. From the 15 cavities detected by fpocket, we decided to focus only on the ATP cavity and the substrate binding pocket since, due to its druggability and structural properties. As said, they are the two most suitable cavities to host a small molecule inhibitor. We also showed the location of the pockets and the amino acids involved in each cavity (Figure 3).

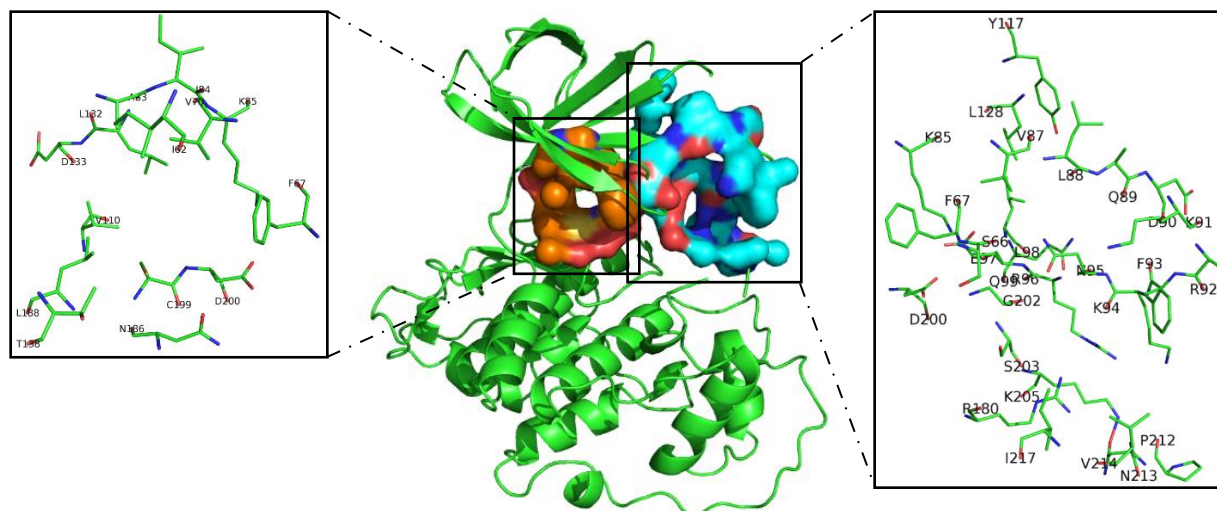


Figure 3. Schematic representation of the structure of GSK3 β (PDB 6B8J) in the central image. In orange and surface the ATP-cavity are shown, while the blue surface represents the substrate binding pocket. In the left zoom image, all the amino acid residues that construct the ATP pocket are shown in sticks. In the right zoom image are shown those the residues that compose the substrate pocket. Letters and numbers correspond to their position in the amino acid sequence and the 6B8J PDB file numbering.

Binding of meridianins and lignarenone B to the ATP and substrate cavities.

To validate the proposed cavities and with the aim of evaluating the behaviour of the two marine molecular families as ATP or non-ATP-competitive inhibitors, docking calculations followed by MD simulations to post-processing were performed. Docking calculations revealed that meridianins bind better than lignarenone B to the ATP and the substrate cavities of GSK3 β , especially to the ATP catalytic pocket. It was also observed that meridianin compounds prefer the ATP cavity rather than the substrate pocket, although they can also bind to both places reasonably well, as does lignarenone B.

After docking calculations, MD simulations were performed in order to validate docking results. One of the main characteristics of proteins is their flexibility, essential to carry out their functions. Docking calculations do not usually consider this, and post-processing the obtaining docking conformations by MDs is a good way to take it into account. In fact, this way to observe the so called-induced fit events that allow the adaptation of the ligand to the target and vice versa (whereas docking, usually rigid, only allows the ligand movement), constitutes a well-established pipeline to study ligand-protein binding [58–60].

After MD simulations, by using molecular mechanics/Generalized Born and surface area (MM/GBSA) alchemical free energy calculations, the binding energy of the simulated ligand-protein

complexes was estimated, ranking the compounds according to it. The obtained energy ranking is more realistic than the one obtained from docking, as the dynamic nature of the protein and the solvent environment are added to the equation (although the obtained energies are far from being directly comparable to those obtained experimentally) [61]. The observed trend at docking results is confirmed after MD, being all the best energies corresponding to the molecules binding over the ATP cavity, which are at least 6 kcal/mol higher than those obtained on the substrate pocket (although most are around 10 kcal/mol and on the particular case of meridianin F, reaching a 18 kcal/mol difference). We obtained the binding energies after each calculation (Table 1).

Table 1. Summary of the binding energy results after docking calculations and after 1 ns MD simulations with molecular mechanics/generalised born surface area MM/GBSA calculations. All energy values are in kcal/mol.

	Substrate cavity		ATP cavity	
	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)
	R0/R1		R0/R1	
Meridianin A	-6.2/-6.2	-22.39	-7.8/-7.8	-28.68
Meridianin B	-6.4/-6.4	-23.60	-7.4/-7.4	-34.23
Meridianin C	-6.5/-6.5	-17.48	-8.0/-8.0	-32.18
Meridianin D	-6.1/-6.2	-25.59	-7.6/-7.5	-35.53
Meridianin E	-6.4/-6.4	-23.86	-7.6/-7.6	-28.18
Meridianin F	-6.7/-6.6	-23.49	-8.1/-8.2	-41.75
Meridianin G	-6.8/-6.8	-16.89	-7.9/-7.9	-29.88
Lignarenone B	-6.0/-6.0	-18.59	-6.7/-6.7	-28.79

The binding mode of the complexes (GSK3 β - marine molecule), when the compounds are bound to the ATP and the substrate binding pocket, were found to be very stable. This stability was evaluated qualitatively by carefully visualizing the generated trajectories and complexes, and it was also assessed quantitatively by the analysis of the hydrogen bonds (HBs) present on each complex along the MD trajectory, as well as by analysing the temperature, kinetic analysis, and root-mean-square deviation (RMSD), as seen in Appendix. Since MDs are dynamic processes, the number of HBs is not constant; they can be continuously forming and breaking or be stable, depending on the system under study. In the case of GSK3 β bound to meridianins A-G and lignarenone B, we found nine important HBs at the ATP cavity. These HBs are established with residues F67, V70, K85, D133, V135, R141, Q185, C199, and D200. Most of the listed aminoacidic residues are configuring the ATP binding pocket (Figure 4) [20,25,26,57]. For the substrate binding pocket, we found 16 important HBs. These were formed with F67, K85, K86,

L88, Q89, F93, K94, N95, R96, R180, G202, S203, A204, K205, E211, and P212. All these residues are part of the substrate pocket (Figure 6) [46]. In addition to the HBs, we also found that both, meridianins and lignarenone B, can interact with I62, F67, V70, A83, K85, E97, L132, T138, L188, and D200 via hydrophobic contacts on the ATP cavity and on the substrate pocket. Meridianins establish hydrophobic contacts with D90, F93, K94, N95, R96, K205, and I217, while lignarenone B does with D90, F93, N95, and R96. Also, two salt bridges were established on this pocket with residues D90 and E211 when meridianins are bound.

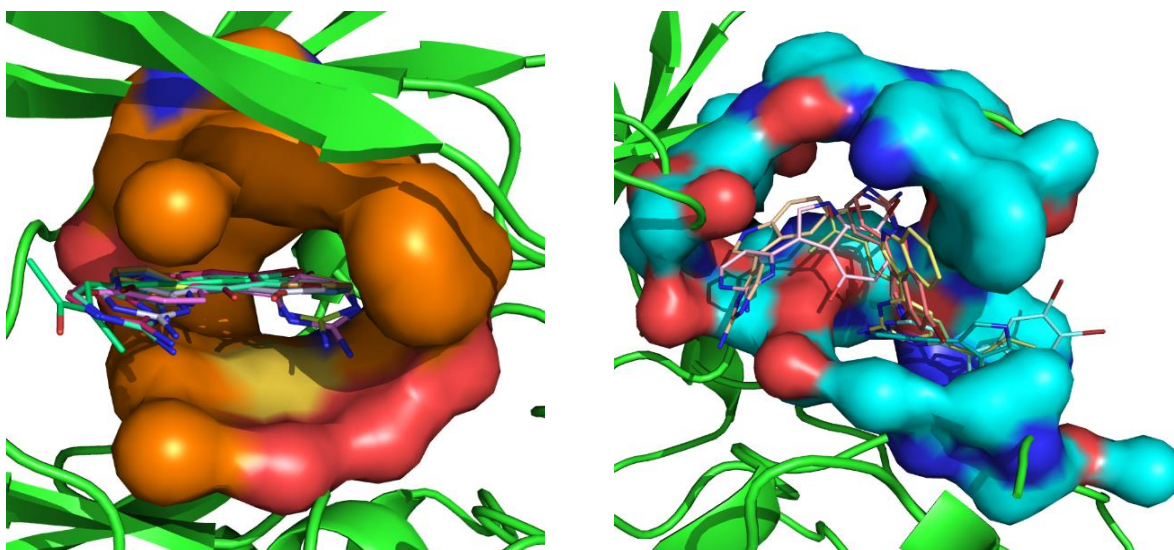


Figure 417. Representation of the two binding cavities ATP and substrate in surface and the binding mode of the marine molecules. On the left, the ATP pocket with all the meridianins and lignarenone B. On the right the substrate pocket also with all the meridianins and lignarenone B. Both images represent the last frame after MD simulation. Meridianin A-G colours: Peach, blue, tan, orange, pink, cyan and yellow. Lignarenone colour: green.

In the ATP cavity, meridianins A-G and lignarenone B are placed on a very similar way, while on the substrate binding pocket the molecular placement of each compound varies with respect to the other, although with some shared features (Figure 4). One fact that could explain this pattern is the different dimension of the pockets. While the substrate pocket has a volume of 1808.60 \AA^3 , the ATP pocket represents a quarter of its volume 404.38 \AA^3 , and thus, due to the size differences between meridianins and lignarenones in a reduced space the position should be different.

The docking post-processing of MDs allows the observation of induced fit events, as mentioned above. The existence or the magnitude of these events can be measured in different ways. For example, qualitatively looking how the binding site slightly changes its conformation, or measuring how the

interaction that stabilizes the ligand over the protein changes with the MD, or analysing the RMSF of the ligand-target complexes along MDs, *i.e.* measuring the amplitude of atom motions during the MD trajectory, and elucidating the flexible regions of the protein [62,63]. In this particular case, we analysed the RMSF of GSK3 β when meridianins and lignarenone B are bound over the ATP or substrate pockets (Figure 5), obtaining homologous results. We found six peaks of fluctuations (corresponding to six different protein regions) with high mobility respect to the baseline. The first one is placed on residues 49-50, the second and the highest one between residues 91 to 94, the third involved the residue 124, the fourth involved residues 148 and 150, the fifth fluctuation occurs on residue 209, and the sixth is localized on residues 290 and 292.

In the RMSF, when the marine molecules are bound to the substrate pocket, one more fluctuation on residue 66 can be observed, while in the ATP fluctuation analysis, focusing on this residue, despite the RMSF values are not zero (in fact they are a little bit higher than 0.20 nm) they do not reach the amplitude of others (going from 0.25 to 0.40/0.50 nm, depending on the molecule bound). Only when meridianin G and lignarenone B are bound, fluctuations reach the 0.25 nm on this 66 residue.

Another pattern that has been detected is the fact that when lignarenone B is bound to any of both pockets, the fluctuation of GSK3 β is in general higher than when meridianins bound, and this could be related to the structure (size, linearity, etc.) of lignarenone B, which could provoke larger changes in the protein conformation to allow an optimal placement, improving the poses obtained during the docking process.

On the other hand, residue K94, present on the substrate pocket and a key component of the cavity, which establishes HBs with the marine molecules during the binding, is highly fluctuating. This can be explained because this residue is placed on the loop of the N-lobe, a very flexible region, which is very exposed to the solvent during the MD simulation [64]. Also, substrate recognition requires GSK3 β residues, F67, Q89, and N95, which facilitate the precise positioning of the substrate within the substrate binding pocket, and provide an insight into the substrate binding and specificity [65]. To do so, the flexibility of the loop is necessary, and this is translated into high RMSF values of these residues and those next to them, as seen on peak 2 and the peak on the residue 66 during the binding of the meridianins and lignarenone B to the substrate pocket (Figure 5).

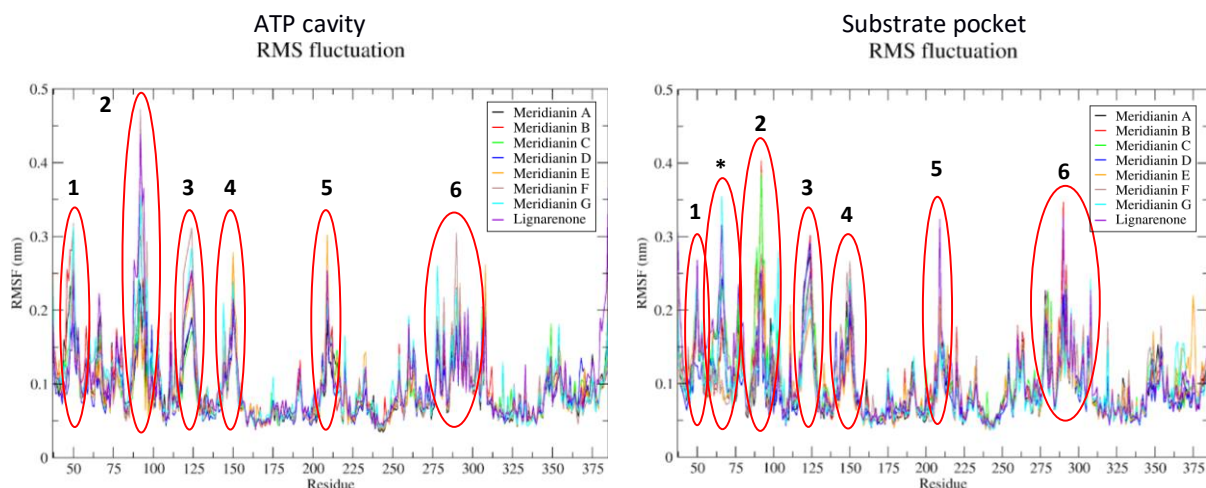


Figure 5. Root-mean-square fluctuation (RMSF) per residue (X-axis) values of each complex, GSK3 β + marine molecules separated per pocket, along the molecular dynamic (MD) simulation. On the left the RMSF of each system were marine molecules are bound to the ATP cavity and on the right when the marine molecules are bound to the substrate pocket. The highest fluctuations (>0.25 nm) detected are highlighted with red circles and those shared between RMSFs are numbered in order of appearance. The asterisk (*) indicates a fluctuation >0.25 nm, only observed on the substrate pocket. The colour code for each system can be seen in the legend box.

Evaluation of the pharmacokinetic properties

In the early stage of the drug discovery process, pharmacokinetics (PK) studies play a key role for developing new molecules, as they can predict the safety and efficacy of potential hit candidates, facilitating the appropriate lead compounds selection, as well as saving investments in terms of money and time in expensive clinical trials [66,67]. These studies are aimed to predict the absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties of potential therapeutic compounds. ADMET properties prediction of meridianins A-G and lignarenone B (Table 2).

Table 2. Summary of absorption, distribution, metabolism, excretion and toxicity properties of meridianins and lignarenone B. Pgp: P-Glycoprotein, BBB: blood-brain barrier, PPB: plasma-protein binding, CYP450: cytochrom P450, OCT2: organic cation transporter 2, hERG: human ether-a-go-go gene.

	Absorption			Distribution			Metabolism	Excretion	Toxicity	
	Mol Weight	logS	Pgp	Caco2	logP	BBB	PPB	CYP450	OCT2 Substrate	hERG
Meridianin A	226.2	-4.2	inactive	High	1.5	NO	High	Yes	No	<4.0
Meridianin B	305.1	-5.0	inactive	High	2.4	NO	High	Yes	No	<4.0
Meridianin C	289.1	-5.6	inactive	High	3.1	NO	High	Yes	No	<4.0
Meridianin D	289.1	-5.6	inactive	High	3.1	NO	High	Yes	No	<4.0
Meridianin E	305.1	-5.0	inactive	High	2.4	NO	High	Yes	No	<4.0
Meridianin F	368.0	-6.2	inactive	High	3.6	NO	High	Yes	No	<4.0
Meridianin G	210.2	-4.5	inactive	High	2.4	NO	High	Yes	No	<4.0
Lignarenone B	212.3	-3.2	inactive	Low	3.6	NO	High	Yes	No	<4.0

In general, all the tested compounds should be optimized to improve some of their properties. According to the results, they can be considered hits, but there is still work to do before they become suitable lead compounds. Meridianins, as shown before, present some solubility problems (logS > 5 indicate not too much solubility), especially meridianin F, as well as shows difficulties to penetrate the blood brain barrier (BBB) [25], a problem that is also shared by lignarenone B. This is also in agreement with the obtained logP values (where values lower than 5 indicate that the compounds have an appropriate hydrophobicity and permeability behavior). However, to become drugs penetrating the central nervous system (CNS), molecules should have a logP around 2 [68]. Some meridianins are almost there, but most of our compounds are far of this optimal value. Another issue is that all the studied compounds present high plasma-protein binding (PPB) probability, indicating that a high percentage of the administrated compound will be found attached to plasma proteins, affecting their diffusion and efficiency (Table 2).

On the other hand, all of the compounds seem to be permeable, especially meridianins, according to Caco2 results (Table 2). Regarding P-glycoprotein (Pgp) binding, none of our compounds was predicted to act over it. The interaction with Pgp has many pharmacological implications that could result in pharmaceutical advantages or contraindications. For instance, Pgp modulation has been suggested as a mechanism to improve CNS pharmacotherapy, but it also plays a major role in the multidrug resistance (MDR) phenomenon in cancer cells, depending on whether binding happens as a substrate or as an inhibitor, and also on the isoform that the compound binds to [69–72]. Thus, none of the molecules tested here seems to interact with Pgp, avoiding possible beneficial but also detrimental effects.

These compounds are able to interact with cytochrom P450 (CYP450), as described in the results (Table 2) and CYP450 drug metabolism can induce clinical effects. On the contrary, any molecule is a substrate of organic cation transporter 2 (OCT2), which means that non-clearance problems and adverse interactions with co-administrated OCT2 inhibitors are expected. Moreover, as another positive point, none of the compounds is an inhibitor of the potassium channels encoded by the human ether-a-go-go gene (hERG). hERG inhibition can lead to fatal pathologies, such as cardiac diseases, because it is the principal cause of the development of acquiring long QT syndrome, fatal arrhythmia, for example [73]. The absence of inhibition of hERG is a good and safe property of both meridianins and lignarenones.

Meridianins and lignarenones differentially increased pGSK3 β Ser9, but not total GSK3 β levels *in vitro*

In order to confirm the predicted interaction between meridianins and lignarenone B towards GSK3 β , we used different doses of both marine molecules, at different times, to pharmacologically inhibit GSK3 β , and comparing Ser9 phosphorylation levels to total levels of GSK3 β as an indication of inhibition. Primary cortical cultures of neurons were treated with vehicle, meridianins, and lignarenones (500 nM) or meridianins and lignarenones (10 μ M) for two time points, 15 and 60 min. Western blot analysis was used to determine protein expression levels of GSK3 β and pGSK3 β Ser9, which is an inhibitory phosphorylation site.

A one-way ANOVA analysis indicated that meridianins treatment significantly increased pGSK3 β levels, both after 15 min ($F_{2,30} = 4.189$, $p = 0.024$) and after 60 min ($F_{2,27} = 6.892$, $p = 0.0038$). Specifically, *post hoc* analysis revealed that at 60 min both doses, 500 nM ($p < 0.05$) and 10 μ M ($p < 0.01$) significantly increased pGSK3 β levels. Similarly, *post hoc* analysis after 15 min of treatment indicated that the dose of 10 μ M ($p < 0.01$), but not the dose of 500nM, exerted significant effects on pGSK3 β levels (Figure 6).

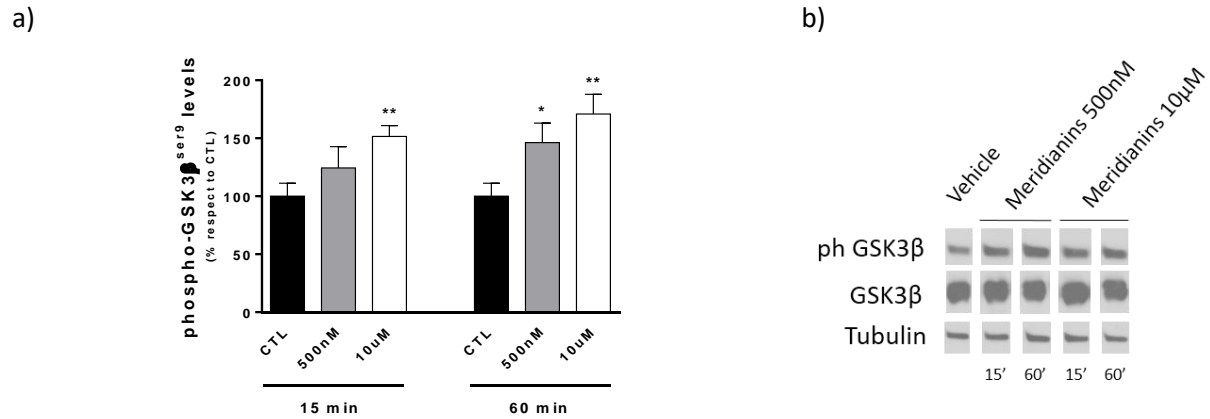


Figure 6. Increased GSK3β phosphorylation at Ser9 residue by meridianins. (a) Total GSK3β in cortical primary cultures treated with meridianins (500nM or 10µM) or vehicle, for 15 and 60 minutes and (b) Densitometric quantification of pGSK3β. Protein levels were normalized with beta-tubulin as loading control. Data were expressed as (mean ± S.E.M). Data were analysed by one-way ANOVA followed by Dunnett's test. *P < 0.05 and **P < 0.01 compared with vehicle. Representative immunoblots are shown. n = 16–9 cultures per condition at 15 min and 16-7 cultures per condition at 60 min.

Regarding lignarenones, one-way ANOVA analysis indicated that only after 15 min there was a significant change on pGSK3β levels ($F_{2,34} = 3.548$, $p < 0.05$). *Post hoc* analysis indicated that this increase was only observed with the dose of 500nM ($p < 0.05$). After 60 min of treatment with lignarenones, neither the 500nM dose nor the 10 µM dose induced changes on pGSK3β levels ($F_{2,28} = 0.5814$, $p = 0.56$) (Figure 7).

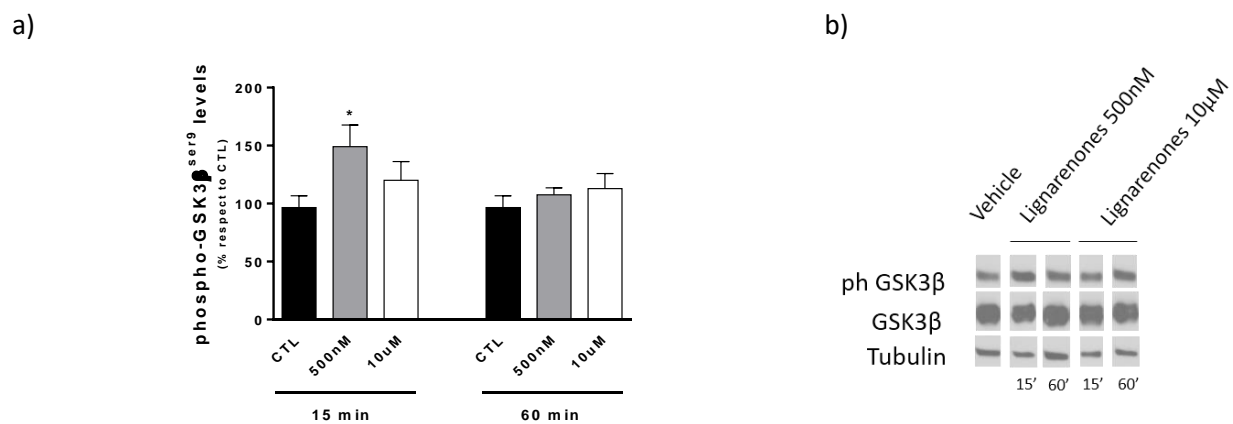


Figure 7. Increased GSK3β phosphorylation at Ser9 residue by lignarenones. (a) Total GSK3β in cortical primary cultures treated with lignarenones (500nM or 10µM) or vehicle, for 15 and 60 min and (b) Densitometric quantification of pGSK3β. Protein levels were normalized with beta-tubulin as loading control. Data were expressed as (mean ± S.E.M). Data were analysed by one-way ANOVA followed by Dunnett's test. *P < 0.05 compared with vehicle. Representative immunoblots are shown. n = 17–10 cultures per condition at 15 min and 17-7 cultures per condition at 60 min.

These results, far to be opposite, are complementary to the previous ones. On one hand, lignarenones have an acute effect within 15 min and then the effect decreases, while, on the other hand, meridianins effects are more sustained over time. In the case of meridianins, the results are more robust in terms of inhibition. Interestingly, the GSK3 β total levels remain stable and this indicates any of these marine molecules affect culture viability.

Meridianins and lignarenones regulate neuritic complexity *in vitro*

To evaluate the possible effects of meridianins and lignarenones in neuronal structural plasticity, primary cortical neurons were treated at 4 DIV with 10 μ M of meridianins and lignarenones (out highest dose). Three days after the treatment, we analyzed morphological characteristics of the imaged neurons stained for MAP2 by using the Sholl analysis. The results of meridianins treatment indicated that the number of intersections in the treated cultures were increased compared with the non-treated cultures (two-way ANOVA analysis; interaction effect, number of dendrites, $F_{7,536} = 55.91$, $p < 0.0001$) (Figure 8).

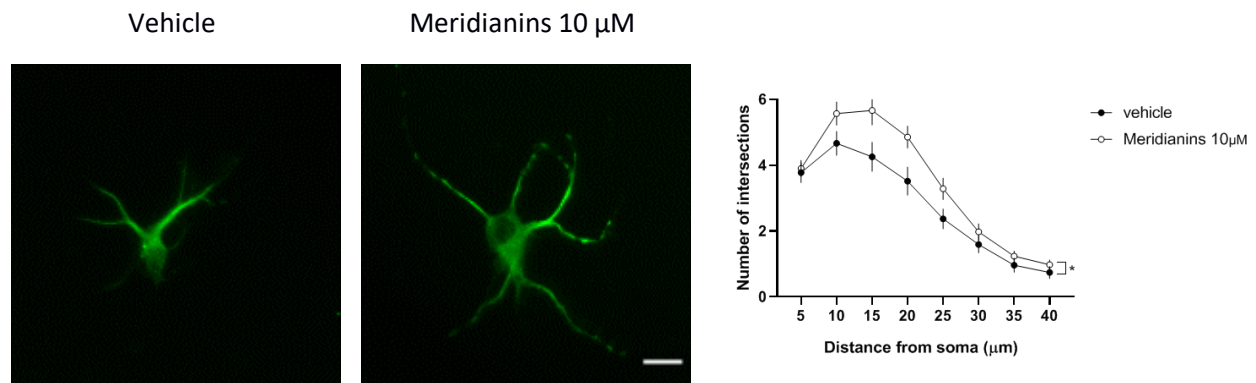


Figure 8. Meridianins up-regulate neuritic complexity *in vitro*. Representative MAP2 images obtained by epifluorescent microscopy from primary cortical neurons (left). Scale bar, 40 μ m. On the right, Sholl analysis from MAP2-positive neurons (two-way ANOVA analysis; interaction effect, number of dendrites, $F_{7,536} = 55.91$, $p < 0.0001$). $n = 45$ -55 neurons per condition.

The results obtained for lignarenones are similar to those obtained for meridianins and indicated that the number of intersections in the treated cultures were increased compared with the non-treated (two-way ANOVA analysis; interaction effect, $F_{7,815} = 7.247$, $p < 0.0001$; group effect, $F_{7,815} = 67.90$, $p < 0.0001$; number of dendrites, $F_{7,815} = 72.51$, $p < 0.0001$) (Figure 9).

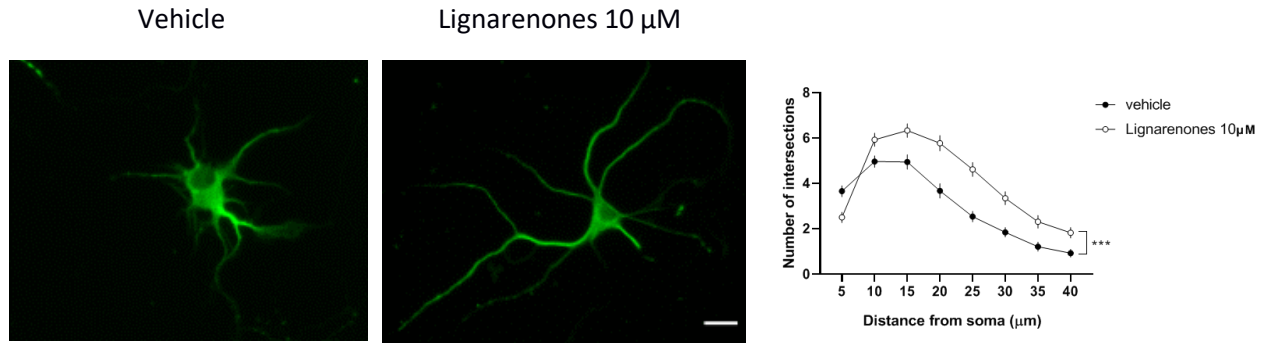


Figure 9. Lignarenones up-regulate neuritic complexity *in vitro*. Representative MAP2 images obtained by epifluorescent microscopy from primary cortical neurons (left). Scale bar, 40 μ m. On the right, Sholl analysis from MAP2-positive neurons (two-way ANOVA analysis; interaction effect, $F_{7,815} = 7.247$, $p < 0.0001$; group effect, $F_{7,815} = 67.90$, $p < 0.0001$; number of dendrites, $F_{7,815} = 72.51$, $p < 0.0001$). $n = 45$ -55 neurons per condition.

In both cases, positive results are obtained because an increase of the neurite outgrowth is clearly observed. These results are in agreement with the previous ones, reported above.

Effect of meridianins and lignarenones on neuronal viability

To elucidate pharmacological effects of meridianins and lignarenones treatments, we then analysed the cell viability on primary cortical cultures treated at 4 DIV at 10 μ M (out highest dose). Even if some previous works did mention the toxicity of this kind of marine molecules at neuronal level (Llorach-Pares et al., 2017), our results showed that neither meridianins nor lignarenones induced changes in cell density in our primary cultures, meaning that cells were viable under the experimental conditions evaluated (one-way ANOVA analysis, $F_{2,13} = 1.600$, $p = 0.2392$) (Figure 10) [25,74].

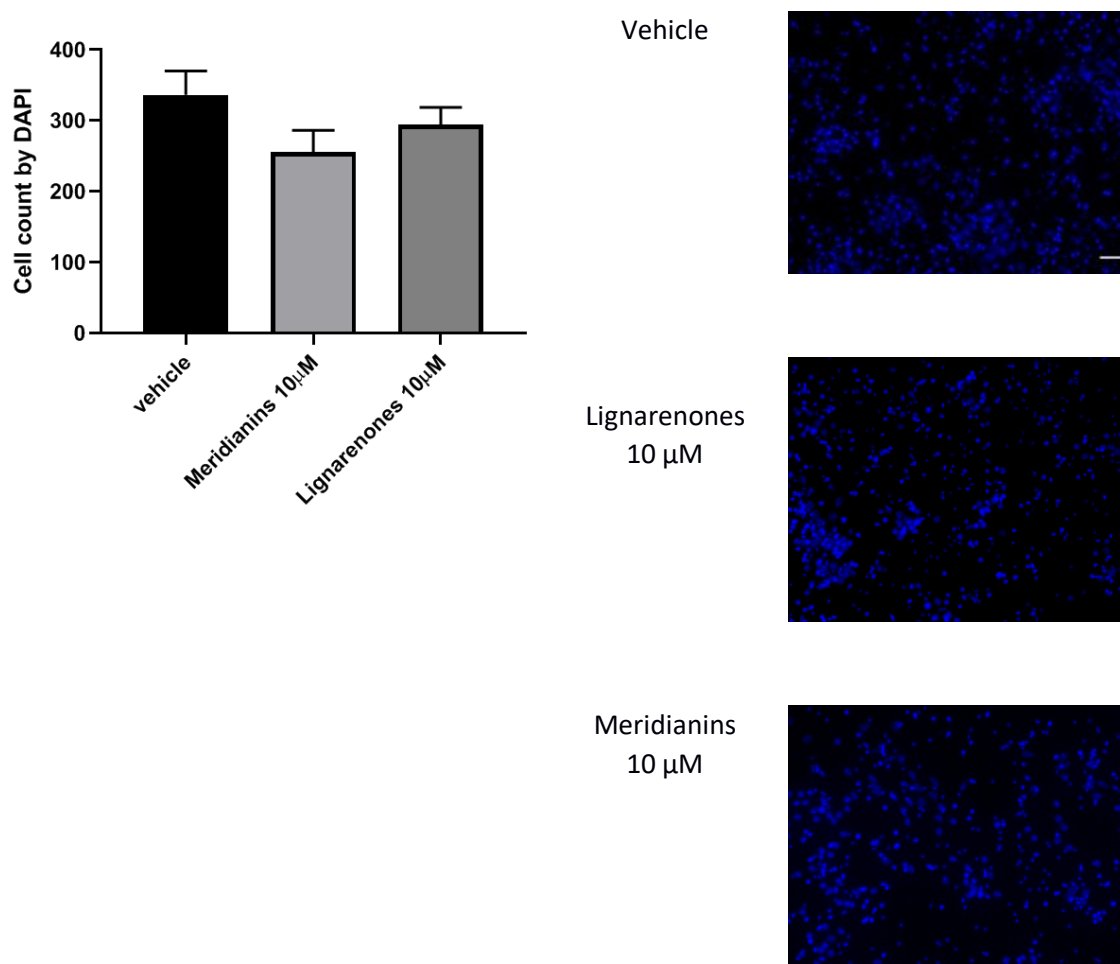


Figure 10. Cell viability is not affected by the highest doses of both meridianins and lignarenones used in the experiments. Representative DAPI images obtained by epifluorescent microscopy from primary cortical neurons in the control treatment (right). Scale bar, 50 μ m. Cell count by DAPI staining shows no significant differences between groups (One-way ANOVA analysis, $F_{2,13} = 1.600$, $p=0.2392$) (left). Data were expressed as (mean \pm S.E.M). $n = 10$ fields counted/6 coverslips per condition.

Materials and Methods

Computational

Target selection and modelling

From all the available structures of human GSK3 β in the Protein Data Bank (PDB), the 3-dimensional (3D) crystallographic structure 6B8J with the co-crystallized ligand 65C was selected, and thus the protein structure modelling from it [53,75]. The meridianins and lignarenone B structures were

modelled from the 2D chemical structure previously published by Nuñez-Pons and co-workers and Cutignano et al., respectively [43,76].

Cavity search

Fpocket software, a protein pocket prediction algorithm was used to identify different cavities on the surface of GSK3 β [77]. Fpocket uses the standard PDB files, without ligands or water molecules, as input, and using alpha spheres predicts the possible binding pockets over the protein surface, ranking them according to a scoring that involves the druggability, the size of the pocket, and its Surface Accessible Surface Area (SASA), among other properties [77].

Docking calculations

Docking calculations were performed over the two top ranked pockets in the cavity search: the ATP and the substrate pocket. The modelled protein structures and the marine molecules were employed as input of the calculations performed over each of the two cavities using Itzamna software tool [78]. More details of this approach can be found at Llorach-Pares et al. (2017 and 2018) [25,26], and Chapters 2-3 of this thesis.

Molecular dynamics simulation

Short MD simulations (1ns), using NAMD software version 2.11, were performed over the top ranked conformations obtained after docking, which were selected based on free binding energy, to post-processing them [79]. Each ligand-target complex was protonated at physiological pH 7.4 and then placed into a TIP3P water cubic box, imposing periodic boundary conditions, in which Na⁺ and Cl⁻ ions were added to neutralize the charge of the system [80]. Electrostatic interactions were calculated by the particle-mesh Ewald method using constant pressure and temperature conditions. Each complex was solvated with a minimum distance of 10Å from the surface of the complex to the edge of the simulation box. The temperature was maintained at 300 K using a Langevin thermostat, and the pressure was maintained at 1 atm using a Langevin Piston barostat. The time step employed was 2 fs. Bond lengths to hydrogens were constrained with the SHAKE algorithm [81]. Before production runs, the structure was energy minimized followed by a slow heating-up phase using harmonic position restraints on the heavy atoms of the protein. Subsequently, the system was energy minimized until volume equilibration, followed by the production run without restraints. The Amber ff99SB-ILDN and the General Amber Force Field (GAFF) set of parameters were used to model the target and the marine molecules, respectively

[82,83]. Ligand GAFF parameters were obtained using Antechamber, whereas the receptor structures were modelled using the leap module of Amber Tools [84,85].

Molecular dynamics analysis

Visual inspection to each trajectory and the HBs occupancies analysis was performed using Visual Molecular Dynamics (VMD) software [86]. Thermodynamic (temperature, potential, kinetic and total energy) and structural (Rg, RMSD and RMSF) analysis were performed using GROMACS simulation package [87,88].

Molecular Mechanics Generalized Born Surface Area

After performing MD simulations to estimate the binding free energy of GSK3 β – marine molecules complexes, Molecular Mechanics Generalized Born Surface Area (MM/GBSA) reweighting techniques were employed [89]. These techniques are widely used and outperform docking results because they are employed using the MD output trajectory as an input, thus taking into account the dynamic behaviour of the protein-ligand complexes [59,90]. MM/GBSA rescoring was performed using the MMPBSA python algorithm contained within the Amber Tools suit [91].

ADMET prediction

ADMETer, a software tool containing support vector regression (SVR) and support vector machine (SVM) ADMET machine-learning (ML) predictive models, was used to assess the ADMET properties of meridianin A-G and lignarenone B [25,92]. More precisely, the models measured logS (molecular aqueous coefficient), logP (octanol/water partition coefficient), Caco2 permeability, blood-brain barrier penetration (BBB), plasma-protein binding (PPB), P-glycoprotein binding (Pgp), Cytocrom P450 (CYPC450), organic cation transporter 2 (OCT2), and hERG mediated toxicity properties.

Graphical representations

Graphical representations were prepared using PyMOL version 1.7 and XMGRACE version 5.1.22 [93,94]. 2D images of marine molecules were prepared using RDKit python library [95].

Experimental

Cortical and striatal primary cultures

Primary cortical neuronal cultures were performed as previously described in the literature [96]. Cortex from E17.5 WT mouse embryos were dissected and gently dissociated with a fire-polished Pasteur pipette. Cells were seeded (50,000 cells/cm² for immunochemical staining and

800,000 cells/cm² for Western blot analysis) onto 24mm culture or 60 mm culture plates pre-coated with 0.1 mg/mL poly-d-lysine (Sigma Chemical Co., St. Louis, MO) and cultured in Neurobasal medium supplemented with B27 (Gibco, Paisley, Scotland, UK, 50x) and GlutaMAX (Gibco, 100x) at 37 °C in a humidified atmosphere containing 5% CO₂. For biochemical assay, 60 mm culture dishes were treated 7 days *in vitro* (DIV) with vehicle (PBS), meridianins (10 μ M), and lignarenones (10 μ M). 15 or 60 min after treatment, cells were washed with cold PBS and lysed for WB analysis. For immunocytochemical staining and morphology analyses, 24mm culture dishes were treated during 4 days *in vitro* (DIV) with vehicle (PBS), meridianins (10 μ M), and lignarenones (10 μ M), and assessed 3 days after.

Immunoblot analysis

Cell samples were collected in cold lysis buffer containing 50 mM Tris base (pH 7.5), 10mM EDTA, 1% Triton X-100, and supplemented with 1 mM sodium orthovanadate, 1mM phenylmethylsulfonyl fluoride, 1 mg/ml leupeptin and 1 mg/ml aprotinin. Samples were centrifuged at 32,000 g for 15 min and the supernatants collected. Following determination of the protein contents by Detergent-Compatible Protein Assay (Bio-Rad, Hercules, CA, USA), protein extracts (15 μ g) were mixed with 5 \times SDS sample buffer, boiled for 5 min, resolved on 15% SDS-PAGE and transferred to nitrocellulose membranes (Whatman Schleicher & Schuell, Keene, NH, USA). After incubation (1 h) in blocking buffer containing 10% non-fat powdered milk in Tris buffered saline-Tween (TBS-T) (50 mM Tris-HCl, 150 mM NaCl, pH 7.4, 0.05% Tween 20), membranes were blotted overnight at 4 °C with primary antibodies. Antibodies used for immunoblot analysis were: GSK3 β (1:1000; Cell Signaling, #9315), phosphoGSK3 β at Ser9 (1:1000; Cell Signaling, #9336xz), and α -Tubulin (1:40,000; Sigma-Aldrich, T9026). The membranes were then rinsed three times with TBS-T and incubated with horseradish peroxidase-conjugated secondary antibody for 1 h at room temperature. After washing for 30 min with TBS-T, the membranes were developed using the enhanced chemilluminescence ECL kit (Santa Cruz Biotechnology). The Gel-Pro densitometry program (Gel-Pro Analyzer for Windows, version 4.0.00.001) was used to quantify the different immunoreactive bands relative to the intensity of the α -tubulin or phospho GSK3Beta band in the same membranes within a linear range of detection for the ECL reagent.

Immunocytochemical staining

Immunochemical staining was performed following standard protocols available [97]. Briefly, primary cortical neuronal cultures were fixed at 7 days *in vitro* (DIV) in 4 % paraformaldehyde for 10 min. After fixation, cells were washed with PBS, incubated 15 min with 0,1M Glycine in PBS, used to block unreacted aldehydes after fixation, which can cause an increase in background fluorescence. After

washes with PBS, cells were permeabilized and blocked during 1h at room temperature with PBS containing 0,3% Triton X-100, 1% bovine serum albumin and 1% donkey normal serum. Then, cells were incubated overnight at 4°C in 0.1 M PBS with 5 % normal horse serum with the proper primary antibody added. MAP2 (1:1000, Sigma-Aldrich, M1406) primary antibodies was used. After primary antibody incubation, cultures were washed with PBS and incubated 2h at room temperature with Alexa Fluor 488-conjugated AffiniPure donkey anti-mouse (1:100; Jackson ImmunoResearch Laboratories, Inc., West Grove, PA). Then, coverslips mounted with Fluoromount containing DAPI onto the surface of a slide after washes with PBS. Immunofluorescence images were taken using an Olympus BX60 epifluorescence microscope, using a 20X objective.

Imaging and analysis

The *in vitro* Sholl analysis was performed with the freeware ImageJ (ImageJ, RRID:SCR 003070). We evaluated 45–55 neurons, all of them MAP2-positive from one primary cortical culture. To estimate the density of dendritic spines, 31–41 dendrites from MAP2-positive neurons (1 or 2 dendrites/neuron) from 3 different cultures were counted.

Statistical analysis

Statistical analysis was performed using one-way ANOVA with the Dunnett's *post hoc* test as appropriate and indicated in the figure legends. Data analysis and graphs were created using Graphpad Prism Software version 6.0. A 95% confidence interval was used and values of $p < 0.05$ were considered as statistically significant. Data is expressed as mean \pm S.E.M.

Conclusions

In this study we have shown that the marine natural products meridianins and lignarenone B are capable of inhibiting the activity of GSK3 β through an ATP competitive and non-competitive, allosteric, mechanism, although it seems that they are preferentially ATP-competitive inhibitors.

Docking and MD studies elucidated the binding mode of each of the studied compounds over GSK3 β . The inhibition can be caused by meridianins or lignarenone B occupying the ATP or the substrate pockets of GSK3 β , or even by the simultaneous binding on the ATP and/or the substrate active site. Experimental analysis confirmed the GSK3 β inhibition predicted *in silico*, as a consequence of growing Ser9 phosphorylation levels. During the inhibitory process the GSK3 β total levels remain stable, suggesting that neither meridianins nor lignarenone B affect their viability. Moreover, the neurite outgrowth increased, supporting the no affection of the structural plasticity. However, even with the

inhibitory behaviour confirmation, further studies are needed to completely depict the binding mode of meridianins and lignarenone B.

Results obtained from both computational and experimental studies allow us to suggest that meridianins and, to a lesser extent, lignarenone B, may inhibit GSK3 β . Therefore, these compounds could be considered as hits, constituting a starting point to develop new future potential therapeutic agents for the treatment of AD. However, they should be optimized because they are not showing good absorption and distribution profiles yet. The solubility should be improved as well as the lipophilicity and the BBB permeability, which are key issues here. Once inside the cells, these compounds have shown good inhibitory profiles and also good permeability towards the cellular membrane, but nevertheless, they should be able to penetrate into the brain. Several strategies can be employed for this, ranging from a proper modification of the chemical structure to improving the nanodelivery, including also the possibility of becoming Pgp or other protein binders that may facilitate their penetrance [98–100].

Our results from *in silico* and *in vitro* experiments constitute a promising starting point for the development of novel anti-AD drugs and further studies should be devoted to improve the specific characteristics of the studied marine compounds.

Acknowledgments: This research was partially supported by an Industrial Doctorate grant from the Generalitat of Catalonia to L.L.-P (DI 2016-051), and the BLUEBIO (CTM2016-78901/ANT) project from the Spanish government. This work is a contribution to the AntEco programme (SCAR). We thank Carlos Jiménez, Laura Núñez-Pons, and Jaime Rodríguez for support with meridianins, and Josep Maria Viñas and his “Estelada” fishing ship and crew for their help in collecting the *Scaphander* samples.

Author Contributions: L.L.-P., A.G., M.S.-M., J.A. and C.A. conceived the study and designed the experiments. L.L.-P., E.R.-U, and A.G. carried out the experiments. M.S.-M., A.N.-C., C.A. and J.A. supervised the experiments. All the authors analysed and discussed the results and elaborated the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

The physics of the all simulations has to be checked before any particular MD analysis. Also, due to its importance and its influence along the simulation, an important value to be checked after an MD is the temperature, a fundamental concept in physics which represents the intensity of the thermal motions of molecules [101]. In our study, after the analysis of all the systems, the stability of the temperature was validated with an average of 297 K \pm 3. Energies, total, potential, and kinetic, must also be taken into account as they physically validate the simulations [102,103]. At this point, the thermodynamic properties of all the complexes could be confirmed and the structural analysis could start. Radius of gyration (Rg) which allow the analysis of the compactness of the protein and is related to the tertiary structure, were analysed when each marine molecule was bound at the ATP and at the substrate binding pocket (Figure A1) [45]. Root-mean-square deviation (RMSD) is used to validate the stability along the simulations and measures the average distance between the atoms of superimposed structures extracted from the MD simulations and a reference structure (Figure A2) [104–106]. The behaviour of each simulation was checked obtaining positive results.

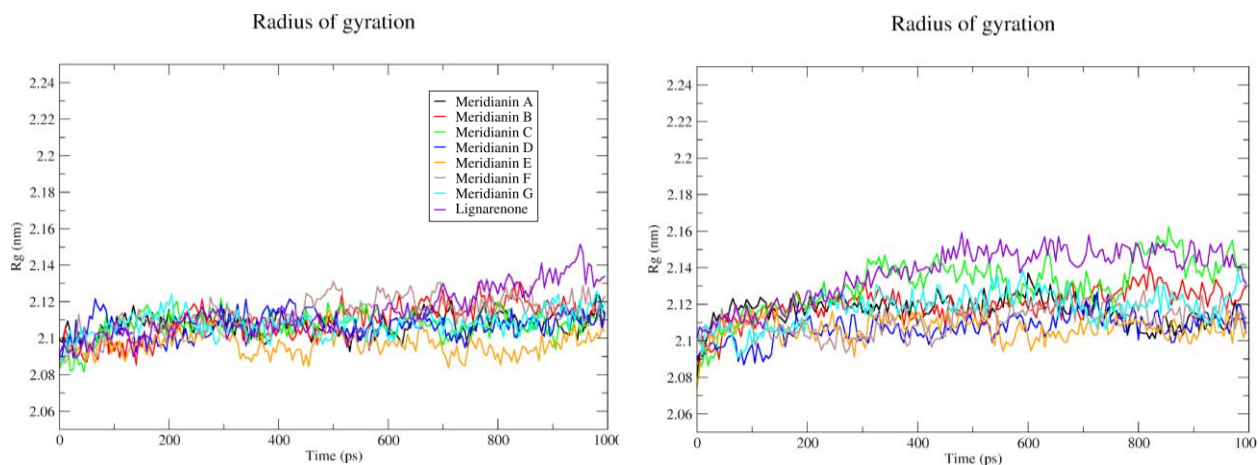


Figure A1. Time evolution of the Radius of gyration (Rg) obtained for each system. On the left, the systems with each marine molecule bound to the ATP cavity, and on the right, when the marine molecules are bound to the substrate pocket. The colour code for each system can be seen in the legend box.

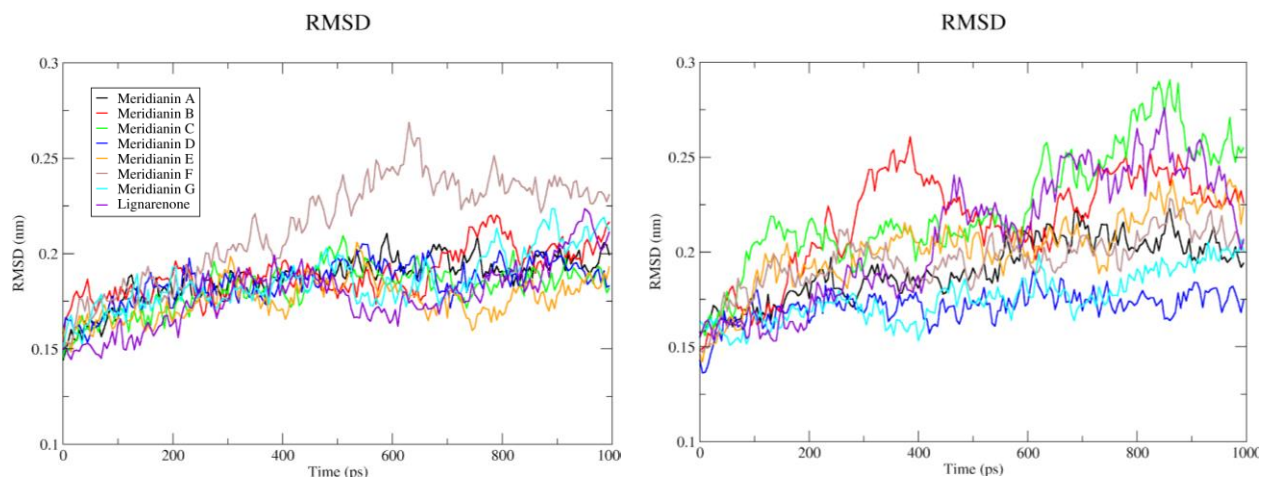


Figure A2. Time evolution of the root-mean-square deviation (RMSD) of the backbone atoms during the molecular dynamics (MD) simulation with respect to the initial structure of GSK3 β . On the left, the RMSD of each system where marine molecules are bound to the ATP cavity, and on the right when the marine molecules are bound to the substrate pocket. The colour code for each system can be seen in the legend box.

References

1. Hanger, D.P.; Hughes, K.; Woodgett, J.R.; Brion, J.-P.; Anderton, B.H. Glycogen synthase kinase-3 induces Alzheimer's disease-like phosphorylation of tau: Generation of paired helical filament epitopes and neuronal localisation of the kinase. *Neurosci. Lett.* 1992, 147, 58–62.
2. Mandelkow, E.-M.; Drewes, G.; Biernat, J.; Gustke, N.; Van Lint, J.; Vandenheede, J.R.; Mandelkow, E. Glycogen synthase kinase-3 and the Alzheimer-like state of microtubule-associated protein tau. *FEBS Lett.* 1992, 314, 315–321.
3. Hooper, C.; Killick, R.; Lovestone, S. The GSK3 hypothesis of Alzheimer's disease. *J. Neurochem.* 2008, 104, 1433–9.
4. Ding, V.W.; Chen, R.H.; McCormick, F. Differential regulation of glycogen synthase kinase 3 β by insulin and Wnt signaling. *J. Biol. Chem.* 2000, 275, 32475–81.
5. Palsgaard, J.; Emanuelli, B.; Winnay, J.N.; Sumara, G.; Karsenty, G.; Kahn, C.R. Cross-talk between insulin and Wnt signaling in preadipocytes: role of Wnt co-receptor low density lipoprotein receptor-related protein-5 (LRP5). *J. Biol. Chem.* 2012, 287, 12016–26.
6. Li, D.-W.; Liu, Z.-Q.; Chen, W.; Yao, M.; Li, G.-R. Association of glycogen synthase kinase-3 β with Parkinson's disease. *Mol. Med. Rep.* 2014, 9, 2043–50.
7. Llorens-Martín, M.; Jurado, J.; Hernández, F.; Avila, J. GSK-3 β , a pivotal kinase in Alzheimer disease. *Front. Mol. Neurosci.* 2014, 7, 46.
8. Lim, N.K.H.; Hung, L.W.; Pang, T.Y.; Mclean, C.A.; Liddell, J.R.; Hilton, J.B.; Li, Q.-X.; White, A.R.; Hannan, A.J.; Crouch, P.J. Localized changes to glycogen synthase kinase-3 and collapsin response mediator protein-2 in the Huntington's disease affected brain. *Hum. Mol. Genet.* 2014, 23, 4051–4063.
9. Kremer, A.; Louis, J. V; Jaworski, T.; Van Leuven, F. GSK3 and Alzheimer's Disease: Facts and Fiction.... *Front. Mol. Neurosci.* 2011, 4, 17.
10. Goedert, M.; Klug, A.; Crowther, R.A. Tau protein, the paired helical filament and Alzheimer's disease. *J. Alzheimers. Dis.* 2006, 9, 195–207.
11. Martin, L.; Latypova, X.; Wilson, C.M.; Magnaudeix, A.; Perrin, M.-L.; Yardin, C.; Terro, F. Tau protein kinases: Involvement in Alzheimer's disease. *Ageing Res. Rev.* 2013, 12, 289–309.
12. Kolarova, M.; García-Sierra, F.; Bartos, A.; Ricny, J.; Ripova, D.; Ripova, D. Structure and Pathology of Tau Protein in Alzheimer Disease. *Int. J. Alzheimers. Dis.* 2012, 2012, 1–13.
13. Kametani, F.; Hasegawa, M. Reconsideration of Amyloid Hypothesis and Tau Hypothesis in Alzheimer's Disease. *Front. Neurosci.* 2018, 12, 25.
14. Doig, A.J.; del Castillo-Frias, M.P.; Berthoumieu, O.; Tarus, B.; Nasica-Labouze, J.; Sterpone, F.; Nguyen, P.H.; Hooper, N.M.; Faller, P.; Derreumaux, P. Why Is Research on Amyloid- β Failing to Give New Drugs for Alzheimer's Disease? *ACS Chem. Neurosci.* 2017, 8, 1435–1437.
15. del Ser, T.; Steinwachs, K.C.; Gertz, H.J.; Andrés, M. V.; Gómez-Carrillo, B.; Medina, M.; Vericat, J.A.; Redondo, P.; Fleet, D.; León, T. Treatment of Alzheimer's Disease with the GSK-3 Inhibitor Tideglusib: A Pilot Study. *J. Alzheimer's Dis.* 2012, 33, 205–215.
16. Cummings, J.; Lee, G.; Ritter, A.; Zhong, K. Alzheimer's disease drug development pipeline: 2018. *Alzheimer's Dement. (New York, N. Y.)* 2018, 4, 195–214.
17. Citron, M. Alzheimer's disease: strategies for disease modification. *Nat. Rev. Drug Discov.* 2010, 9, 387–398.
18. Palomo, V.; Soteras, I.; Perez, D.I.; Perez, C.; Gil, C.; Eugenia Campillo, N.; Martinez, A. Exploring the Binding Sites of Glycogen Synthase Kinase 3. Identification and Characterization of Allosteric Modulation Cavities. *J. Med. Chem.* 2011, 54, 8461–8470.
19. Walz, A.; Ugolkov, A.; Chandra, S.; Kozikowski, A.; Carneiro, B.A.; O'Halloran, T. V; Giles, F.J.; Billadeau, D.D.; Mazar, A.P. Molecular Pathways: Revisiting Glycogen Synthase Kinase-3 β as a

- Target for the Treatment of Cancer. *Clin. Cancer Res.* 2017, 23, 1891–1897.
20. Pandey, M.K.; DeGrado, T.R. Glycogen Synthase Kinase-3 (GSK-3)-Targeted Therapy and Imaging. *Theranostics* 2016, 6, 571–93.
 21. Martinez, A.; Gil, C.; Perez, D.I. Glycogen Synthase Kinase 3 Inhibitors in the Next Horizon for Alzheimer's Disease Treatment. *Int. J. Alzheimers. Dis.* 2011, 2011, 1–7.
 22. Russo, P.; Kisialiou, A.; Lamonaca, P.; Moroni, R.; Prinzi, G.; Fini, M. New Drugs from Marine Organisms in Alzheimer's Disease. *Mar. Drugs* 2015, 14, 5.
 23. Ansari, N.; Khodagholi, F. Natural Products as Promising Drug Candidates for the Treatment of Alzheimer's Disease: Molecular Mechanism Aspect. *Curr. Neuropharmacol.* 2013, 11, 414–429.
 24. Dey, A.; Bhattacharya, R.; Mukherjee, A.; Pandey, D.K. Natural products against Alzheimer's disease: Pharmacotherapeutics and biotechnological interventions. *Biotechnol. Adv.* 2017, 35, 178–216.
 25. Llorach-Pares, L.; Nonell-Canals, A.; Sanchez-Martinez, M.; Avila, C. Computer-Aided Drug Design Applied to Marine Drug Discovery: Meridianins as Alzheimer's Disease Therapeutic Agents. *Mar. Drugs* 2017, 15, 366.
 26. Llorach-Pares, L.; Nonell-Canals, A.; Avila, C.; Sanchez-Martinez, M. Kororamides, Convolutamines, and Indole Derivatives as Possible Tau and Dual-Specificity Kinase Inhibitors for Alzheimer's Disease: A Computational Study. *Mar. Drugs* 2018, 16, 386.
 27. Eldar-Finkelman, H.; Martinez, A. GSK-3 Inhibitors: Preclinical and Clinical Focus on CNS. *Front. Mol. Neurosci.* 2011, 4, 32.
 28. Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on natural products for drug design. *Nat. Chem.* 2016, 8, 531–541.
 29. W-H Li, J.; Vederas, J.C. Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science* (80-.). 2009, 325, 161–165.
 30. Harvey, A.L.; Edrada-Ebel, R.; Quinn, R.J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* 2015, 14, 111–129.
 31. Prachayasittikul, V.; Worachartcheewan, A.; Shoombuatong, W.; Songtawee, N.; Simeon, S.; Prachayasittikul, V.; Nantasenamat, C. Computer-Aided Drug Design of Bioactive Natural Products. *Curr. Top. Med. Chem.* 2015, 15, 1780–800.
 32. Montaser, R.; Luesch, H. Marine natural products: a new wave of drugs? *Future Med. Chem.* 2011, 3, 1475–1489.
 33. Blunt, J.W.; Copp, B.R.; Hu, W.-P.; Munro, M.H.G.; Northcote, P.T.; Prinsep, M.R.; Grabley, S.; Yoon, H.D.; Chang, Y.-C.; Hong, Y.-K.; et al. Marine natural products. *Nat. Prod. Rep.* 2007, 24, 31.
 34. Patridge, E.; Gareiss, P.; Kinch, M.S.; Hoyer, D. An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov. Today* 2016, 21, 204–207.
 35. Meijer, L.; Thunnissen, A.-M.; White, A.; Garnier, M.; Nikolic, M.; Tsai, L.-H.; Walter, J.; Cleverley, K.; Salinas, P.; Wu, Y.-Z.; et al. Inhibition of cyclin-dependent kinases, GSK-3 β and CK1 by hymenialdisine, a marine sponge constituent. *Chem. Biol.* 2000, 7, 51–63.
 36. Khanfar, M.A.; Asal, B.A.; Mudit, M.; Kaddoumi, A.; El Sayed, K.A. The marine natural-derived inhibitors of glycogen synthase kinase-3 β phenylmethylene hydantoin: *In vitro* and *in vivo* activities and pharmacophore modeling. *Bioorg. Med. Chem.* 2009, 17, 6032–9.
 37. Bharate, S.B.; Yadav, R.; Battula, S.; Vishwakarma, R.. Meridianins: Marine-Derived Potent Kinase Inhibitors. *Mini-Reviews Med. Chem.* 2012, 12, 618–631.
 38. Gerwick, W.H.; Moore, B.S. Lessons from the Past and Charting the Future of Marine Natural Products Drug Discovery and Chemical Biology. *Chem. Biol.* 2012, 19, 85–98.
 39. Haefner, B. Drugs from the deep: marine natural products as drug candidates. *Drug Discov. Today* 2003, 8, 536–544.
 40. Molinski, T.F.; Dalisay, D.S.; Lievens, S.L.; Saludes, J.P. Drug development from marine natural

- products. *Nat. Rev. Drug Discov.* 2009, 8, 69–85.
41. Llorach-Pares, L.; Nonell-Canals, A.; Sanchez-Martínez, M.; Avila, C. *In silico* studies to find new therapeutic indications for marine molecules. *Molecular docking in marine drug discovery and design*. prep.
 42. Núñez-Pons, L.; Forestieri, R.; Nieto, R.M.; Varela, M.; Nappo, M.; Rodríguez, J.; Jiménez, C.; Castelluccio, F.; Carbone, M.; Ramos-Espla, A.; et al. Chemical defenses of tunicates of the genus *Aplidium* from the Weddell Sea (Antarctica). *Polar Biol.* 2010, 33, 1319–1329.
 43. Cutignano, A.; Avila, C.; Domenech-Coll, A.; d'Ippolito, G.; Cimino, G.; Fontana, A. First Biosynthetic Evidence on the Phenyl-Containing Polyketides of the Marine Mollusc *Scaphander lignarius*. *Org. Lett.* 2008, 10, 2963–2966.
 44. Martínez, A.; Alonso, M.; Castro, A.; Concepción Pérez, A.; Moreno, F.J. First Non-ATP Competitive Glycogen Synthase Kinase 3 β (GSK-3 β) Inhibitors: Thiadiazolidinones (TDZD) as Potential Drugs for the Treatment of Alzheimer's Disease. 2002.
 45. Palomo, V.; Perez, D.I.; Roca, C.; Anderson, C.; Rodríguez-Muela, N.; Perez, C.; Morales-Garcia, J.A.; Reyes, J.A.; Campillo, N.E.; Perez-Castillo, A.M.; et al. Subtly Modulating Glycogen Synthase Kinase 3 β : Allosteric Inhibitor Development and Their Potential for the Treatment of Chronic Diseases. *J. Med. Chem.* 2017, 60, 4983–5001.
 46. Bidon-Chanal, A.; Fuertes, A.; Alonso, D.; Pérez, D.I.; Martínez, A.; Luque, F.J.; Medina, M. Evidence for a new binding mode to GSK-3: Allosteric regulation by the marine compound palinurin. *Eur. J. Med. Chem.* 2013, 60, 479–489.
 47. Borer, B.C.; Taylor, R.J.K. ChemInform Abstract: The Synthesis of Lignarenone A and Lignarenone B, Metabolites of the Mediterranean Opisthobranch Mollusc *Scaphander lignarius*. *ChemInform* 1990, 21, no-no.
 48. Bell, P.T.; Donaldson, W.A. Synthesis of Lignarenone B. *J. Nat. Prod.* 1992, 55, 1669–1671.
 49. Tibiletti, F.; Simonetti, M.; Nicholas, K.M.; Palmisano, G.; Parravicini, M.; Imbesi, F.; Tollari, S.; Penoni, A. One-pot synthesis of meridianins and meridianin analogues via indolization of nitrosoarenes. *Tetrahedron* 2010, 66, 1280–1288.
 50. Walker, S.R.; Czyz, M.L.; Morris, J.C. Concise syntheses of meridianins and meriolins using a catalytic domino amino-palladation reaction. *Org. Lett.* 2014, 16, 708–711.
 51. Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics* 2009, 10, 168.
 52. Burley, S.K.; Berman, H.M.; Bhikadiya, C.; Bi, C.; Chen, L.; Di Costanzo, L.; Christie, C.; Dalenberg, K.; Duarte, J.M.; Dutta, S.; et al. RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.* 2019, 47, D464–D474.
 53. Wagman, A.S.; Boyce, R.S.; Brown, S.P.; Fang, E.; Goff, D.; Jansen, J.M.; Le, V.P.; Levine, B.H.; Ng, S.C.; Ni, Z.-J.; et al. Synthesis, Binding Mode, and Antihyperglycemic Activity of Potent and Selective (5-Imidazol-2-yl-4-phenylpyrimidin-2-yl)[2-(2-pyridylamino)ethyl]amine Inhibitors of Glycogen Synthase Kinase 3. *J. Med. Chem.* 2017, 60, 8482–8514.
 54. Pérot, S.; Sperandio, O.; Miteva, M.A.; Camproux, A.-C.; Villoutreix, B.O. Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. *Drug Discov. Today* 2010, 15, 656–667.
 55. Schmidtke, P.; Barril, X. Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. *J Med Chem* 2010, 53, 5858–5867.
 56. Hamann, M.; Alonso, D.; Martín-Aparicio, E.; Fuertes, A.; Pérez-Puerto, M.J.; Castro, A.; Morales, S.; Navarro, M.L.; del Monte-Millán, M.; Medina, M.; et al. Glycogen Synthase Kinase-3 (GSK-3) Inhibitory Activity and Structure–Activity Relationship (SAR) Studies of the Manzamine Alkaloids. Potential for Alzheimer's Disease. *J. Nat. Prod.* 2007, 70, 1397–1405.

57. Osolodkin, D.I.; Palyulin, V.A.; Zefirov, N.S. Glycogen Synthase Kinase 3 as an Anticancer Drug Target: Novel Experimental Findings and Trends in the Design of Inhibitors. *Curr. Pharm. Des.* 2013, 19, 665–679.
58. Alonso, H.; Bliznyuk, A.A.; Gready, J.E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* 2006, 26, 531–568.
59. Vivo, M. De; Masetti, M.; Bottegoni, G.; Cavalli, A. The Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* 2016.
60. De Vivo, M.; Cavalli, A. Recent advances in dynamic docking for drug discovery. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2017, 7, e1320.
61. Zhang, X.; Perez-Sanchez, H.; Lightstone, F.C. A Comprehensive Docking and MM/GBSA Rescoring Study of Ligand Recognition upon Binding Antithrombin. *Curr. Top. Med. Chem.* 2017, 17, 1631–1639.
62. Fuglebakk, E.; Echave, J.; Reuter, N. Measuring and comparing structural fluctuation patterns in large protein datasets. *Bioinformatics* 2012, 28, 2431–2440.
63. Bornot, A.; Etchebest, C.; de Brevern, A.G. Predicting protein flexibility through the prediction of local structures. *Proteins* 2011, 79, 839–52.
64. Kempner, E.S. Movable lobes and flexible loops in proteins. Structural deformations that control biochemical activity. *FEBS Lett.* 1993, 326, 4–10.
65. Ilouz, R.; Kowalsman, N.; Eisenstein, M.; Eldar-Finkelman, H. Identification of novel glycogen synthase kinase-3 β substrate-interacting residues suggests a common mechanism for substrate recognition. *J. Biol. Chem.* 2006, 281, 30621–30.
66. Ruiz-Garcia, A.; Bermejo, M.; Moss, A.; Casabo, V.G. Pharmacokinetics in Drug Discovery. *J. Pharm. Sci.* 2008, 97, 654–690.
67. Alqahtani, S. *In silico* ADME-Tox modeling: progress and prospects. *Expert Opin. Drug Metab. Toxicol.* 2017, 13, 1147–1158.
68. Hansch, C.; Björkroth, J.P.; Leo, A. Hydrophobicity and central nervous system agents: on the principle of minimal hydrophobicity in drug design. *J. Pharm. Sci.* 1987, 76, 663–87.
69. Wang, W.; Bodles-Brakhop, A.M.; Barger, S.W. A Role for P-Glycoprotein in Clearance of Alzheimer Amyloid β -Peptide from the Brain. *Curr. Alzheimer Res.* 2016, 13, 615–20.
70. Miller, D.S.; Bauer, B.; Hartz, A.M.S. Modulation of P-glycoprotein at the blood-brain barrier: opportunities to improve central nervous system pharmacotherapy. *Pharmacol. Rev.* 2008, 60, 196–209.
71. Chang, K.L.; Pee, H.N.; Yang, S.; Ho, P.C. Influence of drug transporters and stereoselectivity on the brain penetration of pioglitazone as a potential medicine against Alzheimer’s disease. *Sci. Rep.* 2015, 5, 9000.
72. Cirrito, J.R.; Deane, R.; Fagan, A.M.; Spinner, M.L.; Parsadanian, M.; Finn, M.B.; Jiang, H.; Prior, J.L.; Sagare, A.; Bales, K.R.; et al. P-glycoprotein deficiency at the blood-brain barrier increases amyloid-beta deposition in an Alzheimer disease mouse model. *J. Clin. Invest.* 2005, 115, 3285–90.
73. Roden, D.M.; Viswanathan, P.C. Genetics of acquired long QT syndrome. *J. Clin. Invest.* 2005, 115, 2025–32.
74. Vasskog, T.; Andersen, J.H.; Hansen, E.; Svenson, J. Characterization and Cytotoxicity Studies of the Rare 21:4 n-7 Acid and Other Polyunsaturated Fatty Acids from the Marine Opisthobranch *Scaphander lignarius*, Isolated Using Bioassay Guided Fractionation. *Mar. Drugs* 2012, 10, 2676.
75. Desaphy, J.; Bret, G.; Rognan, D.; Kellenberger, E. sc-PDB: a 3D-database of ligandable binding sites--10 years on. *Nucleic Acids Res.* 2014, 43, D399--D404.
76. Núñez-Pons, L.; Carbone, M.; Vázquez, J.; Rodríguez, J.; Nieto, R.M.; Varela, M.M.; Gavagnin, M.; Avila, C.; Núñez-Pons, L.; Carbone, M.; et al. Natural Products from Antarctic Colonial Ascidians of

- the Genera *Aplidium* and *Synoicum*: Variability and Defensive Role. *Mar. Drugs* 2012, 10, 1741–1764.
77. Schmidtke, P.; Le Guilloux, V.; Maupetit, J.; Tuffeery, P. fpocket: Online tools for protein ensemble pocket detection and tracking. *Nucleic Acids Res.* 2010, 38.
 78. Felix, E.; Santamaría-Navarro, E.; Sanchez-Martinez, M.; Nonell-Canals, A. Itzamna Available online: <https://www.mindthebyte.com/>.
 79. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 2005, 26, 1781–1802.
 80. Jorgensen, W.L.; Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *J. Comput. Chem.* 1998, 19, 1179–1186.
 81. Andersen, H.C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* 1983, 52, 24–34.
 82. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 2010, 78, 1950–8.
 83. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* 2004, 25, 1157–1174.
 84. Wang, J.; Wang, W.; Kollman, P.A.; Case, D.A. Antechamber, An Accessory Software Package For Molecular Mechanical Calculations. *Natl. Institutes Heal.*
 85. Case, D.A.; Cheatham, T.E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* 2005, 26, 1668–88.
 86. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Molec. Graph.* 1996, 14, 33–38.
 87. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* 2008, 4, 435–447.
 88. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M.R.; Smith, J.C.; Kasson, P.M.; Van Der Spoel, D.; et al. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 2013, 29, 845–854.
 89. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 2015, 10, 449–61.
 90. Rastelli, G.; Degliesposti, G.; Del Rio, A.; Sgobba, M. Binding Estimation after Refinement, a New Automated Procedure for the Refinement and Rescoring of Docked Ligands in Virtual Screening. *Chem. Biol. Drug Des.* 2009, 73, 283–286.
 91. Miller III, B.R.; McGee Jr., T.D.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA. py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* 2012, 3314–3321.
 92. Vidal, D.; Nonell-Canals, A. ADMET models Available online: <https://www.mindthebyte.com/>.
 93. Yuan, S.; Chan, H.C.S.; Hu, Z. Using PyMOL as a platform for computational drug design. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2017, 7, e1298.
 94. Grace User’s Guide (for Grace-5.1.22) Available online: <http://plasma-gate.weizmann.ac.il/Grace/doc/UsersGuide.html> (accessed on Apr 26, 2019).
 95. RDKit: Open-source cheminformatics.
 96. Anglada-Huguet, M.; Vidal-Sancho, L.; Giral, A.; García-Díaz Barriga, G.; Xifró, X.; Alberch, J. Prostaglandin E2 EP2 activation reduces memory decline in R6/1 mouse model of Huntington’s disease by the induction of BDNF-dependent synaptic plasticity. *Neurobiol. Dis.* 2016, 95, 22–34.

97. Xifró, X.; García-Martínez, J.M.; del Toro, D.; Alberch, J.; Pérez-Navarro, E. Calcineurin is involved in the early activation of NMDA-mediated cell death in mutant huntingtin knock-in striatal cells. *J. Neurochem.* 2008, 105, 1596–1612.
98. Kingwell, K. Drug delivery: New targets for drug delivery across the BBB. *Nat. Rev. Drug Discov.* 2016, 15, 84–85.
99. Banks, W.A. From blood–brain barrier to blood–brain interface: new opportunities for CNS drug delivery. *Nat. Rev. Drug Discov.* 2016, 15, 275–292.
100. Saraiva, C.; Praça, C.; Ferreira, R.; Santos, T.; Ferreira, L.; Bernardino, L. Nanoparticle-mediated brain drug delivery: Overcoming blood–brain barrier to treat neurodegenerative diseases. *J. Med. Chem.* 2016.
101. Liu, B.; Xu, R.; He, X. Kinetic Energy-Based Temperature Computation in Non-Equilibrium Molecular Dynamics Simulation. 2009.
102. Merz, P.T.; Shirts, M.R. Testing for physical validity in molecular simulations. *PLoS One* 2018, 13, e0202764.
103. Jepps; Ayton; Evans Microscopic expressions for the thermodynamic temperature. *Phys. Rev. E. Stat. Phys. Plasmas. Fluids. Relat. Interdiscip. Topics* 2000, 62, 4757–63.
104. Kufareva, I.; Abagyan, R. Methods of protein structure comparison. *Methods Mol. Biol.* 2012, 857, 231–57.
105. Maiorov, V.N.; Crippen, G.M. Significance of Root-Mean-Square Deviation in Comparing Three-dimensional Structures of Globular Proteins. *J. Mol. Biol.* 1994, 235, 625–634.
106. Yu, W.; Mackerell, A.D. Computer-Aided Drug Design Methods. *Methods Mol. Biol.* 2017, 1520, 93–94.

General Discussion

General discussion

Marine benthic invertebrates possess very interesting natural products, with a huge potential for drug discovery (Leal et al., 2012; Avila, 2016; Blunt et al., 2018a). In particular, organisms from underexplored areas, such as Antarctica, hide an enormous arsenal of chemodiversity that is starting to be untapped only recently (Barre, 2010; Núñez-Pons et al., 2015; Blunt et al., 2018a). One of the main problems of these studies is to decide which potential bioactivity or therapeutic use can a new compound acquire, since the limited amount of material usually prevents a wide bioactivity screening. For these reasons, using computational methods is paramount to decide the potential activities to check, given a particular molecular structure. The use of these computational techniques and therefore, the results obtained on the study of the possible therapeutic activity and protein function modulation capability in pathogenesis-related pathways of MNPs from Antarctic and Mediterranean benthic communities, allow covering different steps of the drug discovery pipeline. In fact, along this thesis, I used many different CADD tools and techniques, which were also comparatively studied and implemented. This thesis highlights the **utility of CADD in the process of drug development**, giving valuable insights about the process itself (Chapter 1), about the possible pharmacological properties of a group of MNPs (Chapter 2 and 3), and about allosteric modulation as a possible solution in the search for a treatment of AD (Chapter 4). Additionally, our studies on natural products from ascidians provide **new evidences on the potential therapeutic condition** of these secondary metabolites against a group of kinases related with the tau hyperphosphorylation on Alzheimer's pathology (Chapter 2, 3 and 4). Two of the most important contribution of this thesis are first, **the applicability of an established pipeline step procedure**, on the *in silico* drug discovery process of MNPs, and second, the computational elucidation and subsequent *in vitro* **validation of the inhibitory activity** of meridianin A-G and lignarenone B against GSK3 β .

Here, I discuss the most important findings, which are organized in four sections, providing a comprehensive and general overview of the results obtained along these years in the drug discovery field. The most relevant insights about the results obtained within a general perspective are commented. Recommendations and limitations are treated too on every section, while future perspectives are exposed on the last section of this chapter.

CADD potential in drug discovery

Development of new drugs is a complex process, and according to that, computational approaches are versatile tools which facilitate and accelerate the drug design and development (Prachayasittikul et al., 2015). In my opinion, the best definition of CADD, from an structural point of view, was that by Baig et al., who defined CADD as the **methods used with the aim of simulating interactions between receptors and drugs in order to determine binding affinities** (Baig et al., 2017). The utilization of these methods on MNP is not new (Medina-Franco, 2013; Pereira & Aires-de-Sousa, 2018), but the potential and usefulness of CADD techniques applied to MNP has become, in general, more clear along this thesis. One of the main advantages here is to avoid **wasting natural resources**, which often is not valued enough in the literature. Collecting samples, which are exclusively present in marine habitats, requires huge efforts from an economic, human, technical, and biological perspective, and moreover most of these organisms are usually unculturable (Molinski et al., 2009a). Therefore, the contribution of CADD techniques is very relevant, since no samples are required for any computational analysis. Moreover, CADD provides certain knowledge about the chemistry of the molecules which is unreachable only through *in vitro* experiments, reducing the cost and time, and improving the quality of the research (Macalino et al., 2015).

There are several examples of successful CADD application in different areas and diverse pathologies (Marshall, 1987; Propst & Perun, 1989; Song et al., 2009; Talele et al., 2010; Das, 2017). In this thesis, we focussed on the discovery and design of new compounds for the treatment of AD.

It is an evidence that CADD applications, as any computational prediction, have **limitations**. Most CADD techniques, such the employed in this thesis as, docking, QSAR, VS, VP, pharmacophore modelling, or MD, are based on pre-defined algorithms and scripts, which restrict their confidence and performance. Algorithms updates and high experimental data validations are key aspects to overcome these limitations and to improve accuracy in CADD predictions.

The most common failures on drug discovery rely on the **wrong prediction of ligand-protein complexes**. This “misdocked” predictions can be caused by different aspects but the most probable is the erroneous energy determination or prediction for ligand, protein, or both structures (Verkhivker et al., 2000; Ramírez & Caballero, 2016, 2018). The most frequent validation to overcome these drawbacks, is the **reproducibility** of the correct bound complex (redocking) and the assignment of correct scores values to the best dockings, to establish some correlations between the poses, the known natural ligands, if it is the case, and the measured affinities of the lead (Korb et al., 2012; Talevi, 2018). Along

this thesis, all the docking experiments were performed at least twice, following the redocking principle. With the aim to reduce failure and validate docking predictions, the most typically, extensively, and routine approach is the use of MD simulations, which are used as post-processing techniques, and allow the understanding of the protein motions and the exploration of the conformational landscape (Karplus & Kuriyan, 2005; Sakano et al., 2016). Despite MD simulations are accurate methods, they also present drawbacks. Their major limitations are the **time scale** and the required **refinement of the force fields** (Durrant et al., 2011). MD simulations are usually performed in the order of nanoseconds (ns) to microseconds (μ s). Nevertheless, to detect biological conformational changes on the protein folding, dynamics in the order of milliseconds (ms) are needed (Leelananda et al., 2016). Simulating at this time scale is very time consuming a huge computational time is required. Several methods have been recently introduced on the field with the aim to better explore the conformational space and with restricted time, such as accelerated MD (aMD), metadynamics, replica exchange MD (REMD), and umbrella sampling, among others (Bernardi et al., 2015). However, depending on the study, much smaller time steps are enough. As a docking post-processing tool, few ns or even ps are necessary. In the literature there are different streams that postulate that certain combination of time steps and number of replicas is better than others (Hou et al., 2011; Genheden et al., 2015; Sun et al., 2018). After several tests, in this thesis we decided to employ MD simulations of 1ns of duration. This duration is enough to see how a bad docking pose leaves the binding pocket, in fact, a few hundreds of ps are sufficient. Based on these short simulations, binding free energies and/or binding affinities are estimated. These values can be later used to rank the analysed compounds, helping on the elucidation of the best fit compound to a specific target. Different approaches can be used to infer the binding energy, despite nowadays the most accurate technique seems to be free energy perturbation (FEP), and thermodynamic integration (TI). Along this thesis MM/GBSA calculations are used due to their good balance between accuracy, computational power and time, to compute the binding free energies. This methodology has been widely used for decades and its successful prediction of binding energies and good ranking between compounds is more than demonstrated (Mulakala et al., 2013; Zhang et al., 2017).

Another major limitation in the *in silico* drug discovery field is the **absence of the target structure**, which is not yet experimentally solved. This fact does not allow the exploration of all the proteome (Barril, 2017), although there are techniques like HM used to alleviate this kind of problems. However, despite the advances in structure determination techniques, both experimentally and computationally, nowadays there are a lot of proteins that can not be explored computationally (Barril,

2017). Due to this fact, among others, we have to highlight that we still do not know and are far to **understand the etiology** of most diseases (Gonzalez & Kann, 2012).

Furthermore, ADMET prediction tools also need to be reviewed, despite the continuous improvements and the evolution undergone during the last decades. Solid experimental data are needed because good data is the basis of any good model. The most effective, well established and cost-effective approach to decrease the failure on later stages on the drug discovery process before the molecule is synthesized, and increase safety issues, is to apply **computational ADMET properties prediction in early stages** (Clark, 2005). Computational ADMET prediction helps a lot to reduce experimental **bad ADMET prediction**, responsible of drug failure in later stage of the drug discovery process, but with better model the help will be much higher. Most ADMET methods are based on the similarity principle (which means similar compound, similar behaviour), but today it is known that **toxic effects** can be caused by interactions with other proteins, where the predicted lead compound binds to a similar catalytic cavity but not to the predicted target, this fact is known as off-target (Rudmann, 2013).

To select the best predictive ADMET parameter models, it is crucial to select the right mathematical approach, the best molecular descriptors for a particular ADMET property, and the appropriate size of the set of experimental data related to a particular ADMET property, for the correct validation of the model (Van de Waterbeemd et al., 2003). Along this thesis, **Chapters 1-4**, there is a section describing the PK properties of each molecule studied at every stage. In **Chapter 1** the toxicology prediction was assessed using “classical” QSAR models, but nevertheless, in **Chapters 2-4**, due to the great boom in the last years and its relevance, **machine learning** (ML) based QSAR methods, by using ADMETer and pkCSM software tools (Pires et al., 2015; Vidal & Nonell-Canals, 2017). ML is based on the construction of computational models that can improve discovery and decision making from high-quality data (Vamathevan et al., 2019). Moreover, ML can be applied to develop models to predict chemical properties as absorption, distribution, metabolism, excretion, and toxicity (ADMET) (Heikamp & Bajorath, 2014; Lavecchia, 2015). ADMETer is a software tool containing **support vector machines** (SVMs) models for the prediction of Caco2 permeability, BBB, and PPB, as well as, **support vector regression** (SVR) algorithms for the estimation of LogP and LogS. In recent years, due to their accurate and consistent prediction, SVMs and to a less extent, SVR, have become increasingly popular in the drug discovery field, and are supervised ML algorithms for classification and regression-based prediction of property values as ADMET properties (Hou et al.; Clark, 2005; Shen et al., 2010). The advantages of these predictive models to determine ADMET properties is due to the fact that are highly trustful, but as

any computational model, there are also some limitations. Probably, the main limitation of ML models is the applicability domain, as they rely on the available data sets to generate and train appropriate models, so the prediction only occurs within known frameworks of the training data (Vamathevan et al., 2019).

Protein kinase inhibitors and MNPs pharmacophoric properties

The human kinome, all the catalogued protein kinases encoded by the human genome, includes a total of 518 proteins divided into seven subfamilies (Manning et al., 2002). Considering the study carried out on **Chapter 1**, where we were able to identify the capability of a group of marine molecules to bind proteins related to various pathologies. Concretely, we see a trend that point out that the MNPs studies had for targeting protein kinases, large family of signalling enzymes, as proposed also by other authors (Marston, 2011; Skropeta et al., 2011). From these initial data, our results support the **hypothesis of the potential of MNPs to act as inhibitors of protein kinases**. In that sense, in our study (Chapter 1), we founds that these protein kinases were bound by aplicyanin and meridianin A, two molecules that shared a very similar scaffold. Despite the link between meridianins and protein kinases is not new (Giraud et al., 2011; Bharate et al., 2012), an approach based on structural target perspective was never been carried out before. Hence, a first **validation of the uses and applicability of CADD applied to MNPs** was performed here. In order to evaluate the real capability of meridianin A-G, a computational study was performed along **Chapter 2 and 3**. The decision to continue the study with the indole alkaloid meridianin and not with aplicyanin was based on the existing amount of compound sample. Since we planned a further *in vitro* validation of the computational results, it was necessary to have a small amount of sample.

Interaction of meridianins A-G with GSK3 β , CK1 δ , DYRK1A, and CLK1, was consistently observed in all cases. Meridianins tend to bind at the ATP cavity, more concretely, at the phosphate binding groove, which is a polar region with the existence of arginine and aspartate, two aminoacids that facilitate the creation of HBs with small molecules or inhibitors. This supported the idea that the preferred binding zone of some small inhibitors is the phosphate binding groove (Tahtouh et al., 2012), deeper in the ATP cavity. This is in agreement with the results obtained on **Chapter 2**, where a detailed study of the binding mode of each meridianin, after docking calculations and MD simulations, **validated this binding mode, supporting the capacity of meridianins to act as kinase inhibitors**. However, a preferred position or orientation pattern for the meridianins on each of the studied bindings cannot be

established. This could be due to the different hydrophobic interactions established between each of the seven meridianins and the respective kinase proteins. This particularity could be further studied and analysed with the aim of finding some selectivity between kinases or even as a starting point when designing new inhibitors. Another revealing feature on the study of the inhibitory activity of meridianins was elucidated on **Chapter 2**, where **the presence of bromine on its chemical structure was found to be key**. A trend on the energies obtained after MM/GBSA calculations revealed a correlation between the higher values obtained and the presence of bromine on meridianins indole scaffold. Considering the results obtained so far and with the idea to start providing some light into this issue, in **Chapter 3** we designed different **indole scaffolds derivatives**, inspired by meridianins and kororamides. Their therapeutic relevance (Gul & Hamann, 2005; Klein-Junior et al., 2014) and the presence of halogen atoms (Pauletti et al., 2010; Gribble, 2015) at R₂ and R₃, was also revealed to be very important to increase the activity and potency on the design of new inhibitors against the studied kinases (**Chapter 2 and 3**). Three halogenated atoms, bromine, chlorine, and fluorine were tested, as they are used to increase therapeutic resilience and also they are the most commonly used halogen atoms on drug design (Sirimulla et al., 2013; Filgueira et al., 2014). The studies performed on **Chapter 3**, improved our understanding of the transcendental role of the indole scaffold, and also, how bromine atoms enhance the inhibitory capacity. In addition to that, our results pointed out that the introduction of an aromatic ring to the indole scaffold helps to stronger binding to the aforementioned kinases. The best designed derivative on **Chapter 3** incorporate it. In fact, from all the created indole derivatives, this is, by far, the most chemically similar compound to meridianins. This leads us back to the **hypothesis of the possible inhibitory activity of meridianins itself**. In that chapter, we also tried to discern which one of the two features, the indole scaffold or the halogen substituents was more important in terms of binding energy and binding mode, but it was clear that the combination of both indole scaffold plus brominate atoms is the option that gives better results. **Chapters 2 and 3** provide a deep overall picture of the binding mode and mechanisms of action of meridianins, but also other compounds, such as kororamide A, convolutamine J, and two indole derivative compounds 2a and 2e could be proposed as possible ATP-competitive kinase inhibitors to develop new anti-AD drugs. All these data highlight the great usefulness of computational techniques, since from the chemical structure of a marine molecule (meridianins) we managed to propose at least, **two other marine compounds and two newly designed structures as possible inhibitors of GSK3 β , CK1 δ , DYRK1A and CLK1**.

As expected, though, not all are positive results. The biggest challenge is the **lack of selectivity** of these molecules to the four protein kinases. In the performed studies we were not able to assure any

preference for one molecule to one kinase. Selectivity between kinases is an essential feature that needs to be amended on the further designing of appropriate anti-AD drugs (Davis et al., 2011). In that sense, we found differences between the orthosteric pockets of the kinases that can be exploited in the future.

Kinase inhibitors can target directly the active ATP-binding site or may be redirected to **allosteric sites** (B. Li et al., 2004). Due to the high structure conservation in the ATP binding site throughout the protein family, getting specificity between kinases by binding it, could be particularly difficult. With the aim to solve this inconvenient, on **Chapter 4**, we performed a search over GSK3 β looking for new druggable cavities. The results obtained pointed out the **substrate binding pocket**, a pocket places at the same N-loop, a few Å far from the ATP cavity, as a very good cavity that could lodge small inhibitor molecules. This is in line with other studies (Palomo et al., 2011; Bidon-Chanal et al., 2013), which reinforce the results and the employed methodology. Looking at the literature, there is a hypothesis stating that **binding on an allosteric cavity could also inhibit the activity of GSK3 β** (Palomo et al., 2011, 2017), having a direct impact on the reduction of tau hyperphosphorylation, and thus reverberate on the reduction of NFT formation and AD symptoms. As an allosteric pocket can give to our molecules selectivity, we decide to test the hypothesis. Since all the results we obtained so far pointed out that meridianins are the best compounds among all the MNPs studied, we selected them to perform a deep study trying to validate the aforementioned hypothesis. Also, lignarenone B was included in the study due to its structural similarity to other non-ATP natural inhibitor of GSK3 β (Bidon-Chanal et al., 2013). The results obtained suggested a better binding over the ATP catalytic cavity, instead of into the substrate binding pocket. Different reasons may explain these results, ranging from the large size of the substrate pocket compared to the ATP cavity, to the fact that the substrate pocket is placed in a loop, what is translated to high fluctuations due to the solvent exposure (Nilmeier et al., 2011). From an experimental validation carried out, the inhibitory activity of both molecules was evaluated comparing serine 9 phosphorylation and total GSK3 β levels. In any of the assays was the viability of GSK3 β compromised, breaking the stigma of their toxicological potential (**Chapter 1**), and even neurite outgrowth increased. However, we could not verify *in silico* or *in vitro* whether the inhibition occurs at the binding of the ATP catalytic cavity, at the substrate pocket, or even at both. The existence of bivalent inhibitors is not new (Roskoski, 2016), and it is a very interesting feature that could be exploited in the future.

Pharmacokinetic properties to found hits from marine scaffolds

From the set of molecules studied, those that come from the database and the literature as well as the new compounds designed along this thesis, require further careful, individual analysis and pharmacokinetics properties improvement. They can be considered hits but are far from being leads, among other reasons, because of their pharmacokinetics properties.

As a general trend **absorption properties**, should be improved, especially according to P-Glycoprotein (Pgp) substrate binding. However, the results are not extremely bad, we can consider them moderate good absorption results, but to make them optimal, an improvement is needed. Regarding **distribution properties**, the studied compounds showed high PPB probabilities as well as low VDs, which means these compounds will have a lot of difficulties in diffusing or traversing cell membranes. Moreover, most of the molecules are not supposed to cross the blood brain barrier (BBB) and penetrate the CNS, with the exception of the designed compounds 2a and 2e (**Chapter 4**). As anti-AD drugs, this is an important requisite. The inability to penetrate the central nervous system (CNS), could be alleviated at the same time that absorption properties, as Pgp binding (acting as a substrate not an inhibitor that has the opposite effect) is related to the ability to penetrate the CNS (Sadeque et al., 2000; Breedveld et al., 2006). Moreover, a part of the obvious chemical structure modification to improve the penetration properties, there exist the possibility to facilitate the BBB penetration, for instance, by nanocarriers, making use of nanodelivery techniques (Banks, 2016; Kingwell, 2016; Saraiva et al., 2016). **Metabolic properties** analysis revealed that the analysed compounds are likely to be metabolised by CYP450, which could results in adverse reactions or therapeutic failures, so a careful analysis is needed to obtain safer hit or leads. The **excretion properties** showed good results. Only few compounds seems to interact with OCT2, and that the compounds that interact are in the border-line of the model to be classified as interacting or non-interacting compounds, so non-clearance problems and adverse interactions with co-administrated OCT2 inhibitors are expected. Regarding **toxicology properties**, none of the compounds studied indicated human ether-a-go-go gene (hERG) inhibition. However, despite of that, some hepatotoxicity, mutagenic, and carcinogenicity propensities were detected. Those toxic properties need to be removed further on.

To sum up, the obtained results suggest the necessity to performing H2L optimization cycles, in order to improve the absorption, distribution, metabolism and toxicity of the studied compounds, as well as their selectivity according to the binding results commented before, with the aim of obtaining lead compounds able to become effective anti-AD drugs.

To perform H2L optimization cycles, new experiments will be required. These experiments guided by **medicinal chemistry** experts together with computational tools, will be a good option/workflow to employ in order to improve all of these properties. Medicinal chemistry plays an important role on the H2L optimization process (Hoffer et al., 2018). Due to structural modifications guided by medicinal chemistry, optimizing compound properties could be possible, providing the desired efficacy, at acceptable dose, improving its ADME properties, and minimizing any toxicological feature (Hann & Keseř, 2012). Computational chemistry is potent by itself, but wet lab experiments are always needed. The combination of both dry and wet lab techniques is the most powerful and efficient approach.

Concluding remarks and future perspectives

The main conclusion of this thesis is that marine molecules can be used as therapeutic agents against several diseases, especially AD, and CADD techniques can help develop this research line. We have focused especially in AD, but in fact both our results (Chapter 1) as well as the literature, clearly indicate that this can be exerted to other diseases. Along this thesis we exemplified the potential of CADD applied to marine drug discovery. Both Antarctic and Mediterranean benthic organisms should be considered to be rich sources of MNPs capable to modulate pathogenesis-related pathways. This is not something new, it is a fact widely accepted, but there are very few studies attempting to systematically elucidate the biological profile of a relatively big set of marine compounds. Probably extending the analysis performed in Chapter 1 to a large dataset, as the whole MarineLit database, will help to get knowledge about the possible bioactivities of marine drugs. Probably it would increase even more the usage of MNPs as source of inspiration in drug design.

In this thesis, we also clearly pointed out the benefit and usefulness of computational methods to elucidate the pharmacological potential of MNPs and also to find hits. Although powerful methods, alone are not completely conclusive, when mixed with wet-lab experiments, as in Chapter 4, they constitute a strong pipeline, as they have predictive capabilities.

Final Conclusions

Final conclusions

From the studies carried out during this thesis we can conclude that:

1. Meridianins, Lignarenones, Kororamides and derivatives 2a, and 2e, can be considered hit compounds to fight AD through the inhibition of GSK3 β , CK1 δ , DYRK1A, and CLK1 protein kinases.
2. Indole scaffolds with halogen substituents constitute a good starting point to design hits and lead compounds against the aforementioned kinases.
3. CADD techniques are interesting predictive tools that can help the biological profile elucidation of MNPs as well as to identify them as hits compounds for a particular target and diseases. In combination with wet-lab experiments they are even more useful.
4. The utilization of VP techniques in this thesis allowed for the elucidation of a list of possible targets, relating them to a particular disease, such as cardiovascular and neurodegenerative pathologies, and also, to identify the binding cavity where the marine molecules could exert their activity.
5. The evaluation of drug-likeness properties of our compounds by comparing their binding energies constitutes an interesting step that can be established as another computational validation of the proposed inhibitory activity for other marine molecules.
6. Detailed binding mode study elucidated new insights into the small molecules binding on GSK3 β , CK1 δ , DYRK1A and CLK1, allowing the determination of the interactions established between each marine molecule-target complex, and thus, proved its future applicability on the discovery and design of new leads or hits.
7. Molecules such as meridianins A-G are proposed as possible inhibitors over the specific tau kinase GSK3 β , CK1 δ , and dual-specificity kinases, as DYRK1A and CLK1. This constitutes a promising starting point for the development of novel anti-AD drugs.
8. The natural marine compounds analysed here, or its derivatives, are a very interesting source of inspiration for the discovery of novel leads with potent therapeutic activity against protein kinases involved in the AD pathway.
9. The presence of two bromine atoms at the R₂ and R₃ positions on the indole scaffold of meridianins was revealed to be synonymous of potency. In this sense, and following a rational design, a couple of new designed brominated lead compounds (2a and 2e), were established to

exert the best inhibitory activity, which supports them as a good scaffold to start an optimization process.

10. Other brominated molecules, such as kororamide A-B, with two brominated atoms on its indole scaffold and, perhaps also convolutamine I-J, heterocyclic compounds with three bromine atoms, showed some inhibitory activity against protein kinases too, supporting their potential as anti-AD therapeutic agents.
11. The potential of CADD on the elucidation of meridianins A-G and lignarenones as possible inhibitors of the specific tau kinase GSK3 β was validated by an experimental study, and supported the hypothesis that meridianins and in less maner, lignarenone, could be used on the treatment of Alzheimer's pathologies reducing the pathological hyperphosphorylation of tau and thus reducing the NFT formation.

Catalan Summary

Objectius d'aquesta tesi

L'objectiu principal d'aquesta tesi és **dilucidar la possible activitat terapèutica i la capacitat per modular les funcions de proteïnes que estan relacionades amb una determinada patologia de les molècules marines mitjançant l'ús de diferents eines i tècniques de disseny de fàrmacs assistit per ordinador (DFAO)**. D'acord amb aquest l'objectiu, la present tesi es divideix en tres seccions on s'intenta il·lustrar l'assoliment d'aquests. En la **Secció I**, poso en rellevància com un enfocament computacional podria **millorar el pipeline de descobriment de fàrmacs** (Capítol 1). La **Secció II**, es centra en la **dilucidació de les diferents característiques farmacofòriques dels compostos marins i en un precís estudi d'unió *in silico***, que acaba amb la dilucidació de la capacitat de diferents compostos marins per actuar com a inhibidors de les proteïnes tau quinases (GSK3 β i CK1 δ) i de les proteïnes quinases de doble especificitat (DYRK1A i CLK1), totes elles relacionades amb la malaltia d'Alzheimer (MA), el que representa un prometedor punt de partida per al desenvolupament de nous fàrmacs contra la MA (Capítol 2 i Capítol 3). La **Secció III**, presenta un **estudi computacional i una validació experimental** de l'activitat inhibidora de les meridianines i les lignarenones com a possibles inhibidors de la GSK3 β mitjançant la seva unió a les cavitats del trifosfat d'adenosina (ATP) i/o del substrat, la qual cosa permet proposar-los com a fàrmacs candidats per al tractament de la MA (Capítol 4).

Els objectius específics de cada capítol es resumeixen a continuació:

- Capítol 1. **Estudis *in silico* per trobar noves indicacions terapèutiques per a molècules marines.** L'objectiu principal d'aquest estudi és (I) establir el possible potencial terapèutic de diverses molècules marines mitjançant l'ús de diferents tècniques computacionals, (II) predir i validar el mode d'unió, la forma en la qual interaccionen, molècula marina-diana, (III) dilucidar una llista de possibles dianes, (IV) avaluar els seus efectes adversos en la salut mitjançant un estudi preliminar de predicció de toxicologia; i (V) estimar les propietats com a fàrmacs de cadascuna de les molècules estudiades.
- Capítol 2. **Disseny de fàrmacs assistit per ordinador aplicat a la cerca de possibles fàrmacs marins: Meridianines com a agents terapèutics de la malaltia d'Alzheimer.** El nostre objectiu és (I) ressaltar el poder de les tècniques de DFAO en molècules marines i productes naturals en general, alhora de trobar-hi possibles usos terapèutics; (II) avaluar i informar de l'activitat

inhibitòria trobada en el tunicat marí *Aplidum*: Meridianines A-G, actuant com a inhibidors competitiu de l'ATP en GSK3 β , CK1 δ , DYRK1A, i CLK1; (III) avaluar els seus possibles efectes adversos en la salut mitjançant un estudi preliminar de predicció de propietats farmacocinètiques (PK); i (IV) analitzar les seves propietats farmacològiques així com l'acció d'àtoms halògens en la seva estructura.

- Capítol 3. **Kororamides, Convolutamines, i derivats indòlics com a possibles inhibidors de les proteïnes tau quinases i de les quinases de doble especificitat per a la malaltia d'Alzheimer: un estudi computacional.** L'objectiu és (I) determinar la possible acció terapèutica de kororamides i convolutamines contra la MA mitjançant la inhibició de GSK3 β , CK1 δ , DYRK1A, i CLK1; (II) determinar la importància de l'esquelet indòlic en la inhibició de les quatre quinases estudiades i la importància/efecte dels substituents halògens; (III) dissenyar nous possibles inhibidors de les quatre quinases a partir dels esquelets indòlics de meridianines i kororamides; i (IV) avaluar els efectes adversos de kororamides, convolutamides i els seus derivats en la salut mitjançant un estudi d'administració, distribució, metabolisme, excreció i toxicitat (ADMET).
- Capítol 4. **Meridianines i lignarenones com a potencials inhibidors de GSK3 β i inductors de la plasticitat sinàptica neuronal.** El propòsit aquí és (I) dilucidar la possible activitat inhibidòria de meridianines i lignarenone sobre les cavitats de l'ATP i/o del substrat en la GSK3 β , una diana clau involucrada en la ruta de la MA; (II) explorar llocs d'unió farmacològics en GSK3 β a la recerca de noves cavitats al·lostèriques; (III) predir les propietats farmacològiques; i (IV) validar experimentalment l'activitat inhibidòria de les meridianines i les lignarenones comparant els nivells de fosforilació de la Ser9 i els nivells totals de GSK3 β com a indicador de la inhibició.

Resultats

A continuació es presenta un resum clar i concís dels principals resultats obtinguts en cada un dels articles científics publicats en aquesta tesi.

Capítol 1. Estudis *in silico* per trobar noves indicacions terapèutiques per a molècules marines.

Identificar petites molècules que s'adaptin bé a una cavitat activa és un dels primers passos a seguir en el descobriment de fàrmacs. En aquest estudi pretenem dilucidar una llista de possibles dianes, i el potencial terapèutic, d'un conjunt de molècules marines utilitzant diferents eines computacionals. L'acoblament molecular és un dels instruments del disseny de fàrmacs assistit per ordinador (DFAO) més comuns i que permet l'estudi de les interaccions proteïna-l·ligand, predint a la vegada, tant l'orientació com la postura de la molècula acoblada. La captura dels moviments de les proteïnes és clau per entendre aquestes interaccions proteïna-l·ligand, i la simulació de dinàmica molecular (DM) és la millor eina computacional per fer-ho. Mitjançant la combinació d'aquestes eines computacionals i d'altres, en aquest estudi hem pogut dilucidar el vincle entre un grup de molècules marines i algunes patologies neurodegeneratives i cardiovasculars. A més, hem avaluat els possibles efectes adversos en la salut mitjançant prediccions de toxicologia, i les propietats farmacològiques d'aquest conjunt de molècules marines, proporcionant algunes idees sobre la predicció dels vincles d'unió dels complexos molècules marines-proteïna. En aquest estudi exemplifiquem com es pot aplicar l'ús d'eines computacionals en el camp del descobriment de fàrmacs marins, establint un procediment que es pot seguir en futurs estudis.

Capítol 2. Disseny de fàrmacs assistit per ordinador aplicat a la cerca de possibles fàrmacs marins: Meridianines com a agents terapèutics de la malaltia d'Alzheimer.

Les tècniques de descobriment o disseny de fàrmacs assistits per ordinador (DFAO) permeten la identificació de productes naturals que són capaços de modular les funcions de proteïnes que estan relacionades amb una determinada patologia, constituint una de les línies a seguir més prometedores en el descobriment de fàrmacs. En aquest treball, hem avaluat computacionalment l'activitat inhibidòria trobada en les meridianines A – G, un grup d'alcaloides indòlics marins aïllats del tunicat marí *Aplidium*,

contra diverses proteïnes quinases implicades en la malaltia d'Alzheimer (MA), una patologia neurodegenerativa caracteritzada per la presència de cabdells neurofibril·lars. La pèrdua de l'estabilitat entre les activitats de la tau quinasa i el fosfat és la causa de la hiperfosforilació de la tau i, per tant, la seva agregació i formació de cabdells neurofibril·lars. La inhibició de quinases específiques implicades en la seva via de fosforilació podria ser una de les estratègies clau per revertir la hiperfosforilació de la tau i representaria una aproximació per desenvolupar fàrmacs per pal·liar els símptomes de la MA. Les meridianines s'acoblen al lloc d'unió del trifosfat d'adenosina (ATP) de determinades proteïnes quinases, actuant com a inhibidors competitiu del ATP. Aquests compostos mostren esquelets molt prometedors per dissenyar nous fàrmacs contra la MA, els quals podrien actuar sobre les proteïnes tau, com ara la glycogen synthase kinase-3 beta (GSK3 β) i la casein kinase 1 delta (CK1 δ , CK1D o KC1D), i les quinases de doble especificitat, com la dual specificity tyrosine phosphorylation regulated kinase 1 (DYRK1A) i les quinases cdc2-like (CLK1). Aquest treball té com a objectiu destacar el paper de les tècniques de DFAO en el descobriment de fàrmacs marins i proporcionar informació precisa sobre el mode i la força d'unió de les meridianines contra diverses proteïnes quinases que podrien ajudar en el futur desenvolupament de fàrmacs contra la MA.

Capítol 3. Kororamides, Convolutamines, i derivats indòlics com a possibles inhibidors de les proteïnes tau quinases i de les quinases de doble especificitat per a la malaltia d'Alzheimer: un estudi computacional.

La malaltia d'Alzheimer (MA) s'està convertint en una de les malalties més inquietants i en un problema socioeconòmic en l'actualitat, ja que és una patologia neurodegenerativa que no té tractament, i s'espera que la seva afectació augmenti encara més a causa de l'envelliment de la població. Els tractaments actuals per a la MA només produeixen una modesta millora dels símptomes, tot i que hi ha una constant i permanent investigació de noves estratègies terapèutiques orientades a millorar aquests símptomes, i fins i tot, per curar completament la malaltia. Una característica principal de la MA és la presència de cabdells neurofibril·lars, induïts per una aberrant fosforilació de la proteïna tau associada als microtúbuls, que es troben presents en el cervell dels individus afectats. La glycogen synthase kinase-3 beta (GSK3 β) i la casein kinase 1 delta (CK1 δ , CK1D o KC1D), així com les quinases de doble especificitat, com la dual specificity tyrosine phosphorylation regulated kinase 1 (DYRK1A) i les quinases cdc2-like (CLK1), s'han identificat com les principals proteïnes implicades en aquest procés d'hiperfosforilació. Per això, la inhibició d'aquestes quinases s'ha proposat com una estratègia terapèutica plausible per combatre la MA. En aquest estudi, hem estudiat computacionalment l'activitat

inhibitòria de diferents compostos naturals d'origen marí, així com molècules dissenyades a partir dels seus esquelets, sobre les esmentades proteïnes quinases, trobant alguns nous possibles inhibidors amb potencial aplicació terapèutica.

Capítol 4. Meridianines i lignarenones com a potencials inhibidors de GSK3 β i inductors de la plasticitat sinàptica neuronal.

La glycogen synthase kinase-3 (GSK3) és una proteïna essencial, amb un paper rellevant en moltes patologies com ara la diabetis, el càncer i les malalties neurodegeneratives. En particular, la isoforma GSK3 β està relacionada amb patologies com la malaltia d'Alzheimer (MA). Aquest enzim constitueix un objectiu molt interessant per al descobriment i/o disseny de nous agents terapèutics contra la MA a causa de la seva relació amb la hiperfosforilació de la proteïna tau associada als microtúbuls, i per tant, la seva contribució en la formació de cabdells neurofibril·lars. Els estudis *in silico* van identificar dues famílies de molècules marines, els alcaloides indòlics meridianines, del tunicat *Aplidium*, i les lignarenones, metabòlits secundaris del mol·lusc cefalaspidi amb closca *Scaphander lignarius*, com a possibles inhibidors GSK3 β . L'anàlisi de la superfície de GSK3 β revelà que les dues molècules marines poden actuar sobre les regions d'unió del trifosfat d'adenosina (ATP) i/o del substrat. El predit potencial inhibitori d'aquestes dues molècules marines va ser validat experimentalment *in vitro* mitjançant la comparació dels nivells de fosforilació de Ser9 i els nivells totals de GSK3 β , a la vegada que es va poder observar que les dues molècules potencien la plasticitat sinàptica estructural. Aquests fets ens permeten suggerir que les meridianines i la lignarenone B podrien ser utilitzats com a possibles candidats terapèutics per al tractament de GSK3 β implicada en patologies com ara la MA.

Discussió general

Els invertebrats bentònics marins posseeixen productes naturals molt interessants, amb un potencial enorme per al descobriment de (Leal et al., 2012; Avila, 2016; Blunt et al., 2018a). En particular, els organismes d'àrees poc explorades, com l'Antàrtida, amaguen un enorme arsenal de quimiodiversitat que recentment ha començat a ser explotada (Barre, 2010; Núñez-Pons et al., 2015; Blunt et al., 2018a). Un dels problemes principals d'aquests estudis és predir quina bioactivitat o potencial ús terapèutic pot tenir cada nou compost, ja que la limitada quantitat de material sol evitar un ampli cribratge de bioactivitat. Per aquestes raons, l'ús de mètodes computacionals és primordial per predir les potencials activitats donada una estructura molecular determinada. L'ús d'aquestes tècniques computacionals i, per tant, els resultats obtinguts en l'estudi sobre la possible activitat terapèutica i la capacitat de modulació de la funció proteica en vies relacionades amb patologies dels PNMs recol·lectats en les comunitats bentònica de l'Antàrtida i el Mediterrani, permeten cobrir diferents passos del protocol de descobriment de fàrmacs. De fet, al llarg d'aquesta tesi, he utilitzat moltes eines i tècniques de DFAO, que també s'han estudiat i implementat de manera comparativa. Aquesta tesi posa de manifest la **utilitat del DFAO en el procés de desenvolupament de fàrmacs**, donant valuosos coneixements sobre el propi procés (capítol 1), sobre les possibles propietats farmacològiques d'un grup de PNMs (capítol 2 i 3), i sobre modulació al·lostèrica com a possible solució en la recerca d'un tractament per la MA (capítol 4). A més, els nostres estudis sobre productes naturals d'ascidiacs aporten **noves evidències sobre la possible condició terapèutica** d'aquests metabòlits secundaris contra un grup de quinases relacionats amb la hiperfosforilació de la tau en la patologia d'Alzheimer (capítol 2, 3 i 4). Dues de les contribucions més importants d'aquesta tesi són, en primer lloc, **l'aplicabilitat d'un protocol establert per passos** en el procés de descobriment de fàrmacs *in silico* dels PNMs, i en segon lloc, l'elucidació computacional i posterior **validació *in vitro* de l'activitat inhibidora** de meridianines A-G i lignarenone B contra GSK3 β . Aquí, discutim les troballes més importants, que s'organitzen en quatre seccions, proporcionant una visió global i general dels resultats obtinguts al llarg d'aquests anys en el camp del descobriment de fàrmacs. Es comenten les perspectives més rellevants sobre els resultats obtinguts dins d'una perspectiva general. Les recomanacions i limitacions es tracten també en cada secció, mentre que perspectives futures estan exposades a l'última secció d'aquest capítol.

Potencial de DFAO en el descobriment de fàrmacs

El desenvolupament de nous fàrmacs és un procés complex, i segons això, els enfocaments computacionals són eines versàtils que faciliten i acceleren el disseny i desenvolupament de fàrmacs (Prachayasittikul et al., 2015). En la meua opinió, la millor definició de DFAO, des d'un punt de vista estructural, va ser feta per Baig et al., qui definia DFAO com els **mètodes utilitzats amb l'objectiu de simular interaccions entre receptors i fàrmacs per tal de determinar afinitats vinculants** (Baig et al., 2017). La utilització d'aquests mètodes amb PNM no és nova (Medina-Franco, 2013; Pereira & Aires-de-Sousa, 2018), però la potencialitat i utilitat de les tècniques DFAO aplicades a PNM s'ha vist ressaltada al llarg d'aquesta tesi. Un dels principals avantatges aquí és evitar el **malbaratament de recursos naturals**, que sovint no es valora prou en la literatura. La recollida de mostres, que són exclusivament presents en els hàbitats marins, requereix enormes esforços des d'una perspectiva econòmica, humana, tècnica i biològica, i, a més, la majoria d'aquests organismes solen ser no cultivables (Molinski et al., 2009a). Per tant, la contribució de les tècniques de DFAO és molt rellevant, ja que no es requereixen mostres per a cap anàlisi computacional. A més, el DFAO proporciona certs coneixements sobre la química de les molècules al qual no es pot accedir a través d'experiments *in vitro*, reduint el cost i el temps, i millorant la qualitat de la recerca (Macalino et al., 2015).

Hi ha diversos exemples d'aplicació de DFAO exitoses en diferents àrees i patologies (Marshall, 1987; Propst & Perun, 1989; Song et al., 2009; Talele et al., 2010; Das, 2017). En aquesta tesi, ens vam centrar en el descobriment i disseny de nous compostos per al tractament de la MA.

És una evidència que les aplicacions DFAO, com qualsevol predicció computacional, tenen **limitacions**. La majoria de les tècniques de DFAO, com les emprades en aquesta tesi, ja siguin, acoblament molecular, QSAR, VS, VP, modelització de farmacòfor, o DM, es basen en algorismes i scripts predefinitos, que restringeixen la seva confiança i rendiment. Les actualitzacions d'algorismes i un increment de les validacions de dades experimentals són aspectes clau per superar aquestes limitacions i per millorar la precisió en les prediccions del DFAO.

Els fracassos més comuns en el descobriment de fàrmacs depenen de la **predicció incorrecta dels complexos de lligand-proteïna**. Aquesta errònia predicció pot ser causada per diferents aspectes, però el més probable és l'errònia determinació de l'energia o la mala predicció de lligand, proteïnes, o ambdues estructures (Verkhivker et al., 2000; Ramírez & Caballero, 2016, 2018). La validació més freqüent per superar aquests inconvenients, és la **reproductibilitat** correcta del complex lligand-proteïna (reacoblament) i l'assignació de valors correctes als millors acoblaments, per establir algunes correlacions entre les postures, els lligands coneguts, i si és el cas, i les afinitats mesurades del "lead"

(Korb et al., 2012; Talevi, 2018). Al llarg d'aquesta tesi, tots els experiments d'acoblament es van realitzar com a mínim dues vegades, seguint el principi de reacoblament. Amb l'objectiu de reduir el fracàs i la validació de les prediccions d'acoblament, l'enfocament més habitual, extensiu i rutinari és l'ús de simulacions de DM, que s'utilitzen com a tècniques de post-processament, i permeten comprendre els moviments de proteïnes i l'exploració de l'espai conformacional (Karplus & Kuriyan, 2005; Sakano et al., 2016). Malgrat que les simulacions de DM són mètodes precisos, també presenten inconvenients. Les seves principals limitacions són l'**escala de temps** i el **refinament** requerit dels camps de força (Durrant et al., 2011). Les simulacions de DM normalment es duen a terme en l'ordre de nanosegons (ns) a microsegons (μ s). No obstant això, per detectar canvis conformacionals biològics sobre el plegament de proteïnes, es necessiten dinàmiques en l'ordre de mil·lisegons (ms) (Leelananda et al., 2016). Simular en aquesta escala de temps és consumir molt temps i requereix d'un gran cost computacional. S'han introduït, recentment, diversos mètodes en el camp amb l'objectiu d'explorar millor l'espai conformacional i amb temps restringit, com ara DM accelerada (aDM), metadynamics, rèplica d'intercanvi DM (REMD), i integració termodinàmica, entre d'altres (Bernardi et al., 2015). No obstant això, depenent de l'estudi, amb passos de temps molt més petits es suficient. Com a eina de post-processament d'acoblament, pocs ns o fins i tot ps són necessaris. En la literatura hi ha diferents corrents que postulen que una certa combinació de passos de temps i nombre de rèpliques és millor que d'altres (Hou et al., 2011; Genheden et al., 2015; Sun et al., 2018). Després de diverses proves, en aquesta tesi vam decidir emprar simulacions de DM de 1ns de durada. Aquesta durada és suficient per veure com una postura d'unió dolenta deixa la cavitat, de fet, uns quants centenars de ps són suficients. A partir d'aquestes simulacions curtes, s'estimen les energies lliures d'unió i/o les afinitats. Aquests valors es poden utilitzar posteriorment per classificar els compostos analitzats, ajudant en l'elucidació del millor compost d'acord a un objectiu específic. Es poden utilitzar diferents enfocaments per inferir l'energia d'unió, tot i que avui en dia la tècnica més precisa sembla ser la pertorbació de l'energia lliure (FEP), i la integració termodinàmica (TI). Al llarg d'aquesta tesi es fan servir càlculs MM/GBSA pel seu bon equilibri entre exactitud, potència computacional i temps, per calcular les energies lliures d'unió. Aquesta metodologia ha estat àmpliament utilitzada durant dècades i la seva predicció exitosa d'energies d'unió i la bona classificació entre compostos està més que demostrada (Mulakala et al., 2013; Zhang et al., 2017).

Una altra limitació important en el camp de descobriment de fàrmacs *in silico* és l'**absència de l'estructura d'unió**, que encara no està resolta experimentalment. Aquest fet no permet l'exploració de tot el proteoma (Barril, 2017), encara que hi ha tècniques com els models d'homologia que es poden

utilitzar per reduir aquest tipus de problemes. No obstant això, malgrat els avenços en les tècniques de determinació de l'estructura, tant experimentalment com computacionalment, avui en dia hi ha una gran quantitat de proteïnes que no es poden explorar computacionalment (Barril, 2017). A causa d'aquest fet, entre d'altres, hem de destacar que encara no sabem i estem lluny d'**entendre l'etiologia** de la majoria de malalties (Gonzalez & Kann, 2012).

A més, les eines de predicció d'ADMET també han de ser revisades, malgrat les contínues millores i l'evolució que s'ha experimentat durant les darreres dècades. Calen dades experimentals sòlides perquè unes bones dades són la base de qualsevol bon model. L'enfocament més eficaç, ben establert i rendible per disminuir el fracàs en etapes posteriors en el procés de descobriment de fàrmacs abans que la molècula es sintetitzi, i augmenti els problemes de seguretat, és aplicar la **predicció de propietats de d'ADMET computacional en les primers etapes** (Clark, 2005). La predicció d'ADMET computacional ajuda molt a reduir la predicció de males d'ADMET experimentals, responsables del fracàs de fàrmacs en posteriors etapes del procés de descoberta de fàrmacs, però amb millors models la seva contribució serà molt més alta. La majoria dels mètodes d'ADMET es basen en el principi de similitud (que significa compostos similars, similar comportament), però avui en dia se sap que els **efectes tòxics** poden ser causats per interaccions amb altres proteïnes, on el compost "lead" previst s'uneix a una cavitat catalítica similar, però no a la diana prevista, aquest fet es coneix com off-target (Rudmann, 2013).

Per seleccionar els millors models de paràmetres d'ADMET predictius, és crucial seleccionar l'enfocament matemàtic adequat, els millors descriptors moleculars per a una particular propietat d'ADMET, i la mida apropiada del conjunt de dades experimentals relacionades amb aquesta propietat, per a la correcta validació del model (van de Waterbeemd et al., 2003). Al llarg d'aquesta tesi, capítols 1-4, hi ha un apartat que descriu les propietats PK de cada molècula estudiada en cada etapa. En el capítol 1, la predicció de toxicologia va ser avaluada utilitzant models QSAR "clàssics", però no obstant això, en els capítols 2-4, a causa del gran auge dels últims anys i la seva rellevància, models **machine learning** (ML) basats en mètodes QSAR, es varen emprar mitjançant l'ús d'eines de programari ADMETer i pkCSM (Pires et al., 2015; Vidal & Nonell-Canals, 2017). ML es basa en la construcció de models computacionals que poden millorar la presa de decisions a partir de dades d'alta qualitat (Vamathevan et al., 2019). D'altra banda, es pot aplicar ML per desenvolupar models per predir les propietats químiques com l'absorció, distribució, metabolisme, excreció i toxicitat (ADMET) (Heikamp & Bajorath, 2014; Lavecchia, 2015). ADMETer és una eina de programari que conté **suport de màquines vectorials** (SVMs) per a la predicció de la permeabilitat Caco2, BBB, i PPB, així com, els algorismes de **suport de**

regressió vectorial (SVR) per a l'estimació de la LogP i LogS. En els últims anys, a causa de la seva predicció precisa i consistent, SVMs i en menor mesura, SVR, s'han tornat cada vegada més populars en el camp del descobriment de fàrmacs. Els avantatges d'aquests models predictius per a determinar les propietats d'ADMET es deuen al fet que són molt confiables, però com qualsevol model computacional, també hi ha algunes limitacions. Probablement, la principal limitació dels models ML és el domini de l'aplicabilitat, ja que depenen dels conjunts de dades disponibles per generar i entrenar els models apropiats, de manera que la predicció només es produeix en els marcs coneguts de les dades d'entrenament (Vamathevan et al., 2019).

Inhibidors de les proteïnes kinases i propietats farmacofòriques del PNM

El kinoma humà, totes les proteïnes quinases catalogades codificats pel genoma humà, inclou un total de 518 proteïnes dividides en set subfamílies (Manning et al., 2002). Tenint en compte l'estudi realitzat en el capítol 1, hem pogut identificar la capacitat d'un grup de molècules marines per vincular-se a proteïnes relacionades amb diverses patologies. Concretament, veiem una tendència que assenyalava que els PNM tendeixen a vincular-se amb proteïnes quinases, la gran família d'enzims de senyalització, tal com es proposat també per altres autors (Marston, 2011; Skropeta et al., 2011). A partir d'aquestes dades inicials, els nostres resultats suporten **la hipòtesi del potencial dels PNM per actuar com a inhibidors de les proteïnes quinases**. En aquest sentit, en el nostre estudi (capítol 1), es fundamenta la vinculació aquestes proteïnes quinases amb aplicyanina i meridianina A, dues molècules que comparteixen un esquelet molt similar. Malgrat el vincle entre meridianines i proteïnes quinases no és nou (Giraud et al., 2011; Bharate et al., 2012), mai s'ha dut a terme una aproximació basada en una perspectiva estructural. Per tant, es va realitzar una **primera validació dels usos i aplicabilitat del DFAO aplicat al PNM**. Per tal d'avaluar la capacitat real de meridianines A-G, es va dur a terme un estudi computacional al llarg del capítols 2 i 3. La decisió de continuar l'estudi amb l'alcaloide indòlic meridianina i no amb aplicyanina es basava en la quantitat existent de mostra del compost. Atès que vam planificar una posterior validació *in vitro* dels resultats computacionals, calia tenir una petita quantitat de mostra.

La interacció de les meridianines A-G amb GSK3 β , CK1 δ , DYRK1A i CLK1, es va observar sistemàticament en tots els casos. Les meridianines tendeixen a unir-se a la cavitat de l'ATP, més concretament, al solc fosfatat, que és una regió polar amb l'existència d'arginina i aspartat, dos

aminoàcids que faciliten la creació de ponts d'hidrogen amb molècules petites o inhibidors. Això donava suport a la idea que la zona d'unió preferida d'alguns inhibidors petits és el solc fosfatat (Tahtouh et al., 2012), situat a la part més profunda en la cavitat de l'ATP. Això concorda amb els resultats obtinguts en el capítol 2, on un estudi detallat de la manera d'unió de cada meridianina, després de càlculs d'acoblament i les simulacions MD, va **validar aquest mode d'unió, donant suport a la capacitat de les meridianines per actuar com inhibidors de les proteïnes quinases**. No obstant això, no es pot establir una posició preferida o un patró d'orientació per les meridianines en cadascuna de les unions estudiades. Això podria ser degut a les diferents interaccions hidrofòbiques establertes entre cada una de les set meridianines i les proteïnes quinases respectives. Aquesta particularitat pot ser estudiada i analitzada amb l'objectiu de trobar certa selectivitat entre quinases o fins i tot com a punt de partida en el disseny de nous inhibidors. Un altre tret revelador en l'estudi de l'activitat inhibidora de les meridianines va ser elucidat en el capítol 2, on **la presència de brom en la seva estructura química es va trobar que era clau**. Una tendència a les energies obtingudes després de càlculs MM/GBSA va revelar una correlació entre els valors més alts obtinguts i la presència de brom en l'esquelet indòlic de les meridianines. Tenint en compte els resultats obtinguts fins ara i amb la idea de començar a donar llum a aquesta qüestió, en el capítol 3 hem dissenyat diferents **derivats a partir dels esquelets indòlics**, inspirats en meridianines i kororamides. La seva rellevància terapèutica (Gul & Hamann, 2005; Klein-Junior et al., 2014) i la presència d'àtoms halògens (Pauletti et al., 2010; Gribble, 2015) a les posicions R₂ i R₃, també es va revelar molt important per incrementar l'activitat i potència en el disseny de nous inhibidors enfront de les quinases estudiades (capítol 2 i 3). Es van provar tres àtoms halogenats, brom, clor i fluor, ja que s'utilitzen per augmentar la resiliència terapèutica i també són els àtoms d'halogen més comúment utilitzats en el disseny de fàrmacs (Sirimulla et al., 2013; Filgueira et al., 2014). Els estudis realitzats en el capítol 3, van millorar la nostra comprensió del paper transcendental de l'esquelet indòlic, i també, com els àtoms de brom potencien la capacitat inhibidora. A més d'això, els nostres resultats van assenyalar que la introducció d'un anell aromàtic a l'esquelet indòlic ajuda a la unió més forta a les esmentades quinases. El millors compostos dissenyats en el capítol 3, derivats de l'esquelet indòlic, l'incorporen (2a i 2e). De tots els derivats indòlics creats, aquests són de lluny, els compostos químicament més similars a les meridianines. Això ens porta de nou a la **hipòtesi de la possible activitat inhibidora de les meridianines**. En aquest capítol, també es va tractar de discernir si una de les dues característiques, l'esquelet indòlic o el substitut d'halogen podia ser més important en termes d'energia d'unió i de manera d'unió, però es va evidenciar que la combinació de l'esquelet indòlic més els àtoms de brom és l'opció que dona millors resultats. Els capítols 2 i 3 proporcionen una

imatge global profunda de la forma d'unió i els mecanismes d'acció de les meridianines, però també altres compostos, com ara kororamide A, convolutamine J, i dos compostos derivats indòlics 2a i 2e que podrien ser proposats com a possibles inhibidors de quinases competitiu de l'ATP per desenvolupar nous fàrmacs contra la MA. Totes aquestes dades destaquen la gran utilitat de les tècniques computacionals, ja que a partir de l'estructura química d'una molècula marina (meridianines) hem aconseguit proposar com a mínim altres **dos compostos marins i dues estructures de nova construcció com a possibles inhibidors de GSK3 β , CK1 δ , DYRK1A i CLK1**.

Com s'esperava, però, no tots són resultats positius. El desafiament més gran és la **manca de selectivitat** d'aquestes molècules per les quatre quinases. En els estudis realitzats no vam poder assegurar cap preferència per una molècula a una quinasa. La selectivitat entre quinases és una característica essencial que ha de ser modificada en el disseny de fàrmacs apropiats per a la MA (Davis et al., 2011). En aquest sentit, trobem diferències entre les cavitats ortostèriques de les quinases que es poden explotar en el futur.

Els inhibidors de quinases poden dirigir-se directament al lloc d'unió de l'ATP o pot ser redirigit a **llocs al·lostèrics** (B. Li et al., 2004). A causa de l'alta conservació de estructura en el lloc d'unió de l'ATP al llarg de la família de proteïnes, obtenir l'especificitat entre les quinases per la seva unió, podria ser particularment difícil. Amb l'objectiu de resoldre aquest inconvenient, en el capítol 4, vam realitzar una cerca sobre GSK3 β buscant noves cavitats farmacològiques. Els resultats obtinguts van assenyalar la cavitat d'unió del substracte, una cavitat en el mateix N-loop, a uns pocs Å de distància de la cavitat de l'ATP, com una molt bona cavitat que podria encabir molècules petites amb caràcter inhibitori. Això està en consonància amb altres estudis (Palomo et al., 2011; Bidon-Chanal et al., 2013), que reforcen els resultats i la metodologia emprada. En la literatura, hi ha una hipòtesi que indica **que la unió en una cavitat al·lostèrica també podria inhibir l'activitat de GSK3 β** (Palomo et al., 2011, 2017), tenint un impacte directe en la reducció de la hiperfosforilació de la tau, i per tant resultaria en la reducció de la formació de NFT i així en els símptomes de la MA. Com una cavitat al·lostèrica pot donar a les nostres molècules selectivitat, decidim provar la hipòtesi. Atès que tots els resultats obtinguts fins ara van assenyalar que les meridianines són els millors compostos entre tots els PNM's estudiats, els vam seleccionar per dur a terme un estudi detallat intentant validar la hipòtesi esmentada. A més, la lignarenone B va ser inclosa en l'estudi a causa de la seva similitud estructural amb un altre inhibidor natural no competitiu de l'ATP en la GSK3 β (Bidon-Chanal et al., 2013). Els resultats obtinguts van suggerir un millor lligam sobre la cavitat catalítica de l'ATP, en lloc d'entrar en la cavitat d'unió del substrat. Diferents raons poden explicar aquests resultats, que van des del gran volum de la cavitat del

substrat en comparació amb la cavitat ATP, fins al fet que la cavitat del substrat es col·loca en un loop, fet que es tradueix en elevades fluctuacions a causa de l'exposició al dissolvent (Nilmeier et al. , 2011). A partir d'una validació experimental, l'activitat inhibidora de les dues molècules va ser avaluada comparant la fosforilació de la Serina 9 i els nivells totals de GSK3 β . En cap dels assajos la viabilitat de GSK3 β no es va veure compromesa, trencant l'estigma del seu potencial toxicològic (capítol 1), i fins i tot el creixement de neurites es va veure augmentat. No obstant això, no podríem verificar *in silico* o *in vitro* si la inhibició es produeix en la unió a la cavitat catalítica de l'ATP, a la cavitat del substrat, o fins i tot en ambdues. L'existència d'inhibidors bivalents no és nova (Roskoski, 2016), i és una característica molt interessant que podria ser explotada en el futur.

Propietats farmacocinètiques per trobar hits d'esquelets de molècules marines

A partir del conjunt de molècules estudiades, les que provenen de la base de dades i la literatura, així com els nous compostos dissenyats al llarg d'aquesta tesi, requereixen més cura, anàlisi individual i millora de les propietats farmacocinètiques. Poden considerar-se “hits”, però estan lluny de ser “leads” potencials, entre altres raons, per les seves propietats farmacocinètiques.

Com tendència general les **propietats d'absorció** s'han de millorar, especialment segons l'unió de substrat P-glicoproteïna (PgP). No obstant això, els resultats no són dolents, podent-los considerar moderadament bons, però per fer-los òptims, es necessita una millora. Pel que fa a les **propietats de distribució**, els compostos estudiats van mostrar probabilitats altes de PPB, així com de baixa VDs, el que significa que aquests compostos tindran una gran quantitat de dificultats en la difusió o la travessa de les membranes cel·lulars. D'altra banda, la majoria de les molècules no se suposa que puguin creuar la barrera hematoencefàlica (BBB) i penetrar en el SNC, amb l'excepció dels compostos dissenyats 2a i 2e (capítol 4). Com a medicaments contra la MA, aquest és un requisit important. La incapacitat per penetrar en el sistema nerviós central (SNC), podria ser reduïda al mateix temps que les propietats d'absorció, com l'unió al PgP (actuant com un substrat i no un inhibidor que té l'efecte contrari) es relaciona amb la capacitat de penetrar en el SNC (Sadeque et al., 2000; Breedveld et al., 2006). A més a més, a part de la evident millora en l'estructura química per millorar-ne les propietats de penetració, existeix la possibilitat de facilitar la penetració de BBB, per exemple, per nanocarriers, fent ús de tècniques de nanodelivery (Bancs, 2016; Kingwell, 2016; Saraiva et al., 2016). L'anàlisi de les **propietats metabòliques** va revelar que els compostos analitzats són susceptibles de ser metabolitzats per CYP450,

la qual cosa podria resultar en reaccions adverses o errades terapèutiques, de manera que es necessita un anàlisi acurat per obtenir “hits” o “leads” més segurs. Les **proprietats de l'excreció** van mostrar bons resultats. Només uns pocs compostos semblen interactuar amb OCT2, i els compostos que interactuen estan en el llindar del model per ser classificats com a compostos d'interacció o no-interacció, de manera que s'esperen problemes de no eliminació i d'interaccions adverses amb la co-administració d'inhibidors de OCT2. Pel que fa a les **proprietats toxicològiques**, cap dels compostos estudiats indicava la inhibició de gens d'èter humà (hERG). No obstant això, s'han detectat algunes propensions d'hepatotoxicitat, mutagènica i carcinogenicitat. Aquestes propietats tòxiques han de ser eliminades més endavant.

En resum, els resultats obtinguts suggereixen la necessitat de realitzar cicles d'optimització H2L, per tal de millorar l'absorció, distribució, metabolisme i toxicitat dels compostos estudiats, així com la seva selectivitat segons els resultats d'unió comentats abans, amb l'objectiu d'obtenir compostos “lead” capaç de ser medicaments eficaços contra la MA.

Per dur a terme cicles d'optimització H2L, es requereix de nous experiments. Aquests experiments guiats per experts en **química medicinal** juntament amb eines computacionals, seran una bona opció de treball a emprar per tal de millorar totes aquestes propietats. La química medicinal juga un paper important en el procés d'optimització H2L (Hoffer et al., 2018). A causa de modificacions estructurals guiades per la química medicinal, podria ser possible l'optimització de les propietats dels compostes, proporcionant l'eficàcia desitjada, a dosis acceptables, millorant les seves propietats ADME, i minimitzant qualsevol característica toxicològica (Hann & Keseř, 2012). La química computacional és potent per si mateix, però els experiments de laboratori són sempre necessaris. La combinació de tècniques de laboratori i *in silico* és l'enfocament més potent i eficient.

Observacions finals i perspectives futures

La conclusió principal d'aquesta tesi és que les molècules marines poden ser utilitzades com a agents terapèutics contra diverses malalties, especialment la MD, i tècniques de DFAO poden ajudar a desenvolupar aquesta línia de recerca. Ens hem centrat especialment en la MA, però de fet tant els resultats (capítol 1) com la literatura, indiquen clarament que això pot ser exercit en altres malalties. Al llarg d'aquesta tesi exemplificàvem el potencial de DFAO aplicat al descobriment de fàrmacs marins. Tant els organismes bentònics antàrtics com els mediterranis han de ser considerats com a fonts riques de PNMs capaços de modular les vies relacionades amb patologies. Això no és una cosa nova, és un fet

àmpliament acceptat, però hi ha molt pocs estudis que intentin elucidar sistemàticament el perfil biològic d'un conjunt relativament gran de compostos marins. Probablement estenent l'anàlisi realitzat en el Chapter1 a un gran conjunt de dades, com tota la base de dades MarineLit, ajudarà a obtenir coneixement sobre les possibles bioactivitats de fàrmacs marins. Probablement augmentaria encara més l'ús de PNMs com a font d'inspiració en el disseny de fàrmacs.

En aquesta tesi, també es va assenyalar clarament el benefici i la utilitat dels mètodes computacionals per elucidar el potencial farmacològic dels PNMs i també per trobar "hits". Encara que son mètodes potents, per si sols no són completament concloents, quan es barregen amb experiments de laboratori, com en el capítol 4, constitueixen una protocol complert, ja que tenen capacitats predictives.

Conclusions finals

A partir dels estudis fets durant aquesta tesi, podem concloure que:

1. Meridianines, Lignarenones, Kororamides i derivatives 2a i 2e, es poden considerar compostos “hits” per combatre la MA a través de la inhibició de les GSK3 β , CK1 δ , DYRK1A, i CLK1 proteïnes quinases.
2. Els esquelets indòlics amb els substituïus halogenats, constitueixen un bon punt de partida per dissenyar hits i compostos de “leads” contra les esmentades quinases.
3. Les tècniques de DFAO són eines predictives interessants que poden ajudar a dilucidació del perfil biològic dels PNMs, així com identificar-los com a compostos “hits” per a una diana concreta o malaltia particular. En combinació amb experiments de laboratori són encara més útils.
4. La utilització de tècniques de VP en aquesta tesi va permetre l'elucidació d'una llista de possibles dianes, relacionant-les amb una determinada malaltia, com ara les patologies cardiovasculars i neurodegeneratives, i també, identificar la cavitat d'unió on les molècules marines podrien exercir la seva activitat.
5. L'avaluació de les propietats de semblança a altres fàrmacs dels nostres compostos comparant les seves energies d'unió, constitueix un pas interessant que es pot establir com una altra validació computacional de l'activitat inhibidora proposada per a altres molècules marines.
6. L'estudi detallat del mode d'unió dilucida noves idees sobre l'unió de molècules petites en GSK3 β , CK1 δ , DYRK1A i CLK1, permetent la determinació de les interaccions establertes entre cada molècula marina i la seva diana, i per tant, demostrar la seva futura aplicabilitat en el descobriment i disseny de nous potencials “hits” o “leads”.
7. Molècules com les meridianines a-G són proposades com a possibles inhibidors de les proteïnes tau específiques, GSK3 β i CK1 δ , i les quinases de doble especificitat com DYRK1A i CLK1. Aquest fet constitueix un punt de partida prometedori per al desenvolupament de nous fàrmacs contra la MA.
8. Els compostos naturals marins analitzats aquí, i els seus derivats, són una font d'inspiració molt interessant per al descobriment de nous “hits” potencials amb una potent activitat terapèutica contra proteïnes quinases implicats en la MA.

9. La presència de dos àtoms de brom a les posicions de R2 i de R3 en l'esquelet de les meridianines, va ser revelat com a sinònim de potència. En aquest sentit, i seguint un disseny racional, es van establir un parell de nous compostos brominats (2a i 2e) interessant, capaços d'exercir la millor activitat inhibidora, fet que els valida per considerar-los un bon esquelet per iniciar un procés d'optimització.
10. Altres molècules bromades, com ara kororamides A-B, amb dos àtoms de brom en el seu esquelet indòlic i, potser també convolutamines I-J, compostos heterocíclics amb tres àtoms de brom, van mostrar una activitat inhibidora contra les proteïna quinases, donant suport a les seves potencialitats com a agents terapèutics contra la MA.
11. El potencial del DFAO sobre l'elucidació de les meridianines A-G i lignarenones com possibles inhibidors tau quinasa GSK3 β va ser validat per un estudi experimental, i va recolzar la hipòtesi que les meridianines i en menor mesura, la lignarenone, podrien ser utilitzats pel tractament de les patologies d'Alzheimer, reduint la hiperfosforilació patològica provocada per la tau i reduint així la formació de l'NFT.

General References

General references

- Abdel-Magid, A. F. (2015). Allosteric Modulators: An Emerging Concept in Drug Discovery. *ACS Medicinal Chemistry Letters*, 6(2), 104–107. <https://doi.org/10.1021/ml5005365>
- Acharya, C., Coop, A., Polli, J. E., Mackerell, A. D., & Jr. (2011). Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Current Computer-Aided Drug Design*, 7(1), 10–22. <https://doi.org/10.2174/157340911793743547>
- Adayev, T., Wegiel, J., & Hwang, Y.-W. (2011). Harmine is an ATP-competitive inhibitor for dual-specificity tyrosine phosphorylation-regulated kinase 1A (Dyrk1A). *Archives of Biochemistry and Biophysics*, 507(2), 212–218. <https://doi.org/10.1016/j.abb.2010.12.024>
- Adcock, S. A., & McCammon, J. A. (2006). Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chemical Reviews*, 21062(5), 1589–1615.
- Alonso, H., Bliznyuk, A. A., & Gready, J. E. (2006). Combining docking and molecular dynamic simulations in drug design. *Medicinal Research Reviews*, 26(5), 531–568. <https://doi.org/10.1002/med.20067>
- Amaro, R. E. (2017, September 27). Toward Understanding “the Ways” of Allosteric Drugs. *ACS Central Science*, Vol. 3, pp. 925–926. <https://doi.org/10.1021/acscentsci.7b00396>
- Amaro, R. E., Baudry, J., Chodera, J., Demir, Ö., McCammon, J. A., Miao, Y., & Smith, J. C. (2018, May 22). Ensemble Docking in Drug Discovery. *Biophysical Journal*, Vol. 114, pp. 2271–2278. <https://doi.org/10.1016/j.bpj.2018.02.038>
- Amsler, C. D., McClintock, J. B., & Baker, B. J. (2001). Secondary Metabolites as Mediators of Trophic Interactions Among Antarctic Marine Organisms. *Integrative and Comparative Biology*, 41(1), 17–26. [https://doi.org/10.1668/0003-1569\(2001\)041\[0017:SMAMOT\]2.0.CO;2](https://doi.org/10.1668/0003-1569(2001)041[0017:SMAMOT]2.0.CO;2)
- Anderson, A. C. (2003, September 1). The process of structure-based drug design. *Chemistry and Biology*, Vol. 10, pp. 787–797. <https://doi.org/10.1016/j.chembiol.2003.09.002>
- Andrusier, N., Mashiach, E., Nussinov, R., & Wolfson, H. J. (2008, November 1). Principles of flexible protein-protein docking. *Proteins: Structure, Function and Genetics*, Vol. 73, pp. 271–289. <https://doi.org/10.1002/prot.22170>
- Antonov, A. S., Avilov, S. A., Kalinovsky, A. I., Anastyuk, S. D., Dmitrenok, P. S., Evtushenko, E. V., ... Stonik, V. A. (2008). Triterpene Glycosides from Antarctic Sea Cucumbers. 1. Structure of Liouvillosides A₁, A₂, A₃, B₁, and B₂ from the Sea Cucumber *Staurocucumis liouvillei*: New Procedure for Separation of Highly Polar Glycoside Fractions and Taxonomic Revision. *Journal of Natural Products*, 71(10), 1677–1685. <https://doi.org/10.1021/np800173c>
- Appeltans, W., Ah Yong, S. T., Anderson, G., Angel, M. V., Artois, T., Bailly, N., ... Costello, M. J. (2012).

General References

- The Magnitude of Global Marine Species Diversity. *Current Biology*, 22(23), 2189–2202. <https://doi.org/10.1016/J.CUB.2012.09.036>
- Appleton, D. R., Chuen, C. S., Berridge, M. V., Webb, V. L., & Copp, B. R. (2009). Rossinones A and B, biologically active meroterpenoids from the antarctic ascidian, Aplidium species. *Journal of Organic Chemistry*, 74(23), 9195–9198. <https://doi.org/10.1021/jo901846j>
- Aravindhana, G., Coote, M. L., & Barakat, K. (2017). Molecular dynamics-driven drug discovery: leaping forward with confidence. *Drug Discovery Today*, 22(2), 249–269. <https://doi.org/10.1016/J.DRUDIS.2016.11.001>
- Atanassoff, P. (2000). Ziconotide, a new n-type calcium channel blocker, administered intrathecally for acute postoperative pain*1, *2. *Regional Anesthesia and Pain Medicine*, 25(3), 274–278. [https://doi.org/10.1016/S1098-7339\(00\)90010-5](https://doi.org/10.1016/S1098-7339(00)90010-5)
- Avila, C. (2006). *Molluscan Natural Products as Biological Models: Chemical Ecology, Histology, and Laboratory Culture*. https://doi.org/10.1007/978-3-540-30880-5_1
- Avila, C. (2016). Ecological and Pharmacological Activities of Antarctic Marine Natural Products. *Planta Medica*, 82(9–10), 767–774. <https://doi.org/10.1055/s-0042-105652>
- Avila, C., Taboada, S., & Núñez-Pons, L. (2008). Antarctic marine chemical ecology: what is next? *Marine Ecology*, 29(1), 1–71. <https://doi.org/10.1111/j.1439-0485.2007.00215.x>
- Baig, M. H., Ahmad, K., Rabbani, G., Danishuddin, M., & Choi, I. (2017). Computer Aided Drug Design and its Application to the Development of Potential Drugs for Neurodegenerative Disorders. *Current Neuropharmacology*, 16(6), 740–748. <https://doi.org/10.2174/1570159x15666171016163510>
- Baker, B. J. (Bill J. . (2015). *Marine biomedicine : from beach to bedside*.
- Banks, W. A. (2016). From blood–brain barrier to blood–brain interface: new opportunities for CNS drug delivery. *Nature Reviews Drug Discovery*, 15(4), 275–292. <https://doi.org/10.1038/nrd.2015.21>
- Barre, S. La. (2010). *Marine biodiversity and chemodiversity: a tale of many different stories, with a dicey outcome*.
- Barril, X. (2017, October 3). Computer-aided drug design: time to play with novel chemical matter. *Expert Opinion on Drug Discovery*, 12(10), 977–980. <https://doi.org/10.1080/17460441.2017.1362386>
- Benkendorff, K. (2010). Molluscan biological and chemical diversity: secondary metabolites and medicinal resources produced by marine molluscs. *Biological Reviews of the Cambridge Philosophical Society*, 85(4), 757–775. <https://doi.org/10.1111/j.1469-185X.2010.00124.x>
- Bergquist, P. R. (2001). Porifera (Sponges). In *Encyclopedia of Life Sciences*.

- <https://doi.org/10.1038/npg.els.0001582>
- Bernardi, R. C., Melo, M. C. R., & Schulten, K. (2015). *Enhanced sampling techniques in molecular dynamics simulations of biological systems* ☆. <https://doi.org/10.1016/j.bbagen.2014.10.019>
- Bharate, S. B., Sawant, S. D., Singh, P. P., & Vishwakarma, R. A. (2013). Kinase Inhibitors of Marine Origin. *Chemical Reviews*, *113*(8), 6761–6815. <https://doi.org/10.1021/cr300410v>
- Bharate, S. B., Yadav, R., Battula, S., & Vishwakarma, R. . (2012). Meridianins: Marine-Derived Potent Kinase Inhibitors. *Mini-Reviews in Medicinal Chemistry*, *12*(7), 618–631. <https://doi.org/10.2174/138955712800626728>
- Bhat, R. V., Budd Haeberlein, S. L., & Avila, J. (2004). Glycogen synthase kinase 3: a drug target for CNS therapies. *Journal of Neurochemistry*, *89*(6), 1313–1317. <https://doi.org/10.1111/j.1471-4159.2004.02422.x>
- Bianchi, C. N., & Morri, C. (2000). Marine Biodiversity of the Mediterranean Sea: Situation, Problems and Prospects for Future Research. *Marine Pollution Bulletin*, *40*(5), 367–376. [https://doi.org/10.1016/S0025-326X\(00\)00027-8](https://doi.org/10.1016/S0025-326X(00)00027-8)
- Bidon-Chanal, A., Fuertes, A., Alonso, D., Pérez, D. I., Martínez, A., Luque, F. J., & Medina, M. (2013). Evidence for a new binding mode to GSK-3: Allosteric regulation by the marine compound palinurin. *European Journal of Medicinal Chemistry*, *60*, 479–489. <https://doi.org/10.1016/j.ejmech.2012.12.014>
- Billingsley, M. L., & Kincaid, R. L. (1997). Regulated phosphorylation and dephosphorylation of tau protein: effects on microtubule interaction, intracellular trafficking and neurodegeneration. *The Biochemical Journal*, *323* (Pt 3)(Pt 3), 577–591. <https://doi.org/10.1042/bj3230577>
- Blake, J. (1884). On the connection between Physiological Action and Chemical Constitution. *The Journal of Physiology*, *5*(1), 35–44. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16991361>
- Blunt, J. W., Carroll, A. R., Copp, B. R., Davis, R. A., Keyzers, R. A., & Prinsep, M. R. (2018a). Marine natural products. *Natural Product Reports*, *35*(1), 8–53. <https://doi.org/10.1039/C7NP00052A>
- Blunt, J. W., Carroll, A. R., Copp, B. R., Davis, R. A., Keyzers, R. A., & Prinsep, M. R. (2018b). Marine natural products. *Natural Product Reports*, *35*(1), 8–53. <https://doi.org/10.1039/c7np00052a>
- Branca, C., Shaw, D. M., Belfiore, R., Gokhale, V., Shaw, A. Y., Foley, C., ... Oddo, S. (2017). Dyrk1 inhibition improves Alzheimer's disease-like pathology. *Aging Cell*, *16*(5), 1146–1154. <https://doi.org/10.1111/acel.12648>
- Breedveld, P., Beijnen, J. H., & Schellens, J. H. M. (2006). Use of P-glycoprotein and BCRP inhibitors to improve oral bioavailability and CNS penetration of anticancer drugs. *Trends in Pharmacological*

General References

- Sciences*, 27(1), 17–24. <https://doi.org/10.1016/J.TIPS.2005.11.009>
- Breinbauer, R., Vetter, I. R., & Waldmann, H. (2002). From Protein Domains to Drug Candidates—Natural Products as Guiding Principles in the Design and Synthesis of Compound Libraries. *Angewandte Chemie International Edition*, 41(16), 2878. [https://doi.org/10.1002/1521-3773\(20020816\)41:16<2878::AID-ANIE2878>3.0.CO;2-B](https://doi.org/10.1002/1521-3773(20020816)41:16<2878::AID-ANIE2878>3.0.CO;2-B)
- Brooijmans, N., & Kuntz, I. D. (2003). Molecular Recognition and Docking Algorithms. *Annual Review of Biophysics and Biomolecular Structure*, 32(1), 335–373. <https://doi.org/10.1146/annurev.biophys.32.110601.142532>
- Carbone, M., Núñez-Pons, L., Paone, M., Castelluccio, F., Avila, C., & Gavagnin, M. (2012). Rossinone-related meroterpenes from the Antarctic ascidian *Aplidium fuegiense*. *Tetrahedron*, 68(18), 3541–3544. <https://doi.org/10.1016/j.tet.2012.03.013>
- Carté, B. (1996). Biomedical Potential of Marine Natural Products. *BioScience*, 46(4), 271–286. <https://doi.org/10.2307/1312834>
- Cavalli, A., Salvatella, X., Dobson, C. M., & Vendruscolo, M. (2007). Protein structure determination from NMR chemical shifts. *Proceedings of the National Academy of Sciences of the United States of America*, 104(23), 9615–9620. <https://doi.org/10.1073/pnas.0610313104>
- Cavanagh, J., Fairbrother, W. J., Palmer, I., & Skelton, N. J. (1995). *Protein NMR spectroscopy : principles and practice*. Retrieved from [https://books.google.es/books?hl=en&lr=&id=85rYGWiBJ1kC&oi=fnd&pg=PP1&dq=NMR+spectroscopy+proteins&ots=L5lqzn1Wxu&sig=k3Tqls7ZVxcguH6oXOsY4uJQ_C8#v=onepage&q=NMR spectroscopy proteins&f=false](https://books.google.es/books?hl=en&lr=&id=85rYGWiBJ1kC&oi=fnd&pg=PP1&dq=NMR+spectroscopy+proteins&ots=L5lqzn1Wxu&sig=k3Tqls7ZVxcguH6oXOsY4uJQ_C8#v=onepage&q=NMR+spectroscopy+proteins&f=false)
- Cereto-Massagué, A., Ojeda, M. J., Valls, C., Mulero, M., Garcia-Vallvé, S., & Pujadas, G. (2015). Molecular fingerprint similarity search in virtual screening. *Methods*, 71, 58–63. <https://doi.org/10.1016/J.YMETH.2014.08.005>
- Chen, B., & Frank, J. (2016). Two promising future developments of cryo-EM: capturing short-lived states and mapping a continuum of states of a macromolecule. *Microscopy*, 65(1), 69–79. <https://doi.org/10.1093/jmicro/dfv344>
- Chin, Y., Balunas, M. J., Chai, H. B., & Al., E. (2006). Drug discovery from natural sources. *AAPS*, 8(2), E239–E253. <https://doi.org/https://doi.org/10.1007/BF02854894>
- Citron, M. (2010). Alzheimer's disease: strategies for disease modification. *Nature Reviews Drug Discovery*, 9(5), 387–398. <https://doi.org/10.1038/nrd2896>
- Clark, D. E. (2005). Chapter 10 Computational Prediction of ADMET Properties: Recent Developments

- and Future Challenges. *Annual Reports in Computational Chemistry*, Vol. 1, pp. 133–151. [https://doi.org/10.1016/S1574-1400\(05\)01010-8](https://doi.org/10.1016/S1574-1400(05)01010-8)
- Clarke, A., Barnes, D. K. A., & Hodgson, D. A. (2005). How isolated is Antarctica? *Trends in Ecology & Evolution*, 20(1), 1–3. <https://doi.org/10.1016/J.TREE.2004.10.004>
- Clarke, A., & Crame, J. A. (1989). The origin of the Southern Ocean marine fauna. *Geological Society, London, Special Publications*, 47(1), 253–268. <https://doi.org/10.1144/GSL.SP.1989.047.01.19>
- Clarke, A., & Johnston, N. M. (2003). *Antarctic marine benthic diversity*. 55–57. <https://doi.org/10.1201/9780203180570-8>
- Cragg, G. M., Grothaus, P. G., & Newman, D. J. (2009). Impact of natural products on developing new anti-cancer agents. *Chemical Reviews*, 109(7), 3012–3043. <https://doi.org/10.1021/cr900019j>
- Crous-Bou, M., Minguillón, C., Gramunt, N., & Molinuevo, J. L. (2017). Alzheimer's disease prevention: from risk factors to early intervention. *Alzheimer's Research & Therapy*, 9(1), 71. <https://doi.org/10.1186/s13195-017-0297-z>
- Crum-Brown, A., & Fraser, T. R. (1868). On the Connection between Chemical Constitution and Physiological Action; with special reference to the Physiological Action of the Salts of the Ammonium Bases derived from Strychnia, Brucia, Thebaia, Codeia, Morphia, and Nicotia. *J Anat Physiol*, 2(2), 224–242.
- Cutignano, A., Avila, C., Domenech-Coll, A., d'Ippolito, G., Cimino, G., & Fontana, A. (2008). First Biosynthetic Evidence on the Phenyl-Containing Polyketides of the Marine Mollusc *Scaphander lignarius*. *Organic Letters*, 10(14), 2963–2966. <https://doi.org/10.1021/ol800877f>
- Cutignano, A., Avila, C., Rosica, A., Romano, G., Laratta, B., Domenech-Coll, A., ... Fontana, A. (2012). Biosynthesis and Cellular Localization of Functional Polyketides in the Gastropod Mollusc *Scaphander lignarius*. *ChemBioChem*, 13(12), 1759–1766. <https://doi.org/10.1002/cbic.201200287>
- Das, P. S. (2017). A REVIEW ON COMPUTER AIDED DRUG DESIGN IN DRUG DISCOVERY. *World Journal of Pharmacy and Pharmaceutical Sciences*, 279–291. <https://doi.org/10.20959/wjpps20177-9450>
- Davis, M. I., Hunt, J. P., Herrgard, S., Ciceri, P., Wodicka, L. M., Pallares, G., ... Zarrinkar, P. P. (2011). Comprehensive analysis of kinase inhibitor selectivity. *Nature Biotechnology*, 29(11), 1046–1051. <https://doi.org/10.1038/nbt.1990>
- Dayton, P. K., Mordida, B. J., & Bacon, F. (1994). Polar Marine Communities. *American Zoologist*, Vol. 34, pp. 90–99. <https://doi.org/10.2307/3883821>
- De Broyer, C., & Danis, B. (2011). How many species in the Southern Ocean? Towards a dynamic inventory of the Antarctic marine species. *Deep Sea Research Part II: Topical Studies in*

General References

- Oceanography*, 58(1–2), 5–17. <https://doi.org/10.1016/J.DSR2.2010.10.007>
- De Broyer, C., Koubbi, P., Griffiths, H. ., Raymond, B., C., U. D., A.P., V. de P., ... Y., R.-C. (2016). Biogeographic Atlas of the Southern Ocean. *Scientia Marina*, 80(1), 135–136. <https://doi.org/10.3989/scimar.2016.80n1135>
- De Vivo, M., & Cavalli, A. (2017). Recent advances in dynamic docking for drug discovery. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 7(6), e1320. <https://doi.org/10.1002/wcms.1320>
- De Vivo, M., Masetti, M., Bottegoni, G., & Cavalli, A. (2016). The Role of Molecular Dynamics and Related Methods in Drug Discovery. *Journal of Medicinal Chemistry*. <https://doi.org/10.1021/acs.jmedchem.5b01684>
- Dias, D. A., Urban, S., & Roessner, U. (2012, April 16). A Historical overview of natural products in drug discovery. *Metabolites*, Vol. 2, pp. 303–336. <https://doi.org/10.3390/metabo2020303>
- DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, 47, 20–33. <https://doi.org/10.1016/j.jhealeco.2016.01.012>
- Doig, A. J., del Castillo-Frias, M. P., Berthoumieu, O., Tarus, B., Nasica-Labouze, J., Sterpone, F., ... Derreumaux, P. (2017). Why Is Research on Amyloid- β Failing to Give New Drugs for Alzheimer's Disease? *ACS Chemical Neuroscience*, 8(7), 1435–1437. <https://doi.org/10.1021/acscemneuro.7b00188>
- Dubochet, J. (2012). Cryo-EM-the first thirty years. *Journal of Microscopy*, 245(3), 221–224. <https://doi.org/10.1111/j.1365-2818.2011.03569.x>
- Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biology*, 9, 71. <https://doi.org/10.1186/1741-7007-9-71>
- Eckert, H., & Bajorath, J. (2007). Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches. *Drug Discovery Today*, 12(5–6), 225–233. <https://doi.org/10.1016/j.drudis.2007.01.011>
- Egan, S., Thomas, T., & Kjelleberg, S. (2008). Unlocking the diversity and biotechnological potential of marine surface associated microbial communities. *Current Opinion in Microbiology*, 11(3), 219–225. <https://doi.org/10.1016/J.MIB.2008.04.001>
- Elyashberg, M. (2015, June 1). Identification and structure elucidation by NMR spectroscopy. *TrAC - Trends in Analytical Chemistry*, Vol. 69, pp. 88–97. <https://doi.org/10.1016/j.trac.2015.02.014>
- European Medicines Agency. (2007). Yondelis | European Medicines Agency. Retrieved July 30, 2019,

- from <https://www.ema.europa.eu/en/medicines/human/EPAR/yondelis>
- Faulkner, D., & Ghiselin, M. (1983). Chemical defense and evolutionary ecology of dorid nudibranchs and some other opisthobranch gastropods. *Marine Ecology Progress Series*, *13*, 295–301. <https://doi.org/10.3354/meps013295>
- Figuerola, L., Jiménez, C., Rodríguez, J., Areche, C., Chávez, R., Henríquez, M., ... Vaca, I. (2015). 3-Nitroasterric Acid Derivatives from an Antarctic Sponge-Derived *Pseudogymnoascus* sp. Fungus. *Journal of Natural Products*, *78*(4), 919–923. <https://doi.org/10.1021/np500906k>
- Figuerola, B., Núñez-Pons, L., Moles, J., & Avila, C. (2013). Feeding repellence in Antarctic bryozoans. *Naturwissenschaften*, *100*(11), 1069–1081. <https://doi.org/10.1007/s00114-013-1112-8>
- Filgueira, W., Azevedo, D., Cristina, A., & Leite, L. (2014). *Halogen Atoms in the Modern Medicinal Chemistry: Hints for the Drug Design SANDReS. Statistical Analysis of Docking Results and Scoring Functions View project Development of synthetic larvicidals for Aedes aegypti View project*. <https://doi.org/10.2174/138945010790711996>
- Fischer, E. (1894). *Einfluss der Configuration auf die Wirkung der Enzyme* (Vol. 27).
- Fiser, A. (2010). Template-based protein structure modeling. *Methods in Molecular Biology* (Clifton, N.J.), *673*, 73–94. https://doi.org/10.1007/978-1-60761-842-3_6
- Folmer, F., Jaspars, M., Dicato, M., & Diederich, M. (2008, February 1). Marine natural products as targeted modulators of the transcription factor NF- κ B. *Biochemical Pharmacology*, Vol. 75, pp. 603–617. <https://doi.org/10.1016/j.bcp.2007.07.044>
- Fontana, A., Scognamiglio, G., & Cimino, G. (1997). *Dendrinolide, a New Degraded Diterpenoid from the Antarctic Sponge Dendrilla membranosa*. <https://doi.org/10.1021/NP960712W>
- Franco, L. H., Joffé, E. B. de K., Puricelli, L., Tatian, M., Seldes, A. M., & Palermo, J. A. (1998). Indole Alkaloids from the Tunicate *Aplidium m eridianum*. *Journal of Natural Products*, *61*(9), 1130–1132. <https://doi.org/10.1021/np970493u>
- Frank, J. (2016). Generalized single-particle cryo-EM – a historical perspective. *Microscopy*, *65*(1), 3–8. <https://doi.org/10.1093/jmicro/dfv358>
- Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, *10*(5), 449–461. <https://doi.org/10.1517/17460441.2015.1032936>
- Gimeno, A., Ojeda-Montes, M., Tomás-Hernández, S., Cereto-Massagué, A., Beltrán-Debón, R., Mulero, M., ... Garcia-Vallvé, S. (2019). The Light and Dark Sides of Virtual Screening: What Is There to Know? *International Journal of Molecular Sciences*, *20*(6), 1375.

- <https://doi.org/10.3390/ijms20061375>
- Giraud, F., Alves, G., Debiton, E., Nauton, L., Th Ery, V., Durieu, E., ... Moreau, P. (2011). Synthesis, Protein Kinase Inhibitory Potencies, and in Vitro Antiproliferative Activities of Meridianin Derivatives. *J. Med. Chem*, *54*, 4474–4489. <https://doi.org/10.1021/jm200464w>
- Göckler, N., Jofre, G., Papadopoulos, C., Soppa, U., Tejedor, F. J., & Becker, W. (2009). Harmine specifically inhibits protein kinase DYRK1A and interferes with neurite formation. *FEBS Journal*, *276*(21), 6324–6337. <https://doi.org/10.1111/j.1742-4658.2009.07346.x>
- Gola, J., Obrezanova, O., Champness, E., & Segall, M. (2006). ADMET Property Prediction: The State of the Art and Current Challenges. *QSAR & Combinatorial Science*, *25*(12), 1172–1180. <https://doi.org/10.1002/qsar.200610093>
- Gonzalez, M. W., & Kann, M. G. (2012). Chapter 4: Protein Interactions and Disease. *PLoS Computational Biology*, *8*(12). <https://doi.org/10.1371/journal.pcbi.1002819>
- Gordon, A. L. (1971). *Antarctic polar front zone*. <https://doi.org/10.1029/AR015p0205>
- Grabher, P., Durieu, E., Kouloura, E., Halabalaki, M., Skaltsounis, L., Meijer, L., ... Potterat, O. (2012). Library-based Discovery of DYRK1A/CLK1 Inhibitors from Natural Product Extracts. *Planta Medica*, *78*(10), 951–956. <https://doi.org/10.1055/s-0031-1298625>
- Greener, J. G., & Sternberg, M. J. (2018, June 1). Structure-based prediction of protein allostery. *Current Opinion in Structural Biology*, Vol. 50, pp. 1–8. <https://doi.org/10.1016/j.sbi.2017.10.002>
- Gribble, G. W. (2015). Biological Activity of Recently Discovered Halogenated Marine Natural Products. *Marine Drugs*, *13*(7), 4044–4136. <https://doi.org/10.3390/md13074044>
- Grosso, C., Valentão, P., Ferreres, F., & Andrade, P. B. (2014). Review: Bioactive marine drugs and marine biomaterials for brain diseases. *Marine Drugs*, *12*(5), 2539–2589. <https://doi.org/10.3390/md12052539>
- Gul, W., & Hamann, M. T. (2005). Indole alkaloid marine natural products: An established source of cancer drug leads with considerable promise for the control of parasitic, neurological and other diseases. *Life Sciences*, *78*(5), 442–453. <https://doi.org/10.1016/J.LFS.2005.09.007>
- Haefner, B. (2003). Drugs from the deep: marine natural products as drug candidates. *Drug Discovery Today*, *8*(12), 536–544. [https://doi.org/10.1016/S1359-6446\(03\)02713-2](https://doi.org/10.1016/S1359-6446(03)02713-2)
- Hamelberg, D., Mongan, J., & McCammon, J. A. (2004). Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *Journal of Chemical Physics*, *120*(24), 11919–11929.
- Hamley, I. W. (2012). The Amyloid Beta Peptide: A Chemist’s Perspective. Role in Alzheimer’s and

- Fibrillization. *Chemical Reviews*, 112(10), 5147–5192. <https://doi.org/10.1021/cr3000994>
- Hanger, D. P., Byers, H. L., Wray, S., Leung, K. Y., Saxton, M. J., Seereeram, A., ... Anderton, B. H. (2007). Novel phosphorylation sites in Tau from Alzheimer brain support a role for casein kinase 1 in disease pathogenesis. *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.M703269200>
- Hann, M. M., & Keseř, G. M. (2012, May). Finding the sweet spot: The role of nature and nurture in medicinal chemistry. *Nature Reviews Drug Discovery*, Vol. 11, pp. 355–365. <https://doi.org/10.1038/nrd3701>
- Harvey, A. L., Edrada-Ebel, R., & Quinn, R. J. (2015). The re-emergence of natural products for drug discovery in the genomics era. *Nature Reviews Drug Discovery*, 14(2), 111–129. <https://doi.org/10.1038/nrd4510>
- Hay, M., & Fenical, W. (1996). Chemical Ecology and Marine Biodiversity: Insights and Products from the Sea. *Oceanography*, 9(1), 10–20. <https://doi.org/10.5670/oceanog.1996.21>
- Heikamp, K., & Bajorath, J. (2014, January). Support vector machines for drug discovery. *Expert Opinion on Drug Discovery*, Vol. 9, pp. 93–104. <https://doi.org/10.1517/17460441.2014.866943>
- Henderson, R., & McMullan, G. (2013). Problems in obtaining perfect images by single-particle electron cryomicroscopy of biological structures in amorphous ice. *Microscopy*, 62(1), 43–50. <https://doi.org/10.1093/jmicro/dfs094>
- Henzler-Wildman, K., & Kern, D. (2007). Dynamic personalities of proteins. *Nature*, 450(7172), 964–972.
- Hernández, F., Gómez de Barreda, E., Fuster-Matanzo, A., Lucas, J. J., & Avila, J. (2010). GSK3: A possible link between beta amyloid peptide and tau protein. *Experimental Neurology*, 223(2), 322–325. <https://doi.org/10.1016/j.expneurol.2009.09.011>
- Hernandez, F., Lucas, J. J., & Avila, J. (2012). GSK3 and Tau: Two Convergence Points in Alzheimer's Disease. *Journal of Alzheimer's Disease*, 33(s1), S141–S144. <https://doi.org/10.3233/JAD-2012-129025>
- Hetényi, C., & van der Spoel, D. (2006). Blind docking of drug-sized compounds to proteins with up to a thousand residues. *FEBS Letters*, 580(5), 1447–1450. <https://doi.org/10.1016/J.FEBSLET.2006.01.074>
- Hiltunen, M., van Groen, T., & Jolkonen, J. (2009). Functional Roles of Amyloid-β Protein Precursor and Amyloid-β Peptides: Evidence from Experimental Studies. *Journal of Alzheimer's Disease*, 18(2), 401–412. <https://doi.org/10.3233/JAD-2009-1154>
- Hince, B. (2000). *The Antarctic dictionary: a complete guide to Antarctic English*. Retrieved from <https://www.publish.csiro.au/book/2536/>

General References

- Hoffer, L., Muller, C., Roche, P., & Morelli, X. (2018, September 1). Chemistry-driven Hit-to-lead Optimization Guided by Structure-based Approaches. *Molecular Informatics*, Vol. 37. <https://doi.org/10.1002/minf.201800059>
- Holland, L. Z. (2016, February 22). Tunicates. *Current Biology*, Vol. 26, pp. R146–R152. <https://doi.org/10.1016/j.cub.2015.12.024>
- Hooper, C., Killick, R., & Lovestone, S. (2008). The GSK3 hypothesis of Alzheimer's disease. *Journal of Neurochemistry*, 104(6), 1433–1439. <https://doi.org/10.1111/j.1471-4159.2007.05194.x>
- Hou, T., Wang, J., & Li, Y. ADME evaluation in drug discovery. 8. The prediction of human intestinal absorption by a support vector machine. *Journal of Chemical Information and Modeling*, 47(6), 2408–2415. <https://doi.org/10.1021/ci7002076>
- Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling*, 51(1), 69–82. <https://doi.org/10.1021/ci100275a>
- Huang, D., Zhou, T., Lafleur, K., Nevado, C., & Caflisch, A. (2010). Kinase selectivity potential for inhibitors targeting the ATP binding site: a network analysis. *Bioinformatics*, 26(2), 198–204. <https://doi.org/10.1093/bioinformatics/btp650>
- Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. *British Journal of Pharmacology*, 162(6), 1239–1249. <https://doi.org/10.1111/j.1476-5381.2010.01127.x>
- Iken, K., Avila, C., Ciavatta, M. L., Fontana, A., & Cimino, G. (1998). Hodgsonal, a new drimane sesquiterpene from the mantle of the Antarctic nudibranch *Bathydoris hodgsoni*. *Tetrahedron Letters*, 39(31), 5635–5638. [https://doi.org/10.1016/S0040-4039\(98\)01095-8](https://doi.org/10.1016/S0040-4039(98)01095-8)
- Iken, K., Avila, C., Fontana, A., & Gavagnin, M. (2002). Chemical ecology and origin of defensive compounds in the Antarctic nudibranch *Austrodoris kerguelenensis* (Opisthobranchia: Gastropoda). *Marine Biology*, 141(1), 101–109. <https://doi.org/10.1007/s00227-002-0816-7>
- Illergård, K., Ardell, D. H., & Elofsson, A. (2009). Structure is three to ten times more conserved than sequence-A study of structural response in protein cores. *Proteins: Structure, Function, and Bioinformatics*, 77(3), 499–508. <https://doi.org/10.1002/prot.22458>
- Israilewitz, B., Baudry, J., Gullingsrud, J., Kosztin, D., & Schulten, K. (2001). Steered molecular dynamics investigations of protein function. *Journal of Molecular Graphics and Modelling*, 19(1), 13–25.
- Israilewitz, B., Gao, M., & Schulten, K. (2001). Steered molecular dynamics and mechanical functions of proteins. *Current Opinion in Structural Biology*, 11(2), 224–230. <https://doi.org/10.1016/S0959->

440X(00)00194-9

- Jain, P., Karthikeyan, C., Moorthy, N. S., Waiker, D., Jain, A., & Trivedi, P. (2014). Human CDC2-Like Kinase 1 (CLK1): A Novel Target for Alzheimer's Disease. *Current Drug Targets*, *15*(5), 539–550. <https://doi.org/10.2174/1389450115666140226112321>
- Jhoti, H., & Leach, A. R. (2007). *Structure-based drug discovery*. Springer.
- Kaczanowski, S., & Zielenkiewicz, P. (2010). Why similar protein sequences encode similar three-dimensional structures? *Theoretical Chemistry Accounts*, *125*(3–6), 643–650. <https://doi.org/10.1007/s00214-009-0656-3>
- Kametani, F., & Hasegawa, M. (2018). Reconsideration of Amyloid Hypothesis and Tau Hypothesis in Alzheimer's Disease. *Frontiers in Neuroscience*, *12*, 25. <https://doi.org/10.3389/fnins.2018.00025>
- Karplus, M. (2002). *Molecular Dynamics Simulations of Biomolecules*. <https://doi.org/10.1021/AR020082R>
- Karplus, M., & Kuriyan, J. (2005). Molecular dynamics and protein function. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(19), 6679–6685. <https://doi.org/10.1073/pnas.0408930102>
- Katsila, T., Spyroulias, G. A., Patrinos, G. P., & Matsoukas, M.-T. (2016). Computational approaches in target identification and drug discovery. *Computational and Structural Biotechnology Journal*, *14*, 177. <https://doi.org/10.1016/J.CSBJ.2016.04.004>
- Khanfar, M. A., Asal, B. A., Mudit, M., Kaddoumi, A., & El Sayed, K. A. (2009). The marine natural-derived inhibitors of glycogen synthase kinase-3 β phenylmethylene hydantoin: In vitro and in vivo activities and pharmacophore modeling. *Bioorganic & Medicinal Chemistry*, *17*(16), 6032–6039. <https://doi.org/10.1016/J.BMC.2009.06.054>
- Khotimchenko, Y. (2018, May 2). Pharmacological potential of sea cucumbers. *International Journal of Molecular Sciences*, Vol. 19. <https://doi.org/10.3390/ijms19051342>
- Kim, K. S., Fuchs, J. A., & Woodward, C. K. (1993). Hydrogen exchange identifies native-state motional domains important in protein folding. *Biochemistry*, *32*(37), 9600–9608. <https://doi.org/10.1021/bi00088a012>
- Kingwell, K. (2016). Drug delivery: New targets for drug delivery across the BBB. *Nature Reviews Drug Discovery*, *15*(2), 84–85. <https://doi.org/10.1038/nrd.2016.14>
- Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery*, *3*(11), 935–949. <https://doi.org/10.1038/nrd1549>

General References

- Kiuru, P., D'Auria, M., Muller, C., Tammela, P., Vuorela, H., & Yli-Kauhaluoma, J. (2014). Exploring Marine Resources for Bioactive Compounds. *Planta Medica*, *80*(14), 1234–1246. <https://doi.org/10.1055/s-0034-1383001>
- Klein-Junior, L., Santos Passos, C., Moraes, A., Wakui, V., Konrath, E., Nurisso, A., ... Henriques, A. (2014). Indole Alkaloids and Semisynthetic Indole Derivatives as Multifunctional Scaffolds Aiming the Inhibition of Enzymes Related to Neurodegenerative Diseases – A Focus on Psychotria L. Genus. *Current Topics in Medicinal Chemistry*, *14*(8), 1056–1075. <https://doi.org/10.2174/1568026614666140324142409>
- Klopmand, G. (1992). Concepts and applications of molecular similarity, by Mark A. Johnson and Gerald M. Maggiora, eds., John Wiley & Sons, New York, 1990, 393 pp. Price: \$65.00. *Journal of Computational Chemistry*, *13*(4), 539–540. <https://doi.org/10.1002/jcc.540130415>
- Knippschild, U., Gocht, A., Wolff, S., Huber, N., Löhler, J., & Stöter, M. (2005). The casein kinase 1 family: participation in multiple cellular processes in eukaryotes. *Cellular Signalling*, *17*(6), 675–689. <https://doi.org/10.1016/J.CELLSIG.2004.12.011>
- Koehn, F. E., & Carter, G. T. (2005). The evolving role of natural products in drug discovery. *Nature Reviews Drug Discovery*, *4*(3), 206–220. <https://doi.org/10.1038/nrd1657>
- Kolarova, M., García-Sierra, F., Bartos, A., Ricny, J., Ripova, D., & Ripova, D. (2012). Structure and Pathology of Tau Protein in Alzheimer Disease. *International Journal of Alzheimer's Disease*, *2012*, 1–13. <https://doi.org/10.1155/2012/731526>
- Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., ... Cheatham, T. E. (2000). *Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models*. <https://doi.org/10.1021/AR000033J>
- Koplovitz, G., McClintock, J., Amsler, C., & Baker, B. (2009). Palatability and chemical anti-predatory defenses in common ascidians from the Antarctic Peninsula. *Aquatic Biology*, *7*(1–2), 81–92. <https://doi.org/10.3354/ab00188>
- Korb, O., Olsson, T. S. G., Bowden, S. J., Hall, R. J., Verdonk, M. L., Liebeschuetz, J. W., & Cole, J. C. (2012). Potential and limitations of ensemble docking. *Journal of Chemical Information and Modeling*, *52*(5), 1262–1274. <https://doi.org/10.1021/ci2005934>
- Koshland, D. E. (1963). Correlation of structure and function in enzyme action. *Science (New York, N.Y.)*, *142*(3599), 1533–1541. <https://doi.org/10.1126/SCIENCE.142.3599.1533>
- Kosik, K. S. (1993). The Molecular and Cellular Biology of Tau. *Brain Pathology*, *3*(1), 39–43. <https://doi.org/10.1111/j.1750-3639.1993.tb00724.x>

- Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., & Ferrin, T. E. (1982). A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology*, *161*(2), 269–288. [https://doi.org/10.1016/0022-2836\(82\)90153-X](https://doi.org/10.1016/0022-2836(82)90153-X)
- Lambert, G. (2005). Ecology and natural history of the protochordates. *Canadian Journal of Zoology*, *83*(1), 34–50. <https://doi.org/10.1139/z04-156>
- Lapillo, M., Tuccinardi, T., Martinelli, A., Macchia, M., Giordano, A., & Poli, G. (2019). Extensive reliability evaluation of docking-based target-fishing strategies. *International Journal of Molecular Sciences*, *20*(5). <https://doi.org/10.3390/ijms20051023>
- Laskowski, R. A., Gerick, F., & Thornton, J. M. (2009). The structural basis of allosteric regulation in proteins. *FEBS Letters*, *583*(11), 1692–1698. <https://doi.org/10.1016/J.FEBSLET.2009.03.019>
- Lavecchia, A. (2015). Machine-learning approaches in drug discovery: Methods and applications. *Drug Discovery Today*, Vol. 20, pp. 318–331. <https://doi.org/10.1016/j.drudis.2014.10.012>
- Leal, M. C., Puga, J., Serôdio, J., Gomes, N. C. M., & Calado, R. (2012). Trends in the Discovery of New Marine Natural Products from Invertebrates over the Last Two Decades – Where and What Are We Bioprospecting? *PLoS ONE*, *7*(1), e30580. <https://doi.org/10.1371/journal.pone.0030580>
- Leelananda, S. P., & Lindert, S. (2016). Computational methods in drug discovery. *Beilstein Journal of Organic Chemistry*, *12*(1), 2694–2718. <https://doi.org/10.3762/bjoc.12.267>
- Lesk, A. M., & Chothia, C. (1980). How different amino acid sequences determine similar protein structures: The structure and evolutionary dynamics of the globins. *Journal of Molecular Biology*, *136*(3), 225–270. [https://doi.org/10.1016/0022-2836\(80\)90373-3](https://doi.org/10.1016/0022-2836(80)90373-3)
- Levitt, M., & Sharon, R. (1988). Accurate simulation of protein dynamics in solution. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(20), 7557–7561. <https://doi.org/10.1073/pnas.85.20.7557>
- Li, B., Liu, Y., Uno, T., & Gray, N. (2004). Creating Chemical Diversity to Target Protein Kinases. *Combinatorial Chemistry & High Throughput Screening*, *7*(5), 453–472. <https://doi.org/10.2174/1386207043328580>
- Li, G., Yin, H., & Kuret, J. (2004). Casein kinase 1 delta phosphorylates tau and disrupts its binding to microtubules. *The Journal of Biological Chemistry*, *279*(16), 15938–15945. <https://doi.org/10.1074/jbc.M314116200>
- Li, J., & Vederas, J. C. (2009). Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science*, *325*, 161–165.
- Lindequist, U. (2016). Marine-derived pharmaceuticals - challenges and opportunities. *Biomolecules and*

General References

- Therapeutics*, 24(6), 561–571. <https://doi.org/10.4062/biomolther.2016.181>
- Lionta, E., Spyrou, G., Vassilatis, D., & Cournia, Z. (2014). Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Current Topics in Medicinal Chemistry*, 14(16), 1923–1938. <https://doi.org/10.2174/1568026614666140929124445>
- Liu, J., Hu, Y., Waller, D. L., Wang, J., & Liu, Q. (2012). Natural products as kinase inhibitors. *Natural Product Reports*, 29(3), 392. <https://doi.org/10.1039/c2np00097k>
- Liu, Li, Y., Han, L., Li, J., Liu, J., Zhao, Z., ... Wang, R. (2014). PDB-wide collection of binding data: current status of the PDBbind database. *Bioinformatics*, 31(3), 405–412. <https://doi.org/10.1093/bioinformatics/btu626>
- Llorach-Pares, L., Nonell-Canals, A., Sanchez-Martinez, M., & Avila, C. (2017). Computer-Aided Drug Design Applied to Marine Drug Discovery: Meridianins as Alzheimer's Disease Therapeutic Agents. *Marine Drugs*, 15(12), 366. <https://doi.org/10.3390/md15120366>
- Llorach-Pares, L., Rodriguez, E., Giralt, A., Nonell-Canals, A., Sanchez-Martinez, M., Alberch, J., & Avila, C. (2019). Meridianins and lignarenones as potential GSK3 β inhibitors. A combined computational and experimental study. *In Prep.*
- López-Legentil, S., Turon, X., & Schupp, P. (2006). Chemical and physical defenses against predators in Cystodytes (Ascidacea). *Journal of Experimental Marine Biology and Ecology*, 332(1), 27–36. <https://doi.org/10.1016/j.jembe.2005.11.002>
- Macalino, S. J. Y., Gosu, V., Hong, S., & Choi, S. (2015, September 22). Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research*, Vol. 38, pp. 1686–1701. <https://doi.org/10.1007/s12272-015-0640-5>
- Maggioni, T., Taverna, A., Reyna, P. B., Alurralde, G., Rimondino, C., & Tatián, M. (2018). Deep-sea ascidians (Chordata, Tunicata) from the SW Atlantic: Species richness with descriptions of two new species. *Zootaxa*, 4526(1), 001–028. <https://doi.org/10.11646/zootaxa.4526.1.1>
- Mandelkow, E.-M., Biernat, J., Drewes, G., Gustke, N., Trinczek, B., & Mandelkow, E. (1995). Tau domains, phosphorylation, and interactions with microtubules. *Neurobiology of Aging*, 16(3), 355–362. [https://doi.org/10.1016/0197-4580\(95\)00025-A](https://doi.org/10.1016/0197-4580(95)00025-A)
- Manning, G., Whyte, D. B., Martinez, R., Hunter, T., & Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science (New York, N.Y.)*, 298(5600), 1912–1934. <https://doi.org/10.1126/science.1075762>
- Margalef, R. (1974). *Ecología*. Omega.
- Marshall, G. (1987). Computed-aided drug design. *Annual Review of Pharmacology and Toxicology*, Vol.

- 27, 193–213. <https://doi.org/10.1146/annurev.pa.27.040187.001205>
- Marston, A. (2011). Natural Products as a Source of Protein Kinase Activators and Inhibitors. *Current Topics in Medicinal Chemistry*, 11(11), 1333–1339. <https://doi.org/10.2174/156802611795589575>
- Martin, L., Latypova, X., Wilson, C. M., Magnaudeix, A., Perrin, M.-L., Yardin, C., & Terro, F. (2013). Tau protein kinases: Involvement in Alzheimer's disease. *Ageing Research Reviews*, 12(1), 289–309. <https://doi.org/10.1016/j.arr.2012.06.003>
- Martins, A., Vieira, H., Gaspar, H., & Santos, S. (2014). Marketed marine natural products in the pharmaceutical and cosmeceutical industries: Tips for success. *Marine Drugs*, 12(2), 1066–1101. <https://doi.org/10.3390/md12021066>
- Massova, I., & Kollman, P. A. (2000). Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspectives in Drug Discovery and Design*, 18(1), 113–135. <https://doi.org/10.1023/A:1008763014207>
- McCammon, J. A., Gelin, B. R., & Karplus, M. (1977). A Review of the Chemical Ecology of Antarctic Marine Invertebrates. *American Zoologist*, Vol. 37, pp. 329–342. <https://doi.org/10.2307/3884014>
- McClintock, J. B., & Janssen, J. (1990). Pteropod abduction as a chemical defence in a pelagic antarctic amphipod. *Nature*, 346(6283), 462–464. <https://doi.org/10.1038/346462a0>
- Medina-Franco, J. L. (2013). Advances in computational approaches for drug discovery based on natural products. *Revista Latinoamericana de Química*, 41(2), 95–110. Retrieved from http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0370-59432013000200003
- Mehbub, M. F., Lei, J., Franco, C., & Zhang, W. (2014). Marine sponge derived natural products between 2001 and 2010: trends and opportunities for discovery of bioactives. *Marine Drugs*, 12(8), 4539–4577. <https://doi.org/10.3390/md12084539>
- Meijer, L., Skaltsounis, A.-L., Magiatis, P., Polychronopoulos, P., Knockaert, M., Leost, M., ... Greengard, P. (2003). GSK-3-Selective Inhibitors Derived from Tyrian Purple Indirubins. *Chemistry & Biology*, 10(12), 1255–1266. <https://doi.org/10.1016/J.CHEMBIOL.2003.11.010>
- Meng, X.-Y., Zhang, H.-X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current Computer-Aided Drug Design*, 7(2), 146–157. <https://doi.org/10.2174/157340911795677602>
- Miller, B. R., McGee, T. D., Swails, J. M., Homeyer, N., Gohlke, H., & Roitberg, A. E. (2012). MMPBSA.py: An efficient program for end-state free energy calculations. *Journal of Chemical Theory and Computation*, (8), 3314–3321.
- Millot, C. (1989). La circulation générale en Méditerranée occidentale : Aperçu de nos connaissances et

General References

- projets d'études. *Annales de Géographie*, 98(549), 497–515.
<https://doi.org/10.3406/geo.1989.20925>
- Millot, & Taupier-Letage, I. (2005). *Circulation in the Mediterranean Sea*.
<https://doi.org/10.1007/b107143>
- Moles, J., Núñez-Pons, L., Taboada, S., Figuerola, B., Cristobo, J., & Avila, C. (2015). Anti-predatory chemical defences in Antarctic benthic fauna. *Marine Biology*, 162(9), 1813–1821.
<https://doi.org/10.1007/s00227-015-2714-9>
- Molinski, T. F., Dalisay, D. S., Lievens, S. L., & Saludes, J. P. (2009a). Drug development from marine natural products. *Nature Reviews. Drug Discovery*, 8(1), 69–85. <https://doi.org/10.1038/nrd2487>
- Molinski, T. F., Dalisay, D. S., Lievens, S. L., & Saludes, J. P. (2009b). Drug development from marine natural products. *Nature Reviews Drug Discovery*, 8(1), 69–85. <https://doi.org/10.1038/nrd2487>
- Montaser, R., & Luesch, H. (2011). Marine natural products: a new wave of drugs? *Future Medicinal Chemistry*, 3(12), 1475–1489. <https://doi.org/10.4155/fmc.11.118>
- Morra, G., & Colombo, G. (2018). *Understanding Allosterism to Design New Drugs*.
<https://doi.org/10.1002/9783527806836.ch11>
- Mulakala, C., & Viswanadhan, V. N. (2013). Could MM-GBSA be accurate enough for calculation of absolute protein/ligand binding free energies? *Journal of Molecular Graphics and Modelling*, 46, 41–51. <https://doi.org/10.1016/j.jmgm.2013.09.005>
- Murata, K., & Wolf, M. (2018). Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1862(2), 324–334.
<https://doi.org/10.1016/j.BBAGEN.2017.07.020>
- Newman, D. J., & Cragg, G. M. (2016). Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products*, 79(3), 629–661. <https://doi.org/10.1021/acs.jnatprod.5b01055>
- Newman, D. J., Cragg, G. M., & Snader, K. M. (2000). The influence of natural products upon drug discovery. *Natural Product Reports*, 17(3), 215–234. <https://doi.org/10.1039/a902202c>
- Nilmeier, J., Hua, L., Coutsiás, E. A., & Jacobson, M. P. (2011). Assessing protein loop flexibility by hierarchical Monte Carlo sampling. *Journal of Chemical Theory and Computation*, 7(5), 1564–1574.
<https://doi.org/10.1021/ct1006696>
- Núñez-Pons, L., & Avila, C. (2015). Natural products mediating ecological interactions in Antarctic benthic communities: a mini-review of the known molecules. *Natural Product Reports*, 32(7), 1114–1130. <https://doi.org/10.1039/c4np00150h>
- Núñez-Pons, L., Carbone, M., Vázquez, J., Rodríguez, J., Nieto, R. M., Varela, M. M., ... Avila, C. (2012a).

- Natural products from Antarctic colonial ascidians of the genera Aplidium and Synoicum: variability and defensive role. *Marine Drugs*, 10(8), 1741–1764. <https://doi.org/10.3390/md10081741>
- Núñez-Pons, L., Carbone, M., Vázquez, J., Rodríguez, J., Nieto, R. M., Varela, M. M., ... Avila, C. (2012b). Natural Products from Antarctic Colonial Ascidians of the Genera Aplidium and Synoicum: Variability and Defensive Role. *Marine Drugs*, 10(12), 1741–1764. <https://doi.org/10.3390/md10081741>
- Oshel, P. E., & Steele, D. H. Amphipod Paramphithoe hystrix: a micropredator on the sponge Haliclona ventilabrum. *Marine Ecology Progress Series*, Vol. 23, pp. 307–309. <https://doi.org/10.2307/24816000>
- Ou-Yang, S.-S., Lu, J.-Y., Kong, X.-Q., Liang, Z.-J., Luo, C., & Jiang, H. (2012). Computational drug discovery. *Acta Pharmacologica Sinica*, 33(9), 1131–1140. <https://doi.org/10.1038/aps.2012.109>
- Palomo, V., Perez, D. I., Roca, C., Anderson, C., Rodríguez-Muela, N., Perez, C., ... Martinez, A. (2017). Subtly Modulating Glycogen Synthase Kinase 3 β : Allosteric Inhibitor Development and Their Potential for the Treatment of Chronic Diseases. *Journal of Medicinal Chemistry*, 60(12), 4983–5001. <https://doi.org/10.1021/acs.jmedchem.7b00395>
- Palomo, V., Soteras, I., Perez, D. I., Perez, C., Gil, C., Eugenia Campillo, N., & Martinez, A. (2011). Exploring the Binding Sites of Glycogen Synthase Kinase 3. Identification and Characterization of Allosteric Modulation Cavities. *Journal of Medicinal Chemistry*, 54(24), 8461–8470. <https://doi.org/10.1021/jm200996g>
- Patel, J. S., Berteotti, A., Ronsisvalle, S., Rocchia, W., & Cavalli, A. (2014). Steered molecular dynamics simulations for studying protein-ligand interaction in cyclin-dependent kinase 5. *Journal of Chemical Information and Modeling*, 54(2), 470–480. <https://doi.org/10.1021/ci4003574>
- Paterson, I., & Anderson, E. A. (2005). CHEMISTRY: The Renaissance of Natural Products as Drug Candidates. *Science*, 310(5747), 451–453. <https://doi.org/10.1126/science.1116364>
- Pati, P., Kumar Sahu, B., & Panigrahy, R. C. (2015). Marine molluscs as a potential drug cabinet: an overview. In *Indian Journal of Geo-Marine Science* (Vol. 44).
- Paul, V. J., Puglisi, M. P., & Ritson-Williams, R. (2006). Marine chemical ecology. *Natural Product Reports*, 23(2), 153. <https://doi.org/10.1039/b404735b>
- Pauletti, P. M., Cintra, L. S., Braguine, C. G., da Silva Filho, A. A., Silva, M. L. A. E., Cunha, W. R., & Januário, A. H. (2010). Halogenated indole alkaloids from marine invertebrates. *Marine Drugs*, 8(5), 1526–1549. <https://doi.org/10.3390/md8051526>
- PDBFixer. (2019). PDBFixer. Retrieved from <https://github.com/pandegroup/pdbfixer>

General References

- Pereira, F., & Aires-de-Sousa, J. (2018). Computational Methodologies in the Exploration of Marine Natural Product Leads. *Marine Drugs*, 16(7). <https://doi.org/10.3390/md16070236>
- Perez, D. I., Gil, C., & Martinez, A. (2011). Protein kinases CK1 and CK2 as new targets for neurodegenerative diseases. *Medicinal Research Reviews*, 31(6), 924–954. <https://doi.org/10.1002/med.20207>
- Perkins, R., Fang, H., Tong, W., & Welsh, W. J. (2003). Quantitative structure-activity relationship methods: perspectives on drug discovery and toxicology. *Environmental Toxicology and Chemistry*, 22(8), 1666–1679. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12924569>
- Pires, D. E. V., Blundell, T. L., & Ascher, D. B. (2015). pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *Journal of Medicinal Chemistry*, 58(9), 4066–4072. <https://doi.org/10.1021/acs.jmedchem.5b00104>
- Potterton, A., Husseini, F. S., Southey, M. W. Y., Bodkin, M. J., Heifetz, A., Coveney, P. V., & Townsend-Nicholson, A. (2019). Ensemble-Based Steered Molecular Dynamics Predicts Relative Residence Time of A 2A Receptor Binders. *Journal of Chemical Theory and Computation*, 15(5), 3316–3330. <https://doi.org/10.1021/acs.jctc.8b01270>
- Prachayasittikul, V., Worachartcheewan, A., Shoombuatong, W., Songtawee, N., Simeon, S., Prachayasittikul, V., & Nantasenamat, C. (2015). Computer-Aided Drug Design of Bioactive Natural Products. *Current Topics in Medicinal Chemistry*, 15(18), 1780–1800. <https://doi.org/10.2174/1568026615666150506151101>
- Proksch, P. (1994). Defensive roles for secondary metabolites from marine sponges and sponge-feeding nudibranchs. *Toxicon : Official Journal of the International Society on Toxinology*, 32(6), 639–655. [https://doi.org/10.1016/0041-0101\(94\)90334-4](https://doi.org/10.1016/0041-0101(94)90334-4)
- Propst, C., & Perun, T. (1989). *Computer-Aided Drug Design: Methods and Applications* (1st ed.). Retrieved from <https://dl.acm.org/citation.cfm?id=575797>
- Puglisi, M. P., Sneed, J. M., Sharp, K. H., Ritson-Williams, R., & Paul, V. J. (2014, November 1). Marine chemical ecology in benthic environments. *Natural Product Reports*, Vol. 31, pp. 1510–1553. <https://doi.org/10.1039/c4np00017j>
- Quan, Y., Sang, P., Tao, Y., Fu, Y. X., Zhang, K. Q., Xie, Y. H., & Liu, S. Q. (2014). Protein dynamics and motions in relation to their functions: Several case studies and the underlying mechanisms. *Journal of Biomolecular Structure and Dynamics*, 32(3), 372–393. <https://doi.org/10.1080/07391102.2013.770372>
- Ramírez, D., & Caballero, J. (2016). Is It Reliable to Use Common Molecular Docking Methods for

- Comparing the Binding Affinities of Enantiomer Pairs for Their Protein Target? *International Journal of Molecular Sciences*, 17(4), 525. <https://doi.org/10.3390/ijms17040525>
- Ramírez, D., & Caballero, J. (2018). Is It Reliable to Take the Molecular Docking Top Scoring Position as the Best Solution without Considering Available Structural Data? *Molecules*, 23(5), 1038. <https://doi.org/10.3390/molecules23051038>
- Reich, S. H., & Webber, S. E. (1993). Structure-based drug design (SBDD): Every structure tells a story... *Perspectives in Drug Discovery and Design*, 1(2), 371–390. <https://doi.org/10.1007/BF02174536>
- Reig, E., & Abejón, D. (2009). Intrathecal Non-Opioid Analgesics for the Control of Pain. *Neuromodulation*, 467–481. <https://doi.org/10.1016/B978-0-12-374248-3.00036-7>
- Reyes, F., Fernández, R., Rodríguez, A., Francesch, A., Taboada, S., Ávila, C., & Cuevas, C. (2008). Aplicyanins A-F, new cytotoxic bromoindole derivatives from the marine tunicate *Aplidium cyaneum*. *Tetrahedron*, 64(22), 5119–5123. <https://doi.org/10.1016/j.tet.2008.03.060>
- Rodrigues, T., Reker, D., Schneider, P., & Schneider, G. (2016). Counting on natural products for drug design. *Nature Chemistry*, 8(6), 531–541. <https://doi.org/10.1038/nchem.2479>
- Rosenberg, G. (2014). A New Critical Estimate of Named Species-Level Diversity of the Recent Mollusca*. *American Malacological Bulletin*, 32(2), 308. <https://doi.org/10.4003/006.032.0204>
- Roskoski, R. (2016). Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacological Research*, 103, 26–48. <https://doi.org/10.1016/j.phrs.2015.10.021>
- Rudmann, D. G. (2013). On-target and Off-target-based Toxicologic Effects. *Toxicologic Pathology*, 41(2), 310–314. <https://doi.org/10.1177/0192623312464311>
- Ryoo, S.-R., Jeong, H. K., Radnaabazar, C., Yoo, J.-J., Cho, H.-J., Lee, H.-W., ... Song, W.-J. (2007). DYRK1A-mediated hyperphosphorylation of Tau. A functional link between Down syndrome and Alzheimer disease. *The Journal of Biological Chemistry*, 282(48), 34850–34857. <https://doi.org/10.1074/jbc.M707358200>
- Sadeque, A., Wandel, C., He, H., Shah, S., & Wood, A. J. J. (2000). Increased drug delivery to the brain by P-glycoprotein inhibition. *Clinical Pharmacology & Therapeutics*, 68(3), 231–237. <https://doi.org/10.1067/mcp.2000.109156>
- Sakano, T., Mahamood, M. I., Yamashita, T., & Fujitani, H. (2016). Molecular dynamics analysis to evaluate docking pose prediction. *Biophysics and Physicobiology*, 13, 181–194. https://doi.org/10.2142/biophysico.13.0_181
- Saraiva, C., Praça, C., Ferreira, R., Santos, T., Ferreira, L., & Bernardino, L. (2016). Nanoparticle-mediated

General References

- brain drug delivery: Overcoming blood–brain barrier to treat neurodegenerative diseases. *Journal of Medicinal Chemistry*. <https://doi.org/10.1016/j.jconrel.2016.05.044>
- Scher, H. D., & Martin, E. E. (2006). Timing and climatic consequences of the opening of Drake Passage. *Science (New York, N.Y.)*, *312*(5772), 428–430. <https://doi.org/10.1126/science.1120044>
- Schlitter, J., Engels, M., & Krüger, P. (1994). Targeted molecular dynamics: A new approach for searching pathways of conformational transitions. *Journal of Molecular Graphics*, *12*(2), 84–89. [https://doi.org/10.1016/0263-7855\(94\)80072-3](https://doi.org/10.1016/0263-7855(94)80072-3)
- Schwab, C., DeMaggio, A. J., Ghoshal, N., Binder, L. I., Kuret, J., & McGeer, P. L. (2000). Casein kinase 1 delta is associated with pathological accumulation of tau in several neurodegenerative diseases. *Neurobiology of Aging*, *21*(4), 503–510. [https://doi.org/10.1016/S0197-4580\(00\)00110-X](https://doi.org/10.1016/S0197-4580(00)00110-X)
- Shankar, G. M., Li, S., Mehta, T. H., Garcia-Munoz, A., Shepardson, N. E., Smith, I., ... Selkoe, D. J. (2008). Amyloid- β protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nature Medicine*, *14*(8), 837–842. <https://doi.org/10.1038/nm1782>
- Shen, J., Cheng, F., Xu, Y., Li, W., & Tang, Y. (2010). Estimation of ADME properties with substructure pattern recognition. *Journal of Chemical Information and Modeling*, *50*(6), 1034–1041. <https://doi.org/10.1021/ci100104j>
- Shoichet, B. K. (2004). Virtual screening of chemical libraries. *Nature*, *432*(7019), 862–865. <https://doi.org/10.1038/nature03197>
- Singh, T. J., Grundke-Iqbal, I., & Iqbal, K. (2002). Phosphorylation of τ Protein by Casein Kinase-1 Converts It to an Abnormal Alzheimer-Like State. *Journal of Neurochemistry*, *64*(3), 1420–1423. <https://doi.org/10.1046/j.1471-4159.1995.64031420.x>
- Sirimulla, S., Bailey, J. B., Vegesna, R., & Narayan, M. (2013, November 25). Halogen interactions in protein-ligand complexes: Implications of halogen bonding for rational drug design. *Journal of Chemical Information and Modeling*, Vol. 53, pp. 2781–2791. <https://doi.org/10.1021/ci400257k>
- Šíša, M., Pla, D., Altuna, M., Francesch, A., Cuevas, C., Albericio, F., & Álvarez, M. (2009). Total Synthesis and Antiproliferative Activity Screening of (\pm)-Aplicyanins A, B and E and Related Analogues [†]. *Journal of Medicinal Chemistry*, *52*(20), 6217–6223. <https://doi.org/10.1021/jm900544z>
- Skropeta, D., Pastro, N., & Zivanovic, A. (2011, October). Kinase inhibitors from marine sponges. *Marine Drugs*, Vol. 9, pp. 2131–2154. <https://doi.org/10.3390/md9102131>
- Slabinski, L., Jaroszewski, L., Rodrigues, A. P. C., Rychlewski, L., Wilson, I. A., Lesley, S. A., & Godzik, A. (2007). The challenge of protein structure determination--lessons from structural genomics. *Protein Science: A Publication of the Protein Society*, *16*(11), 2472–2482.

- <https://doi.org/10.1110/ps.073037907>
- Slattery, M. (2009). Bioactive compounds from echinoderms. In *Echinoderms: Durham* (pp. 591–600).
<https://doi.org/10.1201/9780203869543-c88>
- Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. (2014). Computational Methods in Drug Discovery. *PHARMACOLOGICAL REVIEWS Pharmacol Rev*, 66, 334–395.
<https://doi.org/10.1124/pr.112.007336>
- Smyth, M. S., & Martin, J. H. (2000). x ray crystallography. *Molecular Pathology: MP*, 53(1), 8–14.
<https://doi.org/10.1136/mp.53.1.8>
- Snelgrove, P. (2016). An Ocean of Discovery: Biodiversity Beyond the Census of Marine Life. *Planta Medica*, 82(09/10), 790–799. <https://doi.org/10.1055/s-0042-103934>
- Song, C. M., Lim, S. J., & Tong, J. C. (2009). Recent advances in computer-aided drug design. *Briefings in Bioinformatics*, Vol. 10, pp. 579–591. <https://doi.org/10.1093/bib/bbp023>
- Suan, M., & Khanh, B. (2013). Steered Molecular Dynamics-A Promising Tool for Drug Design. *Current Bioinformatics*, 7(4), 342–351. <https://doi.org/10.2174/157489312803901009>
- Sun, H., Duan, L., Chen, F., Liu, H., Wang, Z., Pan, P., ... Hou, T. (2018). Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. *Physical Chemistry Chemical Physics*, 20(21), 14450–14460.
<https://doi.org/10.1039/C7CP07623A>
- Tahtouh, T., Elkins, J. M., Filippakopoulos, P., Soundararajan, M., Burgy, G., Durieu, E., ... Meijer, L. (2012). Selectivity, Cocrystal Structures, and Neuroprotective Properties of Leucettines, a Family of Protein Kinase Inhibitors Derived from the Marine Sponge Alkaloid Leucettamine B. *Journal of Medicinal Chemistry*, 55(21), 9312–9330. <https://doi.org/10.1021/jm301034u>
- Talele, T., Khedkar, S., & Rigby, A. (2010). Successful Applications of Computer Aided Drug Discovery: Moving Drugs from Concept to the Clinic. *Current Topics in Medicinal Chemistry*, 10(1), 127–141.
<https://doi.org/10.2174/156802610790232251>
- Talevi, A. (2018). *Computer-Aided Drug Design: An Overview*. https://doi.org/10.1007/978-1-4939-7756-7_1
- Taylor, R. D., Jewsbury, P. J., & Essex, J. W. (2002). A review of protein-small molecule docking methods. *Journal of Computer-Aided Molecular Design*, 16(3), 151–166.
<https://doi.org/10.1023/A:1020155510718>
- Teilum, K., Olsen, J. G., & Kragelund, B. B. (2009). Functional aspects of protein flexibility. *Cellular and Molecular Life Sciences*, 66(14), 2231–2247. <https://doi.org/10.1007/s00018-009-0014-6>

General References

- Tell, V., & Hilgeroth, A. (2013). Recent developments of protein kinase inhibitors as potential AD therapeutics. *Front Cell Neurosci*, *1118*(16), 5–1. <https://doi.org/10.3389/fncel.2013.00189>
- Thakur, N., Thakur, A., & Müller, W. (2005). Marine natural products in drug discovery. *Natural Product Radiance*, *4*, 471–477.
- Tian, F., Lv, Y., & Yang, L. (2012). Structure-based prediction of protein-protein binding affinity with consideration of allosteric effect. *Amino Acids*, *43*(2), 531–543. <https://doi.org/10.1007/s00726-011-1101-1>
- Tian, Li, Y. L., & Zhao, F. C. (2017, March 1). Secondary metabolites from polar organisms. *Marine Drugs*, Vol. 15. <https://doi.org/10.3390/md15030028>
- Tobi, D., & Bahar, I. (2005). Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(52), 18908–18913. <https://doi.org/10.1073/pnas.0507603102>
- Toledo-Sherman, L. M., & Chen, D. (2002). High-throughput virtual screening for drug discovery in parallel. *Current Opinion in Drug Discovery & Development*, *5*(3), 414–421. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12058617>
- Traxler, P., & Furet, P. (1999). Strategies toward the Design of Novel and Selective Protein Tyrosine Kinase Inhibitors. *Pharmacology & Therapeutics*, *82*(2), 195–206. [https://doi.org/10.1016/S0163-7258\(98\)00044-8](https://doi.org/10.1016/S0163-7258(98)00044-8)
- Tringali, C. (2012). *Bioactive compounds from natural sources : natural products as lead compounds in drug discovery*. CRC Press.
- UNESCO. (2017). Facts and figures on marine biodiversity | United Nations Educational, Scientific and Cultural Organization. Retrieved September 14, 2019, from <http://www.unesco.org/new/en/natural-sciences/ioc-oceans/focus-areas/rio-20-ocean/blueprint-for-the-future-we-want/marine-biodiversity/facts-and-figures-on-marine-biodiversity/>
- Uriz, M. J., Martin, D., Turon, X., Ballesteros, E., Hughes, R., & Acebal, C. (1991). An approach to the ecological significance of chemically mediated bioactivity in Mediterranean benthic communities. *Marine Ecology Progress Series*, *70*, 175–188. <https://doi.org/10.2307/24816773>
- Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., ... Zhao, S. (2019, June 1). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*, Vol. 18, pp. 463–477. <https://doi.org/10.1038/s41573-019-0024-5>
- Van de Waterbeemd, H., & Gifford, E. (2003). ADMET in silico modelling: towards prediction paradise? *Nature Reviews Drug Discovery*, *2*(3), 192–204. <https://doi.org/10.1038/nrd1032>

- Van Donk, E., Ianora, A., & Vos, M. (2011). Induced defences in marine and freshwater phytoplankton: a review. *Hydrobiologia*, *668*(1), 3–19. <https://doi.org/10.1007/s10750-010-0395-4>
- Varney, M. D., Marzoni, G. P., Palmer, C. L., Deal, J. G., Webber, S., Welsh, K. M., ... Morse, C. A. (1992). Crystal-structure-based design and synthesis of benz[cd]indole-containing inhibitors of thymidylate synthase. *Journal of Medicinal Chemistry*, *35*(4), 663–676. <https://doi.org/10.1021/jm00082a006>
- Verkhivker, G. M., Bouzida, D., Gehlhaar, D. K., Rejto, P. A., Arthurs, S., Colson, A. B., ... Rose, P. W. (2000). Deciphering common failures in molecular docking of ligand-protein complexes. *Journal of Computer-Aided Molecular Design*, *14*(8), 731–751. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11131967>
- Vidal, D., & Nonell-Canals, A. (2017). ADMET models. Retrieved from <https://www.mindthebyte.com/>
- Wagman, A. S., Boyce, R. S., Brown, S. P., Fang, E., Goff, D., Jansen, J. M., ... Bussiere, D. E. (2017). Synthesis, Binding Mode, and Antihyperglycemic Activity of Potent and Selective (5-Imidazol-2-yl-4-phenylpyrimidin-2-yl)[2-(2-pyridylamino)ethyl]amine Inhibitors of Glycogen Synthase Kinase 3. *Journal of Medicinal Chemistry*, *60*(20), 8482–8514. <https://doi.org/10.1021/acs.jmedchem.7b00922>
- Wagner, U., Utton, M., Gallo, J. M., & Miller, C. C. (1996). Cellular phosphorylation of tau by GSK-3 beta influences tau binding to microtubules and microtubule organisation. *Journal of Cell Science*, *109*(6).
- Wang, G. (2006). Diversity and biotechnological potential of the sponge-associated microbial consortia. *Journal of Industrial Microbiology & Biotechnology*, *33*(7), 545–551. <https://doi.org/10.1007/s10295-006-0123-2>
- Wegiel, J., Dowjat, K., Kaczmarek, W., Kuchna, I., Nowicki, K., Frackowiak, J., ... Hwang, Y.-W. (2008). The role of overexpressed DYRK1A protein in the early onset of neurofibrillary degeneration in Down syndrome. *Acta Neuropathologica*, *116*(4), 391–407. <https://doi.org/10.1007/s00401-008-0419-6>
- Wegiel, J., Gong, C.-X., & Hwang, Y.-W. (2011). The role of DYRK1A in neurodegenerative diseases. *FEBS Journal*, *278*(2), 236–245. <https://doi.org/10.1111/j.1742-4658.2010.07955.x>
- Wermuth, C. G., Ganellin, C. R., Lindberg, P., & Mitscher, L. A. (1998). Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure and Applied Chemistry*, *70*(5), 1129–1143. <https://doi.org/10.1351/pac199870051129>
- Willett, P. (2006). Similarity-based virtual screening using 2D fingerprints. *Drug Discovery Today*, *11*(23–24), 1046–1053. <https://doi.org/10.1016/j.drudis.2006.10.005>
- Willis, A. J. (1997). Forum. *Functional Ecology*, *11*(2), 268–271. <https://doi.org/10.1111/j.1365->

General References

2435.1997.00081.x

WoRMS. World Register of Marine Species. Retrieved September 2, 2019, from <http://www.marinespecies.org/>

Xu, X., Huang, M., & Zou, X. (2018). Docking-based inverse virtual screening: methods, applications, and challenges. *Biophysics Reports*, 4(1), 1–16. <https://doi.org/10.1007/s41048-017-0045-8>

Yan, R., & Vassar, R. (2014). Targeting the β secretase BACE1 for Alzheimer's disease therapy. *The Lancet. Neurology*, 13(3), 319–329. [https://doi.org/10.1016/S1474-4422\(13\)70276-X](https://doi.org/10.1016/S1474-4422(13)70276-X)

Yang. (2010). Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discovery Today*, 15(11–12), 444–450. <https://doi.org/10.1016/J.DRUDIS.2010.03.013>

Yang, A., Baker, B. J., Grimwade, J., Leonard, A., & McClintock, J. B. (1995). Discorhabdin Alkaloids from the Antarctic Sponge *Latrunculia apicalis*. *Journal of Natural Products*, 58(10), 1596–1599. <https://doi.org/10.1021/np50124a020>

Yoshida, W. Y., Bryan, P. J., Baker, B. J., & McClintock, J. B. (1995). Pteroenone: A Defensive Metabolite of the Abducted Antarctic Pteropod *Clione antarctica*. *The Journal of Organic Chemistry*, 60(3), 780–782. <https://doi.org/10.1021/jo00108a057>

Yu, W., & Mackerell, A. D. (2017). Computer-Aided Drug Design Methods. *Methods in Molecular Biology*, 1520, 93–94. https://doi.org/10.1007/978-1-4939-6634-9_5




Zhang, X., Perez-Sanchez, H., & C. Lightstone, F. (2017). A Comprehensive Docking and MM/GBSA Rescoring Study of Ligand Recognition upon Binding Antithrombin. *Current Topics in Medicinal Chemistry*, 17(14), 1631–1639. <https://doi.org/10.2174/1568026616666161117112604>

Appendix I

A copy of the two chapters of this thesis that are already published, Chapter 2 and Chapter 3, are included below in their original format

Article

Computer-Aided Drug Design Applied to Marine Drug Discovery: Meridianins as Alzheimer's Disease Therapeutic Agents

Laura Llorach-Pares ^{1,2} , Alfons Nonell-Canals ², Melchor Sanchez-Martinez ^{2,*}  and Conxita Avila ^{1,*} 

¹ Department of Evolutionary Biology, Ecology and Environmental Sciences, Faculty of Biology and Biodiversity Research Institute (IRBio), Universitat de Barcelona, 08028 Barcelona, Catalonia, Spain; laura@mindthebyte.com

² Mind the Byte S.L., 08028 Barcelona, Catalonia, Spain; alfons@mindthebyte.com

* Correspondence: melchor@mindthebyte.com (M.S.-M.); conxita.avila@ub.edu (C.A.); Tel.: +34-934020938 (M.S.-M.); +34-934020161 (C.A.)

Received: 6 October 2017; Accepted: 14 November 2017; Published: 27 November 2017

Abstract: Computer-aided drug discovery/design (CADD) techniques allow the identification of natural products that are capable of modulating protein functions in pathogenesis-related pathways, constituting one of the most promising lines followed in drug discovery. In this paper, we computationally evaluated and reported the inhibitory activity found in meridianins A–G, a group of marine indole alkaloids isolated from the marine tunicate *Aplidium*, against various protein kinases involved in Alzheimer's disease (AD), a neurodegenerative pathology characterized by the presence of neurofibrillary tangles (NFT). Balance splitting between tau kinase and phosphate activities caused tau hyperphosphorylation and, thereby, its aggregation and NTF formation. Inhibition of specific kinases involved in its phosphorylation pathway could be one of the key strategies to reverse tau hyperphosphorylation and would represent an approach to develop drugs to palliate AD symptoms. Meridianins bind to the adenosine triphosphate (ATP) binding site of certain protein kinases, acting as ATP competitive inhibitors. These compounds show very promising scaffolds to design new drugs against AD, which could act over tau protein kinases Glycogen synthetase kinase-3 Beta (GSK3 β) and Casein kinase 1 delta (CK1 δ , CK1D or KC1D), and dual specificity kinases as dual specificity tyrosine phosphorylation regulated kinase 1 (DYRK1A) and cdc2-like kinases (CLK1). This work is aimed to highlight the role of CADD techniques in marine drug discovery and to provide precise information regarding the binding mode and strength of meridianins against several protein kinases that could help in the future development of anti-AD drugs.

Keywords: computer-aided drug discovery/design; meridianins; Alzheimer disease; protein kinases; tau protein kinases; dual specificity kinases; marine natural products

1. Introduction

Drug discovery is the process of identifying new molecules with a certain therapeutic activity. This process is very expensive in terms of money and time. Translating basic research to the market (going through drug discovery, preclinical and clinical studies) takes tens of years and costs billions of dollars. The average cost to develop a new molecular entity is estimated to be \$1.8 billion and requires about 13.5 years [1]. However, the usage of computational techniques at various stages of the drug discovery process could reduce that cost [2]. Hence, computer-aided drug discovery/design (CADD) methods are becoming very popular and during the last three decades have played a major role in the development of therapeutically important molecules [3,4]. CADD techniques cover

several aspects of the drug discovery pipeline, ranging from the selection of candidate molecules to the optimization of lead compounds. For instance, virtual profiling (VP) methods can predict the biological profile as well as mechanisms of action (MoA) of a certain molecule; molecular modelling techniques, such as docking and molecular dynamics (MD), can predict ligand–target interactions in terms of binding mode and/or binding strength, allowing discrimination between candidate compounds [5,6]; virtual screening (VS) methods are able to find analogues (similar molecules) for a given compound(s) and/or build compound libraries from an input molecule(s); hit to lead (H2L) optimization techniques are used to design new molecules, improving an existing compound; absorption, distribution, metabolism, excretion and toxicity (ADMET) prediction techniques are able to predict the physicochemical properties of a given compound, i.e., information that can be coupled to H2L techniques in order to design better and safer drugs before synthesizing them.

A common classification of these techniques is based on the nature of the input molecule. In this sense, there are two general types of CADD approaches: structure-based drug design (SBDD) and ligand-based drug design (LBDD). In SBDD, macromolecular three-dimensional (3D) target structures, usually proteins, are analysed with the aim of identifying compounds that could interact (block, inhibit or activate) with them. In LBDD, chemical compounds are analysed in order to, for instance, find chemical analogues, explore their biological and/or toxicological profile, or improve their physicochemical and pharmacological characteristics with the aim of developing drug-like compounds (Figure 1) [7,8].

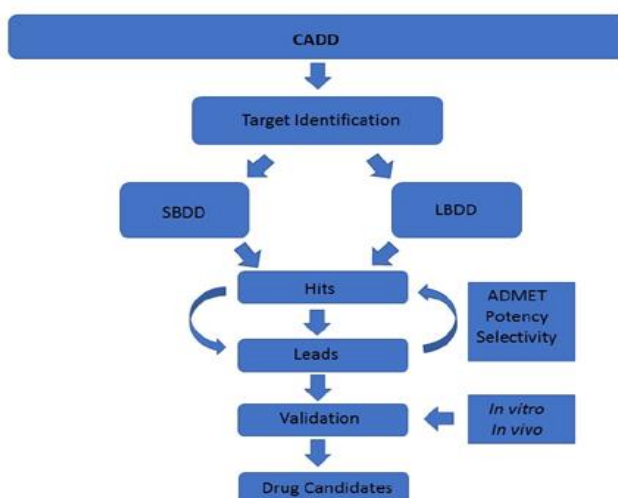


Figure 1. Schematic representation of the computer-aided drug discovery/design (CADD) techniques depicting a drug discovery pipeline.

Historically, most new drugs have been designed from natural products (secondary metabolites) and/or from compounds derived from them [9]. Natural products have thus been a rich source of compounds for drug discovery, and often, feature biologically relevant molecular scaffolds and pharmacophore patterns that have evolved as preferred ligand–protein binding motifs. The United States Food and Drug Administration (US FDA) revealed that between 1981 and 2010, 34% of those medicines approved were based on small molecules from natural products or direct derivatives of them [10,11]. The identification of natural products that are capable of modulating protein functions in pathogenesis-related pathways is one of the most promising lines followed in drug discovery [12]. Therefore, natural products constitute a huge source of inspiration in drug design [13].

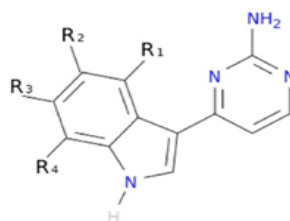
An example is Alzheimer’s disease (AD), a neurodegenerative pathology that constitutes the most common type of dementia (60–80% of the total cases), characterized by the presence of neurofibrillary tangles (NFT) primarily composed of abnormal phosphorylated tau and senile plaques

(SP). Nowadays, despite its high incidence, there is still no specific treatment approved to cure this disease. Tau phosphorylation is regulated by a balance between tau kinase and phosphate activities. Splitting of this balance was considered to cause tau hyperphosphorylation and thereby its aggregation and NTF formation [14,15]. Due to that fact, inhibition of specific tau kinases or kinases involved in tau phosphorylation pathway, could be one of the key strategies to reverse tau phosphorylation and, ultimately, fight AD [16].

The main relevant protein kinases involved in tau phosphorylation have been grouped into two classes: tau protein kinases and dual specificity kinases. The first group contains proteins such as glycogen synthetase kinase-3 beta (GSK3 β), that phosphorylates tau at different sites (specifically at 42 sites, 29 of them phosphorylated in AD brains) and casein kinase 1 delta (CK1 δ), a non-proline-directed protein kinase (non-PDPK) that regulates the microtubule dynamics through tau phosphorylation at 46 sites (25 of them phosphorylated in AD brains). The second group contains proteins such as dual specificity tyrosine phosphorylation regulated kinase 1 (DYRK1) that self-catalyse their autophosphorylation and behave as serine/threonine kinase that phosphorylates tau and the transcription factor cyclic adenosine monophosphate-response element binding (cAMP-CREB), and an evolutionarily conserved group of dual specificity kinases cdc2-like kinases (CLKs), which play an important role in the regulation of ribonucleic acid RNA splicing and are involved in the pathology of AD by phosphorylating the serine residues in arginine-rich (SR) proteins [14,15,17–19].

Among natural products, those of unexplored marine world origin are of great interest in the discovery of novel chemical structures, since they harbour most of the biodiversity of the world [20,21]. For instance, compounds from marine invertebrates may possess interesting pharmacological activities. Examples include Porifera, Cnidaria, Bryozoa, Mollusca and Tunicata [22,23]. However, although very interesting and useful from a pharmacological point of view, obtaining these compounds is difficult, both from technical and biological points of view; technically, because specimens have to be collected by hand using scuba diving or by trawling (both expensive, logistically difficult, and time consuming), and biologically, due to their marine habitats and due to the fact that they are usually unculturable [23]. All these factors, together with the adequate implementation of the Nagoya Protocol and the bioavailability of marine natural products, result in CADD contributions being highly relevant, since no biological sample is needed to perform an *in silico* analysis [24]. This also alleviates some of the marine drug discovery difficulties, such as the quantity of natural product necessary to be used in further clinical studies.

To exemplify and highlight the power of CADD techniques in marine drug discovery, as part of an ongoing study of bioactive marine molecules from benthic invertebrates, in this paper we evaluated and reported the inhibitory activity found in meridianins A–G (Figure 2), a group of marine indole alkaloids consisting of an indole framework connected to an aminopyrimidine ring, isolated from specimens of the tunicate genus *Aplidium*, against various protein kinases involved in AD.



Meridianin A	R ₁ = OH, R ₂ = H, R ₃ = H, R ₄ = H
Meridianin B	R ₁ = OH, R ₂ = H, R ₃ = Br, R ₄ = H
Meridianin C	R ₁ = H, R ₂ = Br, R ₃ = H, R ₄ = H
Meridianin D	R ₁ = H, R ₂ = H, R ₃ = Br, R ₄ = H
Meridianin E	R ₁ = OH, R ₂ = H, R ₃ = H, R ₄ = Br
Meridianin F	R ₁ = H, R ₂ = Br, R ₃ = Br, R ₄ = H
Meridianin G	R ₁ = H, R ₂ = H, R ₃ = H, R ₄ = H

Figure 2. Structures of meridianins A–G.

2. Results

2.1. Virtual Profiling

In a previous VP study (unpublished data not shown here), we observed that meridianins could bind to diverse targets involved in different diseases associated with aging or neurodegenerative pathologies, such as AD and Parkinson's disease, cancer and cardiovascular diseases (Figure 3). The found targets are of special interest as they are involved in several diseases that affect millions of people worldwide, having a huge social incidence and also, in most cases, there is no cure for them. Regarding AD, the most common disease in which meridianins could have a therapeutic role according to our results (Figure 3), GSK3 β , CK1 δ , DYRK1A and CLK1 (four kinases involved in it) could be targeted. This finding can be easily checked in the literature, confirming that meridianins can bind to these kinases. Moreover, it can also be confirmed that the target analysis results are trust worthy, since not only is the involvement of meridianins in AD disease found in the literature, but the role of meridianins as anti-cancer agents can also be easily checked [25,26].

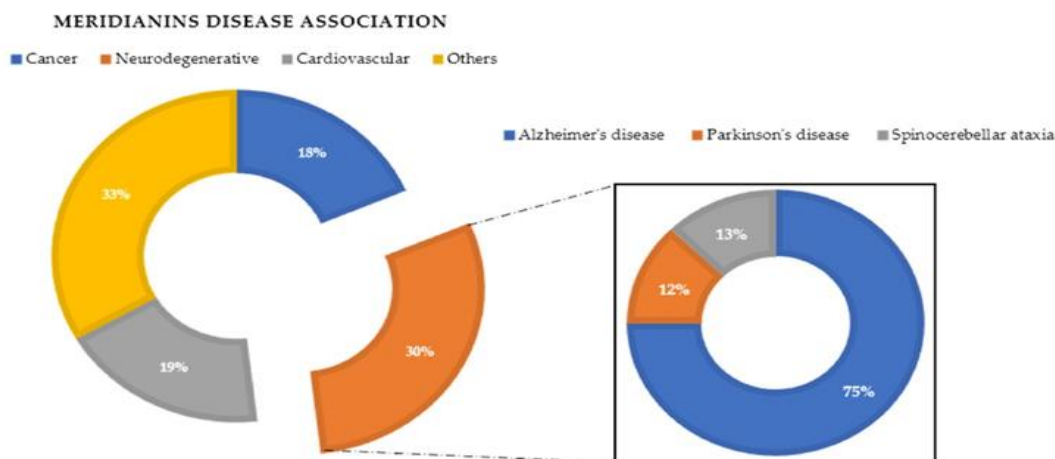


Figure 3. Disease association performed after a virtual profiling (VP) using meridianin A as a seed.

With these results in hand, the four kinases GSK3 β , CK1 δ , DYRK1A and CLK1 were selected for further analysis due to the prevalence of AD as the most common meridianin therapeutic target.

2.2. Structural and Sequence Analysis

Despite that the structural catalytic domains of most kinases are highly conserved, sequence alignment studies revealed some differences (Figure 4). The kinase catalytic domain, referred to as the hinge region, is divided into two lobes; the N-terminal mostly consists of β -sheets, whereas the C-terminal lobe is mainly helical. According to various authors, the adenosine triphosphate (ATP)-binding pocket of protein kinases can be divided into five regions: adenine region, sugar pocket, hydrophobic regions (I and II) and the phosphate-binding region [27–29]. In addition to this division, some recent studies have identified another important region: the glycine-rich loop, which is defined by the GxGxxG motif and is highly conserved among the protein kinase family. This region is suggested to significantly contribute to the potency and selectivity of binding inhibitors [29–31]. The glycine-rich loop and the hydrophobic pocket are placed in the so-called N-terminal region, while the sugar pocket and phosphate-binding region are located over the so-called C-terminal region. The adenine region is situated in the middle of these regions. We have found that meridianins are able to bind to all these regions, with a different binding strength depending on their chemical structure.

			GRR	HP	
P49841	GSK3B_HUMAN	35	SKVTTVVATFGQSPDRPQEVSYDTKY	NGSFGVVYQAKLCD-SGELV	IKVLQDKRF 93
P48730	KC1D_HUMAN	1	-----MELRVGNRYRLGRK	SSGSGFDIYLGTDIA-AGEEV	ILECVKTKH 46
Q13627	DYR1A_HUMAN	142	DDNYDYIVKNG---EKWMDRVEIDSLIGKGS	SKVKAAYDRV-EQEWV	IIKNNKAF 196
P49759	CLK1_HUMAN	144	DEEGHLICQSG---DVLRSARYEIVDTLGEQA	SKVECIDHKAGGRHV	ILVKNVDRY 199
AR					
P49841	GSK3B_HUMAN	94	KNRELQIMRKL DHCNIVRLRYFFYS SGEKKD---	EV-YLNLVLDYVPETVYRVARHYSRA	149
P48730	KC1D_HUMAN	47	-PQLHIESKIYKMMQGG-----VG	IPTIRWCGAEGDYNVMV MELLGPSLEDLF---	NFC 96
Q13627	DYR1A_HUMAN	197	LNQAQIVRLLELMNKHDTMKYYI	HLKRHF MFRN-HLCLV	MSYNLYDLLRNTNFR 255
P49759	CLK1_HUMAN	200	CEAARSIQVLEHLNTTDPNSTFR	QMLEWFEHHG-HICIV	ELGLSTYDFIKENGF 258
SP					
P49841	GSK3B_HUMAN	150	KQTLPEVIYVKLYMYQLERSLAYIHSF--	GICHRDIKPNL	LDPD----- 192
P48730	KC1D_HUMAN	97	SRKFSLKTVLLADQMISRIEYIHSKN--	FIHRDVKPNF	MGLGK----- 141
Q13627	DYR1A_HUMAN	256	--GVSLNLTTRKFAQCMCTALLFLATPELSI	IHCIDLKPNL	LCNPK----- 299
P49759	CLK1_HUMAN	259	--PFRLDHIRKMAVQICKSVNFLHSNK--	LTHTDLKPENI	IFVQSDYTEAYNPKIKRDER 314
PBP					
P49841	GSK3B_HUMAN	193	---TAVKELC	FGSAKQLVRGEPN--V-----	SYICSRYYRAPELIFGATDYTSSIDVW 241
P48730	KC1D_HUMAN	142	---GNLVYII	FGLAKKYRDARTHQHIPYRENKLTGTARYAS	INTHLGI-EQSRRDDLE 197
Q13627	DYR1A_HUMAN	300	---RSAIKI	FGSSCQL--GQRI--Y-----	QYIQSRFYRSPEVLLGM-PYDLAIDMW 345
P49759	CLK1_HUMAN	315	TLINPDKV	FGSATYD--DEHH--S-----	TLVSTRHYRAPEVILAL-GWSQPCDWW 363

Figure 4. Amino acid sequence alignment of GSK3 β , CK1 δ , DYRK1A and CLK1. In the image, only the ATP-binding pocket residues are shown. In blue, the key residues are conserved between all kinases. Green shows those conserved residues between tau protein kinases GSK3 β and CK1 δ , and red shows those conserved in dual specificity kinase DYRK1A and CLK1. Key residues refer to the residues implied in the binding of all the meridianins shared by the different targets and that are evolutionary conserved. The orange boxes represent the diverse region of the adenosine triphosphate (ATP) binding pocket. GRR: glycine-rich region; HP: hydrophobic pocket; AR: adenine region; SP: sugar pocket; PBP: phosphate binding pocket.

As explained above, we analysed two classes of protein kinases, specifically four members of them. The core catalytic regions are conserved among all as they belong to the same enzyme subclass (EC 2.7) and protein family (protein kinase). However, this protein family is divided into subfamilies: serine-threonine protein kinases (EC 2.7.11), dual-specificity kinase (EC 2.7.12), protein-histidine kinases (EC 2.7.13) and other protein kinases (EC 2.7.99). Thus, it seems logical that the binding site may be more conserved among subfamilies, and even more so in lower classifications (sub-subfamilies such as tau protein kinases (EC 2.7.11.26) and dual specificity kinase (EC 2.7.12.1)) than among the whole family. Analysing our results, we have confirmed this trend. Several key residues (associated with the substrate/inhibitor binding mode and/or conforming the pocket(s)) are conserved between the four studied proteins (Figure 4), but a higher identity is observed by pairs. GSK3 β and CK1 δ share more catalytic residues between them than with DYRK1A and CLK1, and vice versa. This observation agrees with the finding of a common binding pattern between the four protein kinases plus another pattern per each subfamily.

2.3. In Silico Binding and Interaction Analysis

Meridianins bind to the ATP binding pocket of each of the selected targets, acting as ATP competitive inhibitors. Binding energies obtained after docking and MD simulations (summarized in Table A1) show a reasonably similar binding strength between the diverse meridianins and even among the four kinases. Despite that fact, it could be observed that meridianin F tends to show higher energies than the rest of the compounds. Moreover, in general, meridianins present better binding interaction energies against CK1 δ , DYRK1A and CLK1 than GSK3 β . It must be said that these differences are hardly noticeable and cannot constitute a unique and definitive prioritization tool.

The binding mode per meridianin and target (that slightly changes between each complex) is summarized in Tables A2–A5. Comparing the interacting residues with the identified binders (summarized in Table A6), it is clearly observed that meridianins could behave as inhibitors of the analysed kinases. Moreover, analysing the observed binding mode together with the identified binders and the conserved residues (Figure 4, Tables A6 and A7), as mentioned above, some patterns of the

general binding of meridianins to protein kinases could be extracted. It has to be highlighted that the majority of the residues found in these patterns are identified as binders.

For tau protein kinases, GSK3 β and CK1 δ , 5 binding residues are shared between each of them, whereas for dual specificity kinases, DYRK1A and CLK1, 12 are conserved. Moreover, there are four residues conserved along the four analysed targets (Figure 4 and Table A7). Concretely, these residues are an alanine and a lysine placed in the hydrophobic pocket, a leucine in the sugar pocket and an aspartic acid in the phosphate binding region. Regarding tau protein kinases, there is also an isoleucine shared by GSK3 β and CK1 δ . In the case of dual specificity kinases, there are eight other shared binders, specifically, two phenylalanine, three valines, two glutamic acids and one leucine conserved and identified as binders. Analysing the meridianin binding mode by focusing on the conserved amino acids also identified as binders, we have found that two of them, A83 and K85 placed in the hydrophobic pockets, are present in all meridianin binding modes over GSK3 β and CK1 δ (in the latter case, numbered A36 and K38). For DYRK1A, three of the conserved residues are identified as key residues for the binding of all meridianins, specifically V173, L241 and L294, in the same way as for CLK1 (in this case numbering as V175, L244 and L295). In addition to these residues, others were found implicated in the general binding of meridianins not conserved through all the targets (Table A6), specifically, for GSK3 β I62, V70, L132 and D200, for CK1 δ I23, M82, L85, L135 and I148, and for DYRK1A K188, V222, F238, V306 and D307. Finally, CLK1 residues L167 and A189 were identified as key meridianin binders.

Besides the above-mentioned residues, there are other important residues per meridianin and target not present in the observed patterns that have a key role (Table A7), not dependent in a general behaviour but dependent on the particular nature of each meridianin and target (Tables A2–A5).

2.3.1. Glycogen Synthetase Kinase-3 Beta

Meridianins (Figure 5) tend to be placed within adenine (LDYV motif) and the hydrophobic regions, formed by the conserved residues A83 and K85, in the catalytic cleft. The indole scaffold of the meridianins is wrapped by N-terminal I62, F67, V70, A83, K85 and C-terminal T138, Q185, L188, D200 residues together with the LDYV motif in the hinge adenine region. Core interaction residues stabilize meridianins by establishing hydrophobic contacts with I62, V70, A83, K85, L132, D200 and hydrogen bonds with I62, K85, D200. The observed results further suggest that meridianins establish interactions over the glycine-rich loop on GSK3 β , defined by the GNGSFG motif, as well as with D200, a residue present in the phosphate pocket. The fact that meridianins bind to I62, V70, A83, K85, L132, L188 and D200, previously identified as binders, highlights meridianins inhibitory nature against GSK3 β .

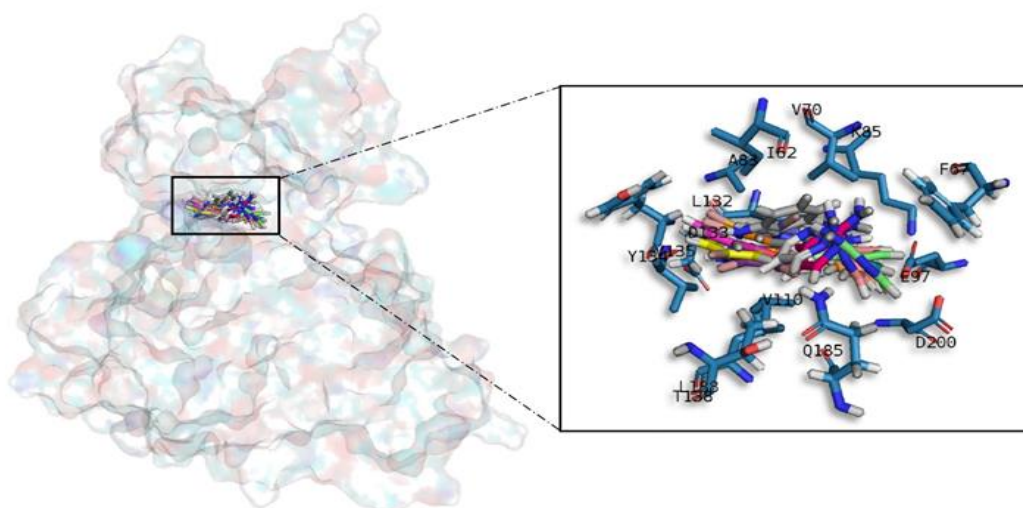


Figure 5. Meridianins A–G superposition over GSK3 β . Labelled ligand-active site amino acid residues involved in binding and the binding position of each meridianin models are enlarged.

2.3.2. Casein Kinase 1 Delta

All meridianin structures (Figure 6) share common interactions occupying the adenine region formed by the MELL motif. Meridianins are stabilized in the hinge catalytic region, establishing hydrogen bonds with A36, K38, M82, L85 and hydrophobic contacts with I23, K38, M82, L85, L135, and I148. Interestingly, it has also been observed that the indole group of the higher ranked poses has additional interactions with N-terminal I15, Y24, A36 and C-terminal D149 residues. It is important to remark that meridianins bind to the previously identified binder residues I23, A36, K38, M82, L85, L135 and I148, a fact that highlights meridianins inhibitory nature against CK1 δ .

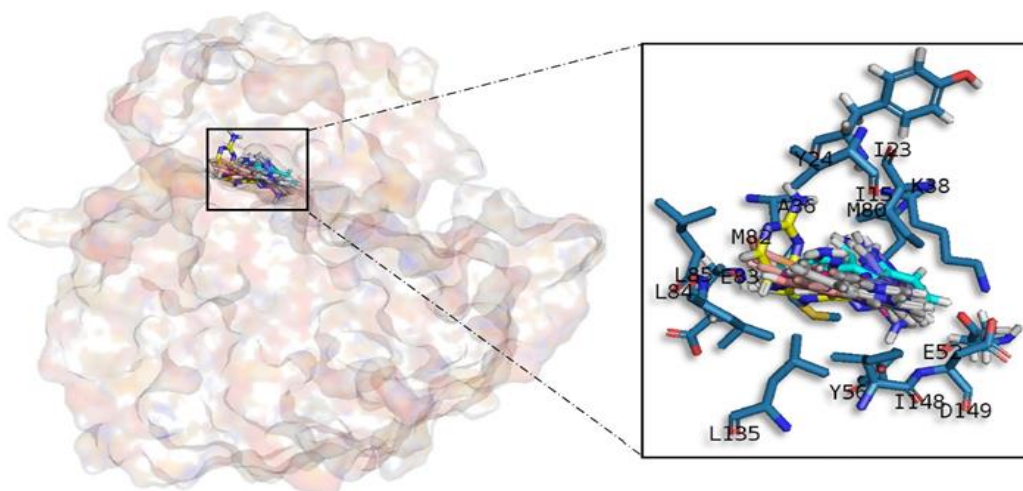


Figure 6. Meridianins A–G superposition over CK1 δ . Labelled ligand-active site amino acid residues involved in binding and the binding position of each meridianin model are enlarged.

2.3.3. Dual Specificity Tyrosine Phosphorylation Regulated Kinase 1

Meridianins are placed on the C-terminal region over the phosphate and sugar pockets as well as the adenine motif FEML (Figure 7). Despite the fact that meridianins seem to interact with the N-terminal residue V173 and the hydrophobic pocket residue K188, the rest of the key interactions are established with residues placed over the C-terminal side. Meridianins establish hydrogen bonds with K188, L241 and V307 as well as hydrophobic contacts with V173, K188, V222, F238, L241, L294, V306 and V307. Moreover, they perform π -cation and π - π stacking interactions with F238, which belongs to the adenine motif. The inhibitory effect of meridianins against DYRK1A is confirmed by the fact that all of them bind to V173, K188, V222, F238, L241, L294, V306 and V307, i.e., residues previously identified as binders.

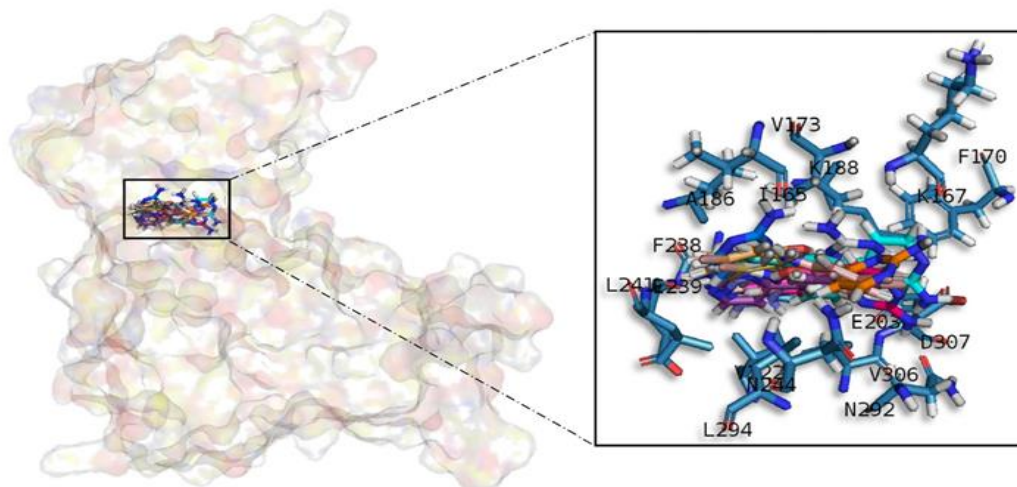


Figure 7. Meridianins A–G superposition over DYRK1A. Labelled ligand-active site amino acid residues involved in binding and the binding position of each meridianin model are enlarged.

2.3.4. Cdc2-Like Kinases

Meridianin A–G conformations against CLK1 differ by pose, as can be observed in the superimposition shown below (Figure 8). In fact, over this target is where meridianins displayed a more different conformation between the family members. In general, all poses tend to be located near the glycine rich loop and the hydrophobic pocket, interacting with the adenine motif FELL through L244 by a hydrogen bond interaction. The different poses were well stabilized into the hinge catalytic pocket by establishing hydrogen bonds interactions with L167 and L244 and hydrophobic contacts with L167, V175, A189, L244 and L295, all of them previously identified as binders, a fact that underline their inhibitory nature against CLK1.

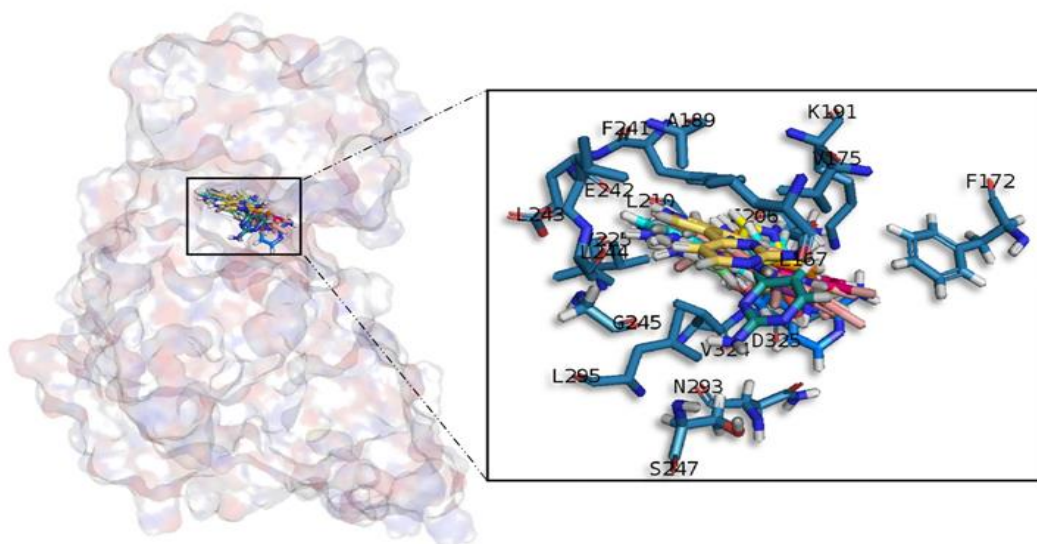


Figure 8. Meridianins A–G superposition over CLK1. Labelled ligand-active site aminoacid residues involved in binding and the binding position of each meridianin model are enlarged.

2.4. Selectivity

Since the results of *in silico* binding showed good interactions against the four studied targets, we wanted to know whether meridianins could be selective inhibitors of the studied protein families. Thus, we conducted a selectivity test consisting of analysing the meridianins binding over eight kinases (seven protein kinases and one non-protein kinase) with the aim of observing meridianin binding preference. This test included meridianins and three compounds derived from them, previously proposed as kinase inhibitors with a good selectivity for DYRK1A and CLK1 [25,32]. Our results show that meridianins and the derived compounds are able to bind to all the studied protein kinases, suggesting that they are not selective among them, although, for isocitrate dehydrogenase cytoplasmic (IDH1) and cGMP-dependent protein kinase 1 (PRKG1), slightly lower binding energies can be observed. Moreover, although compounds 1-3 tend to better interact with DYRK1A and CLK1, large differences are not observed in binding affinity between meridianins and their derived compounds (Table A8). In that sense, the derived compounds show a selectivity for DYRK1A and CLK1 respect to GSK3 β and CK1 δ , but not to all the tested kinases. Together, our results revealed the necessity to increase the selectivity of the meridianins and their, herein analysed, derived compounds.

2.5. Pharmacokinetic Properties

Due to the importance of pharmacokinetics (PK) studies during drug discovery, the whole set of meridianins and the three meridianins derived compounds were analysed, studying the ADMET properties for each molecule (Tables A9–A11).

2.5.1. Absorption Properties

In Caco-2 permeability, two different models were used as in the first one (ML model), compounds 1 and 2 cannot be evaluated because they are out of the applicability domain (OAD). All the analysed molecules have high permeability according to our proprietary model; while using pkCSM meridianin G and compounds 2 and 3 show low permeability values, but they are almost considered as high (>0.9). LogS values confirm good solubility in water and good bioavailability for each compound. Intestinal absorption shows quite good percentages (absorbance >88%) for all the studied compounds, as molecules showing values lower than 30% would be considered to be poorly absorbed. Both the

P-glycoprotein (Pgp) substrate and I/II inhibitor models show good concordance, and all of the studied molecules have been predicted to be Pgp substrates, and any of them could act as an inhibitor. The last absorption property studied was skin permeability, and results show values >-2.76 , which means reasonable low skin permeability.

2.5.2. Distribution Properties

Log P values were calculated. The steady state volume of distribution (VD_{ss}) show by the studied molecules is low, as all are above 2.81 L/Kg, Log VD_{ss} > 0.45 . For plasma protein binding (PPB) property, all the studied compounds have a probability of binding $>90\%$. Blood-brain barrier (BBB) permeability results show poor permeability for all meridianins and the three derived compounds. Compounds with a blood-brain permeability-surface area product (logPS) >-2 are considered to penetrate the central nervous system (CNS), and in that sense, compounds **2** and **3** could be considered as penetrants as they show slightly better results, i.e., logPS values of -1.88 and -1.99 , respectively. However, they are on the border, and the general tendency of all of them is to show poor penetration.

2.5.3. Metabolism Properties

Cytochrome P450 interaction reveal that all the molecules in the studied sets are likely to be metabolised. All of the analysed compounds are able to inhibit the CYP1A2 isoform. Besides meridianin F and compounds **1** and **2** can also inhibit the CYP2C19 isoform, and compounds **1** and **3** the CYP3A4 isoform. Moreover, compound **2** can act as a substrate of the CYP3A4 isoform (Table A11).

2.5.4. Excretion Properties

None of the analysed compounds is a substrate of organic cation transporter 2 (OCT2), which means that non-clearance problems and adverse interactions with co-administrated OCT2 inhibitors are expected. Moreover, total clearance was measured.

2.5.5. Toxicology Properties

Regarding the maximum recommended tolerated dose (MRTD), our results show that only meridianins A, B and E have high (greater than $0.477 \log(\text{mg/kg/day})$) MRTD values, which means that a higher dose could be administrated, while the other compounds show lower values. AMES toxicity predicts mutagenic and carcinogenic characteristics and the results reveal that only meridianins A, B and E have no apparent toxicity. The human ether-a-go-go gene (hERG) I and II inhibitor method show that any of the studied molecules is likely an hERG inhibitor. Hepatotoxicity results point out that meridianins B and F may be associated with disrupted normal function of the liver. Skin sensitisation results show no adverse effects for dermally applied products. In summary, based on all analysed compounds, only meridianins A and E seem to be non-toxic and administrable with a possible high dose without presenting adverse toxic effects.

3. Discussion

CADD techniques have an enormous potential in drug discovery, especially when they originate from marine natural products, as they do not waste natural resources. As mentioned, there are numerous different methodologies enclosed within the term CADD [2,4]. Usually the methodology is chosen based on its applicability, advantages/drawbacks, previous studies in the field, and also the expertise of the authors. In that sense, general methods such as docking, MD or ligand similarity searches have been developed, as well as more specific techniques such as disease or target models [33–44]. Each technique requires a specific input and gives a specific output, aiming to solve one step of the drug discovery pipeline (Figure 1). However, although individual CADD methods can provide insight and solve many questions, their power is their strength when combined, as we

show here. With the techniques employed in this study, we have mostly covered the drug discovery process able to be copied computationally. The methodologies we show in this work, as well as the way and the order in which we have used them, are addressed to cover a plausible general pipeline, which in our opinion is of general interest regarding marine molecules discovery. In previous years, many resources have been invested in biodiscovery (for instance, European funded projects such as PharmaSea, MaCuMBA, SeaBiotech, BlueGenics or MicroB3) and some lead compounds have been designed, but a lot of information remains stored [45–49]. Using CADD techniques, this information could be easily analysed and, potentially, employed to find drug candidates. In summary, we have shown how starting from a molecule, we were able to provide lead compounds (although in this case we provide insights to construct them instead of fully designed compounds) against a certain disease. In that sense, and as we have commented above, we exemplified the role of CADD tools applied to marine drug discovery in general, and in this particular case, analysing the role of meridianins in AD, even more specifically, against four protein kinases involved in its pathology.

The four protein kinases studied here were previously described by other authors as meridianin targets [25,32,50,51]. This constitutes an excellent validation of our computational, blind, approach to identify the biological profile of meridianins. However, although in the literature the possible anti-AD activity of meridianins was reported and several compounds have been designed from them [25,32,50,51], several aspects have not been taken into account and analysed, from a target-based (structural) perspective, as we have done here.

A common observed feature of protein kinases inhibitors is that most of them usually interact with the phosphate binding groove, in the innermost part of the pocket. This is a rich polar region, with groups such as arginine or aspartate, that consequently can create hydrogen bonds with small molecules acting as inhibitors [52]. We observed that meridianins also show this trend, supporting their already mentioned general kinase inhibitory capacity. This, in addition to the fact that most of the meridianin binding residues are previously described as binders of known inhibitors, as well as the enzymatic assays that validated meridianin binding against the four studied kinases, also reinforce their tau protein and dual specificity kinase inhibitory capacity. As mentioned above, to exert this inhibitory capacity, meridianins show general binding trends against protein kinases in general and the studied targets in particular, but also specific features related to the nature of each of the targets. The understanding of these interactions (meridianin–target) and the identification of which of these characteristics are the most important to obtain good interactions is key in the design of meridianin-derived kinase inhibitors.

It was observed that for GSK3 β , the best scored meridianins C, D, E, and F (Table A1) establish hydrophobic contacts within the aminopyrimidine ring, revealing that this scaffold could be important in having optimal interactions. This highlights the fact that the most important interactions between GSK3 β and meridianins were on the glycine rich loop and the hydrophobic and phosphate pockets. For CK1 δ , analysing our *in silico* binding results, we observed that for the best scored meridianins C, D and F (Table A1), it seems that to increase the affinity of the ligand on this receptor, the aminopyrimidine moiety should be oriented towards the top of the hydrophobic pocket at the N-terminal region. Also, key interactions were observed in the adenine and sugar-phosphate pockets. Regarding DYRK1A, meridianins mostly tend to be located over phosphate and sugar pockets as well as the adenine motif FEML rather than the glycine rich loop. Best scored meridianins B, C, E, and F (Table A1), share similar conformations but with different orientation with respect to the rest of the analysed meridianins, a fact that could be exploited for future developments together with meridianins preferential placement over the phosphate and sugar pocket. For CLK1, our molecular modelling studies have revealed that the best interacting meridianins B, C, D and F (Table A1) tend to be located near the glycine rich loop and the sugar pocket.

In general, the orientation of meridianin indole scaffolds differs from one complex to another. Its preferential positioning is directed by hydrophobic interactions and steric effects, due to the aminopyrimidine ring position. In some models, it occupies hydrophobic region I, similar to many

potent serine/threonine or tyrosine kinase inhibitors [27]. It must also be mentioned that for GSK3 β and CLK1, the preferred meridianin binding zones were located over the glycine rich loop (N-terminal). Nevertheless, over CK1 δ and DYRK1A, meridianins tend to be located over the sugar and phosphate region (both over the C-terminal region), correlating this fact with the slightly highest interacting energy observed after *in silico* binding experiments (Table A1). This could establish a new insight into future development of inhibitors.

Another interesting feature observed with respect to the meridianin binding mode is the presence of bromine. When present, interaction energies seem to be higher. The perfect example is meridianin F, which has two Br at R₂ and R₃, and has the best interaction energies for each of the studied targets with respect to the rest of meridianins. Emphasizing this issue, a pattern was observed within the two classes of kinases. For CK1 δ meridianins C (Br = R₂), D (Br = R₃) and F present the best interaction energies. In GSK3 β , meridianins D and F are among the three best interacting compounds. On DYRK1A, meridianins B (Br = R₃), C and F are three of the four best interacting compounds and in CLK1, meridianins B, C, D and F are the ones that show the best energies. All these facts led us to hypothesize that Br on R₂ and R₃ on meridianins could be synonymous with potency and has to be taken into account for the design of new lead compounds against tau and dual-specificity kinases, in particular, and protein kinases in general. Interestingly, the most promising meridianin-derived compounds already designed (by Bharate and co-workers and Giraud and co-workers), are bromine-iodo derivatives (compounds 2 and 3) and noniodinated bromine analogues (compound 1) (Figure 9) [25,32]. This fact supports our hypothesis about the influence of Br in the potency of binding showed by meridianins. According to our binding results, the derived compounds do not interact with target kinases stronger than do the meridianins. Therefore, we hypothesized that to design more potent inhibitors, the presence of Br atoms is key, but it is not enough. Playing with the different orientations and binding residues implicated in the observed patterns in meridianins-kinase binding should be also taken into account.

As protein kinases are a wide family of proteins involved in many cellular events, being selective against the desired ones is key, probably even more important than having a potent inhibitor, to avoid undesired effects. In that sense, our results show that both meridianins and the compounds reported by Bharate and co-workers, as well as Giraud and co-workers, could bind to different protein kinases with a similar strength [25,32]. In addition to that, the reported selectivity of the derived compounds for DYRK1A and CLK1 respect to GSK3 β and CK1 δ is observed, but it is not extensible to all the tested kinases. Going deeply into the results (Table A8), it could be observed that for IDH1 and PRKG1, the binding energies are slightly lower in comparison with the other targets. This fact is very relevant and could be explained because IDH1 is not a protein kinase. We put it in the pool of tested targets to see if out of the studied family, some selectivity could be observed. Regarding PRKG1, despite that it is a protein kinase member, the employed 3D structure contains an amino acid sequence that does not cover the kinase region. It was included to see what happened if despite being a protein kinase family member, the catalytic hinge region was not present. These findings allowed us to hypothesize that, despite meridianins do not show specific selectivity against any of the protein kinases tested, they do have a preferred binding to protein kinases. Moreover, this study validates the hypothesis that meridianins can act as protein kinase inhibitors. However, the low selectivity observed with respect to meridianins indicate that none of them is selective enough to properly act as AD therapeutic agent, even if able to inhibit the desired kinases. Although they could be a good starting point to design new drugs against AD, their selectivity should still be improved. To achieve that improvement, the presence of Br atoms is not enough. A rational design based on the structural differences and binding patterns observed along all meridianins should be carried out to obtain selective compounds that could have options to become an anti-AD drug. In that sense, the analysed derived compounds constitute an excellent example of how to improve meridianins to become therapeutic agents, but a new design is needed to overcome broader selectivity issues.

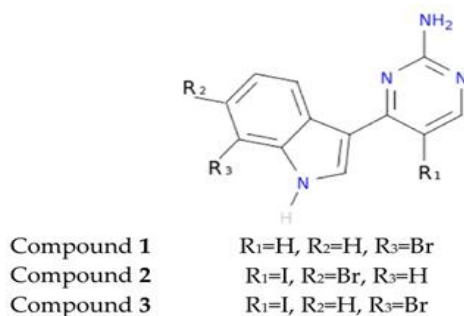


Figure 9. Structures of the three compounds. Selected from Bharate et al. Table 4—Compounds 68–70 and Giraud et al. Figure 2 —Compounds 30,33,34 [25,32].

Potency and selectivity are important characteristics of a drug, but fulfilling certain ADMET requirements is also very important. The characterization of ADMET for the molecules being pursued as potential drug candidates is essential, as clinical failures of about 50% of the drugs under investigation are due to their inadequate ADMET attributes. In this regard, we have analysed the behaviour of all the studied meridianins and also the three compounds designed by Giraud and co-workers to evaluate if the implemented modifications improve the properties of the meridianins (Tables A9–A11) [25,32].

Meridianins and the three derived compounds show a potentially high, oral and intestinal, absorbance as well as reasonable low skin permeability. Probably one of the most relevant findings is that any of the studied compounds is able to cross the BBB by itself, which is essential for a drug that should act in the brain. Good penetration was not shown in the CNS in general. In addition to CNS entrance, the Pgp that seems to play a role in amyloid beta (A β) transport across the BBB and its modulation (inhibition) has been designed as a mechanism to improve CNS pharmacotherapy [53–56]. Unfortunately, any of the studied compounds has been predicted as an inhibitor, but as a substrate, which reinforces their inability to cross the gate into the CNS. Also, in relation to distribution properties, high PPB probabilities were observed as well as a low VD_{ss}, which means these compounds will have a lot of difficulties in diffusing or traversing cell membranes.

These compounds are also able to interact with cytochrom P450, acting as inhibitors and even substrates of some isoforms, as described in the results. As it is well known that CYP450 drug metabolism can induce clinical effects, these properties should be carefully analysed in order to design lead compounds from the herein studied molecules [57]. Moreover, toxicology predictors show that the studied molecules tend to have bad toxic effects, except meridianins A and E, for which no toxicity was predicted and the maximum tolerated dose increases with respect to the rest of the studied compounds.

Together, the obtained results suggest the necessity of performing a H2L optimization, in order to improve the absorption, distribution, metabolism and toxicity of the studied compounds, as well as their selectivity, with the aim of obtaining lead compounds able to become effective anti-AD drugs.

4. Materials and Methods

4.1. Virtual Profiling

VP techniques are computational tools aimed to elucidate the biological profile of a given molecule, for instance, therapeutic indications or targets of a chemical compound could be estimated. These techniques can be ligand- or target-based. Ligand-based approaches are able to automatically evaluate very large libraries or databases of compounds containing diverse information, for example, compound–target–bioactivities associations, using a chemical structure as a seed. As a result, similar molecules (restricted by a cut-off) are found and for instance, plausible targets to the input molecule selected. In this study, meridianin A was used as a seed. To run LBDD experiments,

Cabrakan and Hurakan (Mind the Byte SL, Barcelona, Spain) software tools were employed [58,59]. Cabrakan is a two-dimensional (2D) ligand-based VP tool that compares molecules, through the use of 2D fingerprints, over a reference database and the assignment of biological activity. It allows the identification of similar chemical compounds (analogues) to the input molecule. Hurakan is a three-dimensional (3D) VP tool that compares a query molecule with the structures present in a reference database using Comparative Molecular Similarity Indices Analysis (CoMSIA) fields on a 3D grid. Hurakan can compare molecules according to their relationship with their environment, thus obtaining biomimetic compounds with different chemical structures. ChEMBL, which contains around 1,300,000 chemical compounds with detailed information including target data, was employed as the reference database [60]. A target was counted once when it appeared as both 2D and 3D hit during ligand-based VP experiments.

Here, we have employed similarity search based techniques, as they are simple, fast and accurate. However, they have the limitation imposed by the reference database employed. If there are no similar molecules to the input compound in the database, no results will be returned. This limitation is shared with other LBDD techniques such as quantitative structure–activity relationship (QSAR) or quantitative structure–property relationship (QSPR). The choice of these software tools and not another ones is based basically on the deep knowledge we have about the algorithm, the database and their performance.

Target-based approaches are able to, through knowledge of the 3D structures, evaluate huge databases that contain cavity information of these structures and after a binding site identification, docking calculations can be performed. As a result, the binding energy of every possible interaction is returned, which allows the classification and prediction of the best targets. In this study, meridianin A was used as a seed. Ixchel (Mind the Byte SL, Barcelona, Spain) is a structure-based VP tool that performs docking calculations of a molecule (spatial data file (SDF) or simplified molecular input line entry specification (SMILE) file) against an in-house developed database comprising almost 9000 protein cavities (binding-sites) curated from Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) according to UniProt Knowledgebase (UniProtKB) human entries [61–63].

To run target (or virtual) profiling experiments related to SBDD, docking is the most used technique. MD simulations or related techniques could be also employed, but they are much too computationally expensive for these kinds of techniques, with docking the preferred option. There are several variants of the docking techniques, but as we have commented for LBDD, the main limitation is the reference database. In our case, we have selected a technique whose algorithm is well known and it also incorporates a curated database of which we have a deep understanding. A deep knowledge of the employed techniques is basic and based on that, we have selected Ixchel to run our experiments.

4.2. Structure Modelling

The meridianin structures were modelled from the 2D chemical structure published by Núñez-Pons, Avila and co-workers [26]. The three meridianins derived compounds used for the selectivity test were modelled from Giraud and co-workers and Bharate et al. [25,32].

Prior to any calculation, all the structures of the selected targets, for the binding and the selectivity analysis, were modelled from their crystal structures available from the Protein Data Bank (RCSB PDB). All of them represent human targets. As obtaining good structures is crucial, the best 3D structures were selected; the structures and chains that cover the maximum amino acid region sequence, in general, and the binding region of each of the selected targets in particular.

GSK3 β was modelled from the crystallographic 3D structure with a PDB ID 3PUP that contains the crystallographic ligand OS1. It is stored in the PDB database as a homodimer, but only chain B was considered for further studies since GSK3 β biological assembly is in monomeric form [31]. CK1 δ was modelled from the 3D crystallographic structure corresponding to the entry 4KBK that contains the crystallographic ligand 1QG. Only chain B, since it is naturally a monomer, was considered for further

studies [64]. DYRK1A was modelled from the crystal 3D structure with a PDB ID 4AZE that contains the crystallographic ligand 3RA. In the PDB database, we found 3 chains (A, B and C), but only chain A was considered for further studies as DYRK1A biological assembly is in a monomeric form [52]. CLK1 was modelled from the crystallographic 3D structure with a PDB code 2VAG with V25 as a crystallographic ligand. As this protein is naturally a monomer, there is only one chain in the PDB database, so further studies were performed against chain A [52].

To test selectivity, for all the PDB crystallographic structures selected, chain A was used in all cases. Structures were modelled from their respective crystallographic 3D structure: Fibroblast growth factor receptor 1 (FGFR1); 1AGW containing SU2 as a ligand, cAMP-dependent protein kinase catalytic subunit alpha (PRKACA); 2GU8 containing 796 as a ligand, hexokinase-2 (HK2); 2NZZ containing BG6 as a ligand, dual specificity mitogen-activated protein kinase 1 (MAP2K1); 3DY7 containing ATP as a ligand, phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit gamma isoform (PIK3CG); 3IBE containing L64 as a ligand, PRKG1; 3OCP containing CMP as a ligand, serine/threonine-protein kinase N1 (PKN1); 4OTI containing MI1 as a ligand and one non protein-kinase IDH1; 4I3K containing NDP as a ligand.

To test the binding of meridianins and their selectivity, molecular modelling experiments were performed using the 3D structural models of meridianins A–G, and the models generated from the crystallographic structures available in the PDB (PDB ID 3PUP, 4KBK, 4AZE and 2VAG, respectively) and the PDB ID structures 1AGW, 2GU8, 2NZZ, 3DY7, 3IBE, 3OCP, 4OTI and 4I3K, respectively.

4.3. Docking Calculations

Docking calculations constitute a simulation method, which predicts the preferred orientation of one molecule (ligand) to a second (target). When only the movements of the first molecule are allowed, the docking is considered classical or rigid; when both molecules are allowed to move, docking is considered flexible. Generally, docking, without any other specification, refers to classical (rigid) docking [7]. Docking, in the context of small-molecule drug discovery, concerns the study of binding process of small molecules (ligands) and targets (proteins), i.e., a candidate binding mode (pose) is predicted when ligand and receptor bind to each other. Scoring functions allow us to classify and rank, based on their calculated binding energies, the most favourable pose. In that sense, flexible docking has advantages over the rigid version of the technique. The dynamics is an intrinsic characteristic of proteins, necessary to carry out any of their functions. Flexibility incorporation within the binding mode prediction is key to obtain results capable of being correlated with experimental data. However, not all are advantages, as the predicted binding energies could worsen. The inclusion of additional degrees of freedom to simulate protein flexibility could increase the difficulty of accurately predicting the free energy of binding. This complication could arise because more contributions to the free energy must be considered, for instance, the interaction between flexible residues and the core of the protein, and typically, these additional contributions also introduce additional inaccuracies [65].

Another option to add flexibility is the post-processing of docking results, which means, for instance, docking validation and/or refinement by MD simulations. Rigid docking can predict the optimal placement of a ligand within the binding site of a receptor, but not all the key interactions between the ligand and receptor are usually depicted accurately. Hence, MD simulations can optimize the predicted binding mode and also check the stability of the docked complex, as a bad docking pose will generate an unstable MD trajectory, during which the ligand could even leave the binding site [34,36]. In this study, we have employed a pipeline aimed to simulate a flexible docking protocol in a similar way to other studies reported in the literature, in that we post-processed the obtained docking poses [66]. We selected this approximation as this two-step protocol constitutes a (probably the most) practical and convenient approach to address the docking problem [67]. It is in general less computational expensive and provides the results that we need in an accurate way, comparable to “real” flexible docking methodologies (such as ensemble-based or flexible induced-fit docking).

In general, using MD as a post-processing tool, a smaller fraction of the conformational space is usually covered, but without the several limitations that affect sampling and scoring algorithms for docking.

All docking calculations were performed using Itzamna and Kin software tools (Mind the Byte SL, Barcelona, Spain) [68,69] to perform classical and blind docking calculations, respectively. Itzamna is used to carry out docking calculations and needs the structure of the molecule to dock, as well as the cavity where it should be placed as an input. Kin is a software tool designed to perform blind docking calculations. It involves a cavity search and a (best) cavity selection prior to performing the binding calculation; a difference of Itzamna is that the docking cavity is given as an input to the calculation. When the employed crystal structures were co-crystallized with a ligand, the cavity defined by the ligand was employed. As mentioned above, the modelled structures of the meridianins and the selected targets were used. Two runs were carried out for each calculation to avoid false positives.

Results obtained from docking calculations were ranked based on their calculated binding affinities, and the best poses summarized in Tables A1 and A8.

4.4. Molecular Dynamics Simulations

One of the principal tools in the computational studies of biomolecules are MD simulations, a theoretical method for studying the physical movements of atoms and molecules. MD calculates the time dependent behaviour of a molecular system, which means that atoms and molecules are allowed to interact for a fix period of time, giving a view of the dynamic evolution of the system.

Short (1 nanosecond (ns)) MD simulations were performed using NAMD program version 2.11 over the best-docked complexes, which were selected based on ΔG_{bind} [70]. The Amber ff99SB-ILDN and the General Amber Force Field (GAFF) set of parameters were employed for modelling receptors and ligands, respectively [71,72]. The election of these force-fields was based on the fact that both have been extensively tested, being two of the most used for protein and protein-ligand simulations [71–74]. It has been shown that ff99SB-ILDN correlates consistently well with experimental data, and the GAFF force-field can conveniently and quickly produce reasonable ligand (especially organic molecules) parameters. Moreover, as amber force-fields, both are compatible, giving combined satisfactory results in several studies. Ligand GAFF parameters were obtained using Antechamber, whereas the receptor structures were modelled using the leap module of Amber Tools [75,76]. Simulations were carried out in explicit solvent using the TIP3P water model with the imposition of periodic boundary conditions via a cubic box [77]. Electrostatic interactions were calculated by the particle-mesh Ewald method using constant pressure and temperature conditions. Each complex was solvated with a minimum distance of 10 Å from the surface of the complex to the edge of the box. Temperature was kept at 300 Kelvin (K) using a Langevin Piston barostat. The time step employed was 2 femtoseconds (fs). Bond lengths to hydrogens were constrained with the SHAKE algorithm [78]. Before production runs, the system was energy minimized. Next, the solvent surrounding the protein was equilibrated at the target temperature using harmonic position restraints on the heavy atoms. Finally, the system was submitted to a slow heating-up phase, from 0 to 300 K. For the production run, all position restraints were removed.

4.5. Molecular Mechanics/Generalized Born Surface Area (MM/GBSA)

The so-called reweighting techniques are computational approaches to estimate the alchemical free energy of interaction (ΔG_{bind}) between small ligands and biological macromolecules. In the literature, MM/GBSA is usually employed to estimate ligand-binding affinities based on docking or MD simulations to get a more realistic view of the interaction of docked complexes. The obtained energies are more realistic than the docking interaction values, allowing a better ranking of the analysed compounds, although they cannot be biologically comparable. In our case and following similar approaches, we applied reweighting techniques, specifically MM/GBSA, over the generated MD trajectories for post-processing docking results [34,66,79].

MM/GBSA rescoring was performed using the MMPBSA python algorithm contained within the Amber Tools suite [80]. The snapshots generated in the 1 ns MD simulation were imputed

into the post-simulation MM/GBSA calculation of binding free energy. MM/GBSA was chosen over other techniques such as molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA), linear interaction energy (LIE), thermodynamics integration (TI) or free energy perturbation (FEP) because of its good balance between accuracy and computational cost.

Rigorous thermodynamic pathway approaches, such as TI or FEP, provide more accurate predicting binding free energies, whereas LIE, MM/GBSA and MM/PBSA constitute the so-called end-point methods that in general are less accurate. Each of these methods has its own strengths and limitations, and their computational requirements and speed are inversely correlated with their accuracy. TI and FEP, which outperform end-point approaches, are very useful, especially for ranking molecules inside a chemical series. Consequently, and regardless of their computational cost but given the computational advances, these techniques are gradually being more frequently used in the drug discovery pipeline, especially in guiding lead optimisation. However, in this study, our aim is not to provide a detailed library of lead compounds, and thus we have employed a less rigorous, but very popular approach in SBDD, alternative as the MMGBSA approach. The main problem of these techniques could be that the efficacy of the method is usually system dependent. However, it is generally accepted that they outperform docking results, so a better ranking of the analysed compounds will be always obtained, although, as commented above, the obtained binding energies could be far from being experimentally comparable.

4.6. Interaction Analysis

To analyse the key residues of the active site involved in the inhibitor binding, we examined the obtained binding modes after molecular modelling studies with already known binders of each of the targets. These binders (residues that have been revealed as necessary for the binding of known substrates/inhibitors) were identified through an evidence-based interaction analysis. It was carried out through a bibliographical search plus a database analysis. The bibliographical search was conducted using several studies in which inhibitors against the selected kinases were identified describing each compound binding mode [25,31,32,50–52,64,81,82]. The database search was done using an in-house, recently constructed database. It was built by crossing ChEMBL and the RCSB PDB [62], and it contains all PDB structures per UniProtKB ID with active compounds (by now there are only PDBs with compounds not competing against cofactors). Moreover, the database also contains the residues to which each active compound (per PDB) is bound. Thus, it allows the user, after docking or an MD calculation, to easily check whether the analysed molecules behave as a binder.

4.7. Sequence Analysis

The four targets were aligned using the UniProtKB clustal omega interface from the amino acid sequence associated with each UniProtKB entry.

4.8. Selectivity Analysis

Docking calculations of meridianins, as well as the three selected compounds (derived from them and described in the literature), against twelve protein kinases were performed. These meridianins derived molecules were obtained from the papers of Bharate et al. and Giraud et al. [25,32], and have shown interesting inhibitory concentration (IC_{50}) values in the micro and sub-micromolar range, and a good selectivity for DYRK1A and CLK1. We selected them to see how the selectivity was taken into account in the design of these compounds as they strongly resemble the original meridianins scaffolds that we suspect are not selective enough.

To test the selectivity, we choose seven protein kinases, specifically, FGFR1, PRKACA, HK2, MAP2K1, PIK3CG, PRKG1 (for which the selected crystal structure do not contain the catalytic hinge), PKN1 and one non-protein kinase, IDH1. Thus, we tested if the selected compounds are selective between different protein kinases, belonging to different subfamilies, and between protein and non-protein kinases. Moreover, we explored if without the catalytic hinge, binding could be produced.

4.9. ADMET Properties Prediction

For the meridianins and the three derived compounds, ADMET properties prediction was carried out using proprietary machine-learning (ML) models and the pkCSM webserver [83,84]. The proprietary ML models covered logS (molecular aqueous coefficient), logP (octanol/water partition coefficient), Caco2 permeability, BBB penetration and PPB. The first two models were generated by super vector regression (SVR) techniques and the last three employed support vector machines (SVM). For training and testing the models, ChEMBL (logS, logP, Caco2) and Huuskonen (logS) datasets were employed, and for BBB and PPB, the datasets described by Muehlbacher et al. and Zhu and coworkers [85–87]. The pkCSM webserver allows the prediction of PK properties based on (I) compound general properties (including molecular properties, toxicophores and pharmacophore) and (II) distance-based graph signatures. Given an input molecule, both sources of information are used to train and test machine learning-based predictors. The webserver is composed of 28 (not all employed in this work) regression and classification ML models that have been generated and trained against 30 datasets (described at Pires et al.) [84].

The use of proprietary models, some of which are also covered by pkCSM, is because these methods, similar to other such as VS or VP, strongly rely on the employed reference dataset. As we have a deeper knowledge of our methods, we prefer to use them when possible. Only for Caco2 did we employ both models, ours and the pkCSM model, because for two compounds, our model is not good enough to make a reliable prediction (they are out of the applicability domain as they are too different with respect to the molecular fragments contained in the dataset employed to generate and train the model. If less than 90% of the molecular fragments in that the input molecule can be decomposed are not in the database, the prediction is not done). pkCSM predicted properties for all the compounds; however, it does not indicate if a prediction is out of the applicability domain.

In summary, we have analysed 21 ADMET properties, 5 of which were studied with our proprietary ML models and 17 with pkCSM. One of these properties, Caco2, was analysed twice using both our proprietary model and the pkCSM model.

4.9.1. Absorption Properties

Caco2 permeability, LogS, intestinal absorption (human), P-glycoprotein substrate, P-glycoprotein I/II inhibitor and skin permeability. Caco-2 permeability is used to predict the absorption of orally administered drugs. A high permeability is assessed when the predicted value is >0.90 for the pkCSM model, or high (H), in the proprietary model. LogS reflects the solubility of the molecule in water at 25 °C and also reflects the bioavailability of a given compound; it is represented by the logarithm of the molar concentration (log mol/L). Intestinal absorption indicates the portion of compounds absorbed through the human intestine; a molecule with an absorbance (intestinal absorption) of less than 30% is considered to be poorly absorbed. Pgp acts as a biological barrier by extruding toxins and xenobiotics out of cells, although it could have other, transport mediated, functions in certain tissues and organs. The predictor assesses whether a given compound is likely to be a substrate of Pgp. Pgp I and II inhibitors have significant PK implications for Pgp substrate, and the predictor will determine the inhibitory effect of a given compound against Pgp I/II, which could have advantages that can be exploited therapeutically, or result in contraindications. Skin permeability predicts if a given compound is likely to be skin permeable ($\log K_p > -2.5$).

4.9.2. Distribution

LogP, VDss, PPB, BBB and CNS permeability. LogP allows us to estimate the distribution of a drug within the body (lipophilicity). VDss, which is the theoretical volume that the total dose of a drug would need to be uniformly distributed to give the same concentration as in blood and plasma, is considered low if $\log VD_{ss} < -0.15$ and high if >0.45 (the higher the VD, the greater the drug distribution in tissue rather than plasma). PPB estimates the probability ($>90\%$ is considered high) that

a given molecule binds to a plasma protein, the less bound a drug is, the more efficiently it can traverse cell membranes or diffuse. BBB permeability describes the ability of a drug to cross into the brain. The predictor describes whether a compound is able to cross the BBB. CNS permeability measures blood brain permeability surface-area (logPS), and it is similar to BBB but more direct, as it lacks the systemic distribution effects that may distort brain penetration. Compounds with a logPS >-2 are considered to penetrate CNS, while those with logPS <-3 are considered unable to penetrate.

4.9.3. Metabolism

CYP450. Cytochrom P450 isoforms are important detoxification enzymes in the body and are essential for the metabolism of many medications. Drugs can be inhibitors of CYP450, blocking its metabolic activity, or can be metabolised (substrate) by them. CYP metabolism predictor assess whether a given molecule is likely to be metabolised or not and act as inhibitor of specific isoforms of CYP450; a specific inhibitor of CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4 and/or substrate of CYP2D6 and CYP3A4.

4.9.4. Excretion

Renal OCT2 substrate and Total Clearance. OCT2 is a renal uptake transporter that plays an important role in disposition and renal clearance of drugs and endogenous compounds. The OCT2 substrate predictor indicates if a given molecule is likely to be an OCT2 substrate, which provides not only clearance-related information but potential contraindications. Total clearance is related to bioavailability and is also important for determining dosing rates to achieve steady-state concentrations, and the predictor measures their value in log(mL/min/kg).

4.9.5. Toxicology

MRTD, AMES toxicity, hepatotoxicity, skin sensitization, hERG I/II inhibitors. MRTD provides an estimated of the toxic dose threshold of chemicals in humans, and results less than or equal to 0.477 log(mg/kg/day) are considered low, and high when greater than 0.477 log(mg/kg/day). AMES toxicity indicates if a compound could be mutagenic and therefore may act as a carcinogen. hERG I and II inhibitor predictors determine if a given compound is likely to be a hERG I/II inhibitor as the inhibition of potassium channels encoded by hERG could result in fatal pathologies (for instance it is the principal cause of the development of acquiring long QT syndrome, fatal arrhythmia) and the withdrawal of many substances from the pharmaceutical market. Hepatotoxicity predicts if a given molecule is likely to be associated with disrupted normal function of the liver. Skin permeability predicts if a given compound is likely to be associated with skin sensitisation.

4.10. Graphical Representations

Graphical representations of protein-ligand complexes were prepared using PyMOL version 1.7 [88] and PLIP version 1.3.0 [89].

5. Conclusions

Meridianins can be classified as kinase inhibitors and can be used as a starting point to design and develop novel anti-AD drugs. It has been demonstrated, *in silico* and *in vitro*, that they are able to bind specific tau (GSK3 β and CK1 δ) and dual-specificity (DYRK1A and CLK1) protein kinases. However, they are not selective enough to constitute a therapeutic treatment against AD by themselves. In fact, as they are demonstrated to be protein kinase inhibitors, they could probably inhibit several kinases involved in different diseases [90]. In any case, they could serve as a starting scaffold to design new anti-AD drugs. To achieve that, a rational design taking advantage of the differences found in the binding patterns against different protein-kinases subfamilies, has to be carried out. In that sense, the presence of Br on R₂ and the R₃ position over the meridianin indole

scaffold could be synonymous with potency. Besides, it seems that exploiting the C-terminal region (sugar and phosphate pocket) rather than the N-terminal side, could increase the strength of the interactions exerted by meridianins, and probably the potency shown by the designed compounds. However, although potency is important, and maintaining the presence of Br seems to be fairly accomplished [25,32], the selectivity between protein-kinase subfamilies is a crucial point to design proper anti-AD drugs, and even anti-cancer drugs. Meridianins are not selective enough and should be improved to gain functionality and applicability. In addition, their measured ADMET properties indicate that they should be optimized in order to become a drug or at least a drug-lead compound. Therefore, the above-mentioned rational design in order to improve the potency and selectivity of meridianins should include H2L optimization cycles. The showed toxicity should be removed, and compounds interaction with Cytochrom P450 carefully analysed and, given the case, eliminated or modulated. Moreover, their distribution properties should be improved, lowering the PPB and VDss, to be able to diffuse and penetrate into cells easily. Besides, a mechanism to cross the BBB should be found and in that sense, modifying each compound to be Pgp inhibitors could be a possible strategy, although there are other mechanisms to overcome the BBB, including other protein binding and nanodelivery, that could be also exploited [91–93].

Regarding meridianins specifically and CADD methods in general, we can conclude that these techniques, despite their drawbacks, are very helpful in drug discovery, constituting a powerful tool that could save time and money in experiments. Our study with meridianins is an example of this, since we have been able to find plausible targets, that in the case of AD and cancer we have already validated through the literature. The key role that these techniques could have in drug discovery is even higher for the discovery and development of marine drugs, since no sample is needed to run these virtual experiments. Moreover, since these methods could point out the best direction to follow and in which targets expand the low sample amount that usually is available, these are crucial technologies to maximize the success of marine prospection, as well as to protect biodiversity.

Acknowledgments: This research was partially supported by an Industrial Doctorate grant from the Generalitat of Catalonia to L.L.-P (DI 2016-051).

Author Contributions: C.A. and M.S.-M. conceived the study and designed the experiments. L.L.-P. carried out the experiments whereas A.N.-C., C.A. and M.S.-M. supervised them. All the authors analyzed and discussed the results as well as wrote the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Summary of classical rigid docking and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the two best models selected per meridianins A–G. All energies values are kcal/mol.

GSK3 β		CK1 δ		DYRK1A		CLK1	
Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)
R0/R1		R0/R1		R0/R1		R0/R1	
A	-7.3/-7.3 -6.6/-6.1	A	-6.9/-6.9 -6.8/-6.8	A	-7.4/-7.3 -7.5/-7.4	A	-8.9/-8.9 -7.8/-7.8
B	-7.3/-7.2 -6.8/-6.7	B	-6.4/-6.4 -5.6/-5.5	B	-7.7/-6.9 -7.3/-7.9	B	-8.5/-8.5 -8.0/-8.0
C	-7.6/-7.6 -7.4/-7.5	C	-6.9/-6.9 -6.9/-6.7	C	-8.2/-8.2 -7.6/-7.6	C	-8.5/-8.5 -8.1/-8.1
D	-7.7/-7.7 -7.0/-6.9	D	-7.0/-7.0 -6.8/-6.6	D	-7.9/-7.9 -7.5/-7.6	D	-8.6/-8.6 -8.1/-8.1
E	-7.3/-7.3 -7.5/-7.5	E	-7.0/-7.0 -7.0/-7.0	E	-7.5/-7.4 -7.6/-7.4	E	-9.0/-8.8 -7.9/-7.9
	-26.43 -24.95		-32.25 -30.10		-28.00 -31.43		-27.49 -30.70
	-29.11 -29.25		-35.06 -32.30		-37.38 -34.03		-34.14 -30.38
	-28.54 -31.44		-38.85 -35.84		-31.95 -35.90		-33.31 -34.92
	-31.19 -30.01		-38.69 -38.06		-29.47 -34.59		-33.58 -35.90
	-31.26 -28.43		-35.20 -34.97		-35.62 -32.55		-26.39 -31.63

Table A1. Cont.

GSK3 β		CK1 δ		DYRK1A		CLK1	
Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)	Binding Energy (kcal/mol)	MM/GBSA (kcal/mol)
R0/R1		R0/R1		R0/R1		R0/R1	
F	-7.9/-7.9 -35.18	F	-7.2/-7.3 -38.55	F	-8.0/-7.8 -39.99	F	-8.7/-8.7 -37.71
	-7.7/-7.9 -34.73		-7.1/-7.1 -38.93		-7.8/-7.7 -39.91		-8.5/-8.5 -37.61
G	-7.3/-7.3 -24.04	G	-6.8/-6.8 -31.92	G	-8.1/-8.1 -30.17	G	-9.1/-9.1 -27.95
	-7.2/-7.2 -27.03		-6.9/-6.9 -32.94		-8.1/-8.1 -30.52		-8.7/-8.7 -29.88

To avoid false positives, each docking calculation was performed twice (R0 and R1).

Table A2. GSK3 β residues that interacted with meridianins (each represented by letters A–G) after molecular dynamics (MD) simulations. Those residues involved in all meridianin binding are bold and were considered important binding residues.

Residues	A	B	C	D	E	F	G
I62	X	X	X	X	X	X	X
F67		X		X	X		X
V70	X	X	X	X	X	X	X
A83	X	X	X	X	X	X	X
K85	X	X	X	X	X	X	X
E97	X			X	X	X	X
V110	X	X	X		X	X	X
L132	X	X	X	X	X	X	X
D133	X			X		X	X
Y134	X	X	X	X		X	
V135	X	X	X		X		X
T138		X	X		X		
Q185				X			
L188	X	X	X	X	X	X	
D200	X	X	X	X	X	X	X

Table A3. CK1 δ residues that interacted with meridianins (each represented by letters A–G) after MD simulations. Those residues involved in all meridianin binding are in bold and were considered important binding residues.

Residues	A	B	C	D	E	F	G
I15	X	X	X	X	X		
I23	X	X	X	X	X	X	X
Y24		X					
A36	X	X	X	X	X	X	X
K38	X	X	X	X	X	X	X
E52		X	X		X		
Y56	X			X		X	X
M80				X		X	X
M82	X	X	X	X	X	X	X
E83	X	X			X		
L84	X	X		X			
L85	X	X	X	X	X	X	X
L135	X	X	X	X	X	X	X
I148	X	X	X	X	X	X	X
D149	X	X	X	X	X		X

Table A4. DYRK1A residues that interacted with meridianins (each represented by letters A–G) after MD simulations. Those residues involved in all meridianin binding are in bold and were considered important binding residues.

Residues	A	B	C	D	E	F	G
I165	X	X		X	X	X	X
K167				X			
F170	X		X	X	X	X	
V173	X	X	X	X	X	X	X
A186	X	X	X		X	X	X
K188	X	X	X	X	X	X	X
E203							X
V222	X	X	X	X	X	X	X
F238	X	X	X	X	X	X	X
E239		X			X	X	X
L241	X	X	X	X	X	X	X
N244				X			
N292					X		
L294	X	X	X	X	X	X	X
V306	X	X	X	X	X	X	X
D307	X	X	X	X	X	X	X

Table A5. CLK1 residues that interacted with meridianins (each represented by letters A–G) after MD simulations. Those residues involved in all meridianin binding are in bold and were considered important binding residues.

Residues	A	B	C	D	E	F	G
L167	X	X	X	X	X	X	X
F172	X	X	X		X	X	
V175	X	X	X	X	X	X	X
A189	X	X	X	X	X	X	X
K191	X	X	X	X	X		X
E206			X				X
L210			X				
V225	X	X	X	X		X	X
F241	X	X	X	X		X	X
E242			X	X	X	X	
L243	X		X	X			
L244	X	X	X	X	X	X	X
G245					X		
S247					X		
N293							X
L295	X	X	X	X	X	X	X
V324	X	X	X	X	X		X
D325	X	X	X	X	X		X

Table A6. Binder columns represent those residues identified after a bibliographic and database research and that interacted with other inhibitors. In shared columns are those residues involved with all meridianins binding per target. Residue number corresponds to each Protein Data Bank (PDB) number.

GSK3 β		CK1 δ		DYRK1A		CLK1	
Binders	Shared	Binders	Shared	Binders	Shared	Binders	Shared
I62	I62					L167	L167
V70	V70	I23	I23	V173	V173	V175	V175
A83	A83	A36	A36			A189	A189
K85	K85	K38	K38	K188	K188	K191	
				V222	V222	V225	
L132	L132	M82	M82	F238	F238	F241	
V135		L85	L85	L241	L241	L244	L244
L188		L135	L135	L294	L294	L295	L295
		I148	I148	V306	V306	V324	
D200	D200	D149		D307	D307		

Table A7. Residues involved in all meridianins binding to GSK3 β , CK1 δ , DYRK1A and CLK1. Residue number corresponds to each PDB number.

GSK3 β	CK1 δ	DYRK1A	CLK1
I62	I15	I165	L167
		K167	
F67		F170	F172
V70	I23	V173	V175
	Y24		
A83	A36	A186	A189
K85	K38	K188	K191
E97			
	E52	E203	E206
	Y56		L210
	M80		
V110		V222	V225
L132	M82	F238	F241
D133	E83	E239	E242
Y134	L84		L243
V135	L85	L241	L244
T138		N244	G245
Q185			S247
		N292	N293
L188	L135	L294	L295
	I148	V306	V324
D200	D149	D307	D325

Table A8. Summary of classical rigid docking of the best model selected per meridianin A–G and the derived compounds **1**, **2** and **3**, against others protein kinases and one non-kinase (IDH1).

GSK3β		CK1δ		DYRK1A		CLK1		FGFR1		PRKACA	
Binding Energy (kcal/mol)		Binding Energy (kcal/mol)		Binding Energy (kcal/mol)		Binding Energy (kcal/mol)		Binding Energy (kcal/mol)		Binding Energy (kcal/mol)	
R0/R1		R0/R1		R0/R1		R0/R1		R0/R1		R0/R1	
A	-7.3/-7.3	A	-6.9/-6.9	A	-7.4/-7.3	A	-8.9/-8.9	A	-7.1/-7.1	A	-8.5/-8.5
B	-7.3/-7.2	B	-6.4/-6.4	B	-7.7/-6.9	B	-8.5/-8.5	B	-6.7/-6.7	B	-8.5/-8.5
C	-7.6/-7.6	C	-6.9/-6.9	C	-8.2/-8.2	C	-8.5/-8.5	C	-7.1/-7.1	C	-9.0/-9.0
D	-7.7/-7.7	D	-7.0/-7.0	D	-7.9/-7.9	D	-8.6/-8.6	D	-7.1/-7.1	D	-8.8/-8.8
E	-7.3/-7.3	E	-7.0/-7.0	E	-7.5/-7.4	E	-9.0/-8.8	E	-7.4/-7.4	E	-7.6/-7.6
F	-7.9/-7.9	F	-7.2/-7.3	F	-8.0/-7.8	F	-8.7/-8.7	F	-7.3/-7.3	F	-8.5/-8.5
G	-7.3/-7.3	G	-6.8/-6.8	G	-8.1/-8.1	G	-9.1/-9.1	G	-7.1/-7.1	G	-8.6/-8.6
1	-7.6/-7.6	1	-7.0/-7.0	1	-8.1/-8.1	1	-9.2/-9.2	1	-7.1/-7.1	1	-8.2/-8.2
2	-7.7/-7.7	2	-7.1/-7.1	2	-8.2/-8.2	2	-7.4/-7.4	2	-6.2/-6.2	2	-8.3/-8.3
3	-8.0/-8.0	3	-7.3/-7.3	3	-7.9/-7.9	3	-7.8/-7.8	3	-6.4/-6.4	3	-8.0/-8.0
HK2		MAP2K1		PIK3CG		PRKG1		IDH1		PKN1	
Binding Energy (kcal/mol)		Binding Energy (kcal/mol)		Binding Energy (kcal/mol)		Binding Energy (kcal/mol)		Binding Energy (kcal/mol)		Binding Energy (kcal/mol)	
R0/R1		R0/R1		R0/R1		R0/R1		R0/R1		R0/R1	
A	-7.1/-7.1	A	-7.6/-7.6	A	-6.8/-6.8	A	-6.3/-6.3	A	-5.8/-5.8	A	-7.8/-7.8
B	-6.6/-6.6	B	-7.4/-7.4	B	-7.3/-7.3	B	-6.5/-6.5	B	-6.3/-6.3	B	-7.7/-7.7
C	-7.0/-7.0	C	-7.2/-7.2	C	-7.0/-7.0	C	-6.6/-6.6	C	-5.8/-5.8	C	-8.1/-8.1
D	-6.6/-6.6	D	-7.4/-7.4	D	-7.8/-7.8	D	-6.9/-6.9	D	-5.8/-5.8	D	-7.3/-7.3
E	-6.8/-6.8	E	-6.7/-6.7	E	-7.1/-7.1	E	-6.3/-6.3	E	-5.6/-5.6	E	-7.2/-7.2
F	-6.9/-6.9	F	-7.5/-7.5	F	-7.2/-7.2	F	-6.8/-6.8	F	-6.0/-6.0	F	-7.8/-7.8
G	-6.9/-6.9	G	-7.5/-7.5	G	-7.3/-7.3	G	-6.5/-6.5	G	-6.3/-6.3	G	-7.9/-7.9
1	-6.9/-6.9	1	-7.4/-7.4	1	-7.7/-7.7	1	-6.4/-6.4	1	-5.8/-5.8	1	-7.4/-7.4
2	-8.1/-8.1	2	-7.5/-7.5	2	-7.3/-7.3	2	-5.3/-5.3	2	-6.1/-6.1	2	-7.7/-7.7
3	-7.1/-7.1	3	-7.2/-7.2	3	-7.3/-7.3	3	-6.1/-6.1	3	-6.0/-6.0	3	-7.7/-7.7

All energies values are kcal/mol. To avoid false positives, each docking calculation was performed twice (R0 and R1).

Table A9. Summary of ADMET properties of meridianins (A to G) and the derived compounds extracted from the literature (1–3).

	LogS		Caco2 Permeability		Caco2* Permeability		Intestinal Absorption		Skin Permeability	
	ABSORPTION	A	-4.18	A	H	A	0.99	A	93.38%	A
B		-5.02	B	H	B	1.07	B	92.22%	B	-2.76
C		-5.55	C	H	C	0.95	C	91.77%	C	-2.92
D		-5.55	D	H	D	0.95	D	92.715	D	-2.91
E		-5.04	E	H	E	0.98	E	90.98%	E	-2.74
F		-6.16	F	H	F	0.98	F	91.49%	F	-2.92
G		-4.51	G	H	G	0.86	G	93.44%	G	-2.90
1		-4.18	1	OAD	1	0.93	1	91.41%	1	-2.90
2		-5.02	2	OAD	2	0.8	2	89.89%	2	-2.884
3		-5.55	3	H	3	0.819	3	91.04%	3	-2.895
P-Glycoprotein Substrate		P-Glycoprotein I/II Inhibitor		DISTRIBUTION	LogP		BBB		PPB	
A	Yes	A	No		A	1.53	A	No	A	>90%
B	Yes	B	No		B	2.39	B	No	B	>90%
C	Yes	C	No		C	3.10	C	No	C	>90%
D	Yes	D	No		D	3.10	D	No	D	>90%
E	Yes	E	No		E	2.40	E	No	E	>90%
F	Yes	F	No		F	3.58	F	No	F	>90%
G	Yes	G	No		G	2.44	G	No	G	<50%
1	Yes	1	No		1	3.40	1	No	1	>90%
2	Yes	2	No		2	3.40	2	No	2	>90%
3	Yes	3	No	3	3.10	3	No	3	>90%	
VDss		CNS Permeability		METABOLISM	CYP450 Metabolism *		Total Clearance		Renal OCT2 Substrate	
A	0.25	A	-2.92		A	Yes	A	0.57	A	No
B	0.24	B	-2.92		B	Yes	B	0.30	B	No
C	-0.06	C	-2.81		C	Yes	C	0.09	C	No
D	-0.01	D	-2.82		D	Yes	D	0.14	D	No
E	0.22	E	-2.93		E	Yes	E	0.15	E	No
F	0.07	F	-2.82		F	Yes	F	-0.19	F	No
G	-0.10	G	-2.12		G	Yes	G	0.71	G	No
1	-0.02	1	-2.83		1	Yes	1	-0.07	1	No
2	-0.09	2	-1.88		2	Yes	2	-0.092	2	No
3	-0.09	3	-1.99	3	Yes	3	0.132	3	No	

Caco2 permeability is calculated using proprietary ML model and Caco2* with the pkCSM webserver, as explained in the methods section. CCYP450 metabolism* specific values of interaction with different CYP450 isoforms are listed in Table A11. BBB: blood brain Barrier, PPB: protein-protein binding, VDss: steady state volume of distribution, CNS: central nervous system, OCT2: organic cation transported 2.

Table A10. Summary of toxicity properties of meridianins A–G and the three derived compounds extracted from the literatures (1–3).

	MRTD	AMES Toxicity		hERG I/II Inhibition		Hepatotoxicity		Skin Sensitisation	
A	0.503	A	No	A	No	A	No	A	No
B	0.584	B	No	B	No	B	Yes	B	No
C	-0.107	C	Yes	C	No	C	No	C	No
D	-0.095	D	Yes	D	No	D	No	D	No
E	0.589	E	No	E	No	E	No	E	No
F	-0.088	F	Yes	F	No	F	Yes	F	No
G	-0.086	G	Yes	G	No	G	No	G	No
1	-0.068	1	Yes	1	No	1	No	1	No
2	-0.038	2	Yes	2	No	2	No	2	No
3	-0.058	3	Yes	3	No	3	No	3	No

MRTD: maximum recommended tolerated dose, hERG: human ether-a-go-go gene.

Table A11. Summary of specific values of interaction with different CYP450 isoforms properties of meridianins A–G and the three derived compounds (1–3).

CYP2D6 Substrate		CYP3A4 Substrate		CYP1A2 Inhibitor		CYP2C19 Inhibitor		CYP2C9 Inhibitor		CYP2D6 Inhibitor		CYP3A4 Inhibitor	
A	No	A	No	A	Yes	A	No	A	No	A	No	A	No
B	No	B	No	B	Yes	B	No	B	No	B	No	B	No
C	No	C	No	C	Yes	C	No	C	No	C	No	C	No
D	No	D	No	D	Yes	D	No	D	No	D	No	D	No
E	No	E	No	E	Yes	E	No	E	No	E	No	E	No
F	No	F	No	F	Yes	F	Yes	F	No	F	No	F	No
G	No	G	No	G	Yes	G	No	G	No	G	No	G	No
1	No	1	No	1	Yes	1	Yes	1	No	1	No	1	Yes
2	No	2	Yes	2	Yes	2	Yes	2	No	2	No	2	No
3	No	3	No	3	Yes	3	No	3	No	3	No	3	Yes

CYP: Cytochrome.

References

- Paul, S.M.; Mytelka, D.S.; Dunwiddie, C.T.; Persinger, C.C.; Munos, B.H.; Lindborg, S.R.; Schacht, A.L. How to improve R&D productivity: The pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* **2010**, *9*, 203. [[PubMed](#)]
- Leelananda, S.P.; Lindert, S. Computational methods in drug discovery. *Beilstein J. Org. Chem.* **2016**, *12*, 2694–2718. [[CrossRef](#)] [[PubMed](#)]
- Ou-Yang, S.-S.; Lu, J.-Y.; Kong, X.-Q.; Liang, Z.-J.; Luo, C.; Jiang, H. Computational drug discovery. *Acta Pharmacol. Sin.* **2012**, *33*, 1131–1140. [[CrossRef](#)] [[PubMed](#)]
- Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E.W. Computational Methods in Drug Discovery. *Pharmacol. Rev.* **2014**, *66*, 334–395. [[CrossRef](#)] [[PubMed](#)]
- Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular docking: A powerful approach for structure-based drug discovery. *Curr. Comput. Aided Drug Des.* **2011**, *7*, 146–157. [[CrossRef](#)] [[PubMed](#)]
- Durrant, J.D.; McCammon, J.A. Molecular dynamics simulations and drug discovery. *BMC Biol.* **2011**, *9*, 71. [[CrossRef](#)] [[PubMed](#)]
- Yu, W.; Mackerell, A.D. Computer-Aided Drug Design Methods. *Methods Mol. Biol.* **2017**, *1520*, 93–94.
- Acharya, C.; Coop, A.; Polli, J.E.; Mackerell, A.D., Jr. Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr. Comput. Aided Drug Des.* **2011**, *7*, 10–22. [[CrossRef](#)] [[PubMed](#)]
- Li, J.W.-H.; Vederas, J.C. Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science* **2009**, *325*, 161–165. [[CrossRef](#)] [[PubMed](#)]
- Harvey, A.L.; Edrada-Ebel, R.; Quinn, R.J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* **2015**, *14*, 111–129. [[CrossRef](#)] [[PubMed](#)]
- Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **2016**, *79*, 629–661. [[CrossRef](#)] [[PubMed](#)]
- Prachayasittikul, V.; Worachartcheewan, A.; Shoombuatong, W.; Songtawee, N.; Simeon, S.; Prachayasittikul, V.; Nantasenamat, C. Computer-Aided Drug Design of Bioactive Natural Products. *Curr. Top. Med. Chem.* **2015**, *15*, 1780–1800. [[CrossRef](#)] [[PubMed](#)]
- Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on natural products for drug design. *Nat. Chem.* **2016**, *8*, 531–541. [[CrossRef](#)] [[PubMed](#)]
- Martin, L.; Latypova, X.; Wilson, C.M.; Magnaudeix, A.; Perrin, M.L.; Yardin, C.; Terro, F. Tau protein kinases: Involvement in Alzheimer's disease. *Ageing Res. Rev.* **2013**, *12*, 289–309. [[CrossRef](#)] [[PubMed](#)]
- Kolarova, M.; García-Sierra, F.; Bartos, A.; Riczny, J.; Ripova, D.; Ripova, D. Structure and Pathology of Tau Protein in Alzheimer Disease. *Int. J. Alzheimers Dis.* **2012**, *2012*, 731526. [[CrossRef](#)] [[PubMed](#)]
- Citron, M. Alzheimer's disease: strategies for disease modification. *Nat. Rev. Drug Discov.* **2010**, *9*, 387–398. [[CrossRef](#)] [[PubMed](#)]
- Tell, V.; Hilgeroth, A. Recent developments of protein kinase inhibitors as potential AD therapeutics. *Front. Cell. Neurosci.* **2013**, *7*, 189. [[CrossRef](#)] [[PubMed](#)]
- Dolan, P.J.; Johnson, G.V.W. The role of tau kinases in Alzheimer's disease. *Curr. Opin. Drug Discov. Dev.* **2010**, *13*, 595–603.

19. Lucke-Wold, B.P.; Turner, R.C.; Logsdon, A.F.; Simpkins, J.W.; Alkon, D.L.; Smith, K.E.; Chen, Y.-W.; Tan, Z.; Huber, J.D.; Rosen, C.L. Common mechanisms of Alzheimer's disease and ischemic stroke: The role of protein kinase C in the progression of age-related neurodegeneration. *J. Alzheimers Dis.* **2015**, *43*, 711–724. [PubMed]
20. Montaser, R.; Luesch, H. Marine natural products: A new wave of drugs? *Future Med. Chem.* **2011**, *3*, 1475–1489. [CrossRef] [PubMed]
21. Blunt, J.W.; Copp, B.R.; Keyzers, R.A.; Munro, M.H.G.; Prinsep, M.R. Marine natural products. *Nat. Prod. Rep.* **2017**, *34*, 235–294. [CrossRef] [PubMed]
22. Kiuru, P.; D'Auria, M.; Muller, C.; Tammela, P.; Vuorela, H.; Yli-Kauhaluoma, J. Exploring Marine Resources for Bioactive Compounds. *Planta Med.* **2014**, *80*, 1234–1246. [CrossRef] [PubMed]
23. Molinski, T.F.; Dalisay, D.S.; Lievens, S.L.; Saludes, J.P. Drug development from marine natural products. *Nat. Rev. Drug Discov.* **2009**, *8*, 69–85. [CrossRef] [PubMed]
24. The Nagoya Protocol on Access and Benefit-Sharing. Available online: <https://www.cbd.int/abs/about/default.shtml/> (accessed on 16 May 2017).
25. Bharate, S.B.; Yadav, R.R.; Battula, S.; Vishwakarma, R.A. Meridianins: Marine-Derived Potent Kinase Inhibitors. *Mini-Rev. Med. Chem.* **2012**, *12*, 618–631. [CrossRef] [PubMed]
26. Núñez-Pons, L.; Carbone, M.; Vázquez, J.; Rodríguez, J.; Nieto, R.M.; Varela, M.M.; Gavagnin, M.; Avila, C. Natural Products from Antarctic Colonial Ascidians of the Genera *Aplidium* and *Synoicum*: Variability and Defensive Role. *Mar. Drugs* **2012**, *10*, 1741–1764. [CrossRef] [PubMed]
27. Traxler, P.; Furet, P. Strategies toward the Design of Novel and Selective Protein Tyrosine Kinase Inhibitors. *Pharmacol. Ther.* **1999**, *82*, 195–206. [CrossRef]
28. Huang, D.; Zhou, T.; Lafleur, K.; Nevado, C.; Cafilisch, A. Kinase selectivity potential for inhibitors targeting the ATP binding site: A network analysis. *Bioinformatics* **2010**, *26*, 198–204. [CrossRef] [PubMed]
29. McGregor, M.J. A Pharmacophore Map of Small Molecule Protein Kinase Inhibitors. *J. Chem. Inf. Model.* **2007**, *47*, 2374–2382. [CrossRef] [PubMed]
30. Ebrahim, H.; El Sayed, K. Discovery of Novel Antiangiogenic Marine Natural Product Scaffolds. *Mar. Drugs* **2016**, *14*, 57. [CrossRef] [PubMed]
31. Feng, L.; Geisselbrecht, Y.; Blanck, S.; Wilbuer, A.; Atilla-Gokcumen, G.E.; Filippakopoulos, P.; Kräling, K.; Celik, M.A.; Harms, K.; Maksimoska, J.; et al. Structurally sophisticated octahedral metal complexes as highly selective protein kinase inhibitors. *J. Am. Chem. Soc.* **2011**, *133*, 5976–5986. [CrossRef] [PubMed]
32. Giraud, F.; Alves, G.; Debiton, E.; Nauton, L.; Théry, V.; Durieu, E.; Ferandin, Y.; Lozach, O.; Meijer, L.; Anizon, F.; et al. Synthesis, Protein Kinase Inhibitory Potencies, and in Vitro Antiproliferative Activities of Meridianin Derivatives. *J. Med. Chem.* **2011**, *54*, 4474–4489. [CrossRef] [PubMed]
33. Wang, G.; Zhu, W. Molecular docking for drug discovery and development: A widely used approach but far from perfect. *Future Med. Chem.* **2016**, *8*, 1707–1710. [CrossRef] [PubMed]
34. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. The Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* **2016**, *59*, 4035–4061. [CrossRef] [PubMed]
35. De Vivo, M.; Cavalli, A. Recent advances in dynamic docking for drug discovery. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1320. [CrossRef]
36. Aravindhan, G.; Coote, M.L.; Barakat, K. Molecular dynamics-driven drug discovery: Leaping forward with confidence. *Drug Discov. Today* **2017**, *22*, 249–269.
37. Pagadala, N.S.; Syed, K.; Tuszynski, J. Software for molecular docking: A review. *Biophys. Rev.* **2017**, *9*, 91–102. [CrossRef] [PubMed]
38. Shin, W.-H.; Zhu, X.; Bures, M.; Kihara, D. Three-Dimensional Compound Comparison Methods and Their Application in Drug Discovery. *Molecules* **2015**, *20*, 12841–12862. [CrossRef] [PubMed]
39. Muegge, I.; Mukherjee, P. An overview of molecular fingerprint similarity search in virtual screening. *Expert Opin. Drug Discov.* **2016**, *11*, 137–148. [CrossRef] [PubMed]
40. Alonso, N.; Caamaño, O.; Romero-Duran, F.J.; Luan, F.; Cordeiro, M.N.D.S.; Yañez, M.; González-Díaz, H.; García-Mera, X. Model for High-Throughput Screening of Multitarget Drugs in Chemical Neurosciences: Synthesis, Assay, and Theoretic Study of Rasagiline Carbamates. *ACS Chem. Neurosci.* **2013**, *4*, 1393–1403. [CrossRef] [PubMed]

41. Simpraga, S.; Alvarez-Jimenez, R.; Mansvelter, H.D.; Van Gerven, J.M.A.; Groeneveld, G.J.; Poil, S.-S.; Linkenkaer-Hansen, K. EEG machine learning for accurate detection of cholinergic intervention and Alzheimer's disease. *Sci. Rep.* **2017**, *7*, 5775. [CrossRef] [PubMed]
42. Cummings, J.; Aisen, P.S.; DuBois, B.; Frölich, L.; Jack, C.R.; Jones, R.W.; Morris, J.C.; Raskin, J.; Dowsett, S.A.; Scheltens, P. Drug development in Alzheimer's disease: The path to 2025. *Alzheimers Res. Ther.* **2016**, *8*, 39. [CrossRef] [PubMed]
43. Solomon, K.A.; Sundararajan, S.; Abirami, V. QSAR Studies on *N*-aryl Derivative Activity Towards Alzheimer's Disease. *Molecules* **2009**, *14*, 1448–1455. [CrossRef] [PubMed]
44. Gopi Mohan, C.; Gupta, S. QSAR Models towards Cholinesterase Inhibitors for the Treatment of Alzheimer's Disease. In *Oncology: Breakthroughs in Research and Practice*; IGI Global: Hershey, PA, USA, 2016.
45. PharmaSea. Available online: <http://www.pharma-sea.eu/> (accessed on 2 September 2017).
46. MaCuMBA. Available online: <https://www.macumbaproject.eu/> (accessed on 2 September 2017).
47. SeaBiotech. Available online: <http://spider.science.strath.ac.uk/seabiotech/index.php> (accessed on 2 September 2017).
48. BlueGenics. Available online: <http://www.bluegenics.eu/cms/> (accessed on 2 September 2017).
49. MicroB3. Available online: <https://www.microb3.eu/> (accessed on 2 September 2017).
50. Jain, P.; Karthikeyan, C.; Moorthy, N.S.; Waiker, D.; Jain, A.; Trivedi, P. Human CDC2-Like Kinase 1 (CLK1): A Novel Target for Alzheimer's Disease. *Curr. Drug Targets* **2014**, *15*, 539–550. [CrossRef] [PubMed]
51. Yadav, R.R.; Sharma, S.; Joshi, P.; Wani, A.; Vishwakarma, R.A.; Kumar, A.; Bharate, S.B. Meridianin derivatives as potent Dyrk1A inhibitors and neuroprotective agents. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 2948–2952. [CrossRef] [PubMed]
52. Tahtouh, T.; Elkins, J.M.; Filippakopoulos, P.; Soundararajan, M.; Burgy, G.; Durieu, E.; Cochet, C.; Schmid, R.S.; Lo, D.C.; Delhommel, F.; et al. Selectivity, Cocrystal Structures, and Neuroprotective Properties of Leucettines, a Family of Protein Kinase Inhibitors Derived from the Marine Sponge Alkaloid Leucettamine B. *J. Med. Chem.* **2012**, *55*, 9312–9330. [CrossRef] [PubMed]
53. Wang, W.; Bodles-Brakhop, A.M.; Barger, S.W. A Role for P-Glycoprotein in Clearance of Alzheimer Amyloid β -Peptide from the Brain. *Curr. Alzheimer Res.* **2016**, *13*, 615–620. [CrossRef] [PubMed]
54. Cirrito, J.R.; Deane, R.; Fagan, A.M.; Spinner, M.L.; Parsadanian, M.; Finn, M.B.; Jiang, H.; Prior, J.L.; Sagare, A.; Bales, K.R.; et al. P-glycoprotein deficiency at the blood-brain barrier increases amyloid- β deposition in an Alzheimer disease mouse model. *J. Clin. Investig.* **2005**, *115*, 3285–3290. [CrossRef] [PubMed]
55. Miller, D.S.; Bauer, B.; Hartz, A.M.S. Modulation of P-glycoprotein at the blood-brain barrier: Opportunities to improve central nervous system pharmacotherapy. *Pharmacol. Rev.* **2008**, *60*, 196–209. [CrossRef] [PubMed]
56. Chang, K.L.; Pee, H.N.; Yang, S.; Ho, P.C. Influence of drug transporters and stereoselectivity on the brain penetration of pioglitazone as a potential medicine against Alzheimer's disease. *Sci. Rep.* **2015**, *5*, 9000. [CrossRef] [PubMed]
57. Lynch, T.; Price, A. The Effect of Cytochrome P450 Metabolism on Drug Response, Interactions, and Adverse Effects. *Am. Fam. Physician* **2007**, *76*, 391–396. [PubMed]
58. Santamaria-Navarro, E.; Felix, E.; Nonell-Canals, A. Cabrakan. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
59. Santamaria-Navarro, E.; Nonell-Canals, A. Hurakan. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
60. Bento, A.P.; Gaulton, A.; Hersey, A.; Bellis, L.J.; Chambers, J.; Davies, M.; Krüger, F.A.; Light, Y.; Mak, L.; McGlinchey, S.; et al. The ChEMBL bioactivity database: An update. *Nucleic Acids Res.* **2014**, *42*, 1083–1090. [CrossRef] [PubMed]
61. Felix, E.; Santamaria-Navarro, E.; Sanchez-Martinez, M.; Nonell-Canals, A. Ixchel. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
62. Berman, H.; Henrick, K.; Nakamura, H.; Markley, J.L. The worldwide Protein Data Bank (wwPDB): Ensuring a single, uniform archive of PDB data. *Nucleic Acids Res.* **2007**, *35*, 2006–2008. [CrossRef] [PubMed]
63. The UniProt Consortium Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res.* **2013**, *41*, 43–47.

64. Mente, S.; Arnold, E.; Butler, T.; Chakrapani, S.; Chandrasekaran, R.; Cherry, K.; Dirico, K.; Doran, A.; Fisher, K.; Galatsis, P.; et al. Ligand-protein interactions of selective casein kinase 1 δ inhibitors. *J. Med. Chem.* **2013**, *56*, 6819–6828. [CrossRef] [PubMed]
65. Lill, M.A. Efficient Incorporation of Protein Flexibility and Dynamics into Molecular Docking Simulations. *Biochemistry* **2011**, *50*, 6157–6169. [CrossRef] [PubMed]
66. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, *441*, 1–13. [CrossRef] [PubMed]
67. Alonso, H.; Bliznyuk, A.A.; Gready, J.E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* **2006**, *26*, 531–568. [CrossRef] [PubMed]
68. Felix, E.; Santamaría-Navarro, E.; Sanchez-Martinez, M.; Nonell-Canals, A. Itzamna. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
69. Felix, E.; Nonell-Canals, A. Kin. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
70. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [CrossRef] [PubMed]
71. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950–1958. [CrossRef] [PubMed]
72. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [CrossRef] [PubMed]
73. Martín-García, F.; Papaleo, E.; Gomez-Puertas, P.; Boomsma, W.; Lindorff-Larsen, K. Comparing Molecular Dynamics Force Fields in the Essential Subspace. *PLoS ONE* **2015**, *10*, e0121114. [CrossRef] [PubMed]
74. Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M.P.; Dror, R.O.; Shaw, D.E. Systematic validation of protein force fields against experimental data. *PLoS ONE* **2012**, *7*, e32131. [CrossRef] [PubMed]
75. Wang, J.; Wang, W.; Kollman, P.A.; Case, D.A. Antechamber, An Accessory Software Package For Molecular Mechanical Calculations. *J. Comput. Chem.* **2005**, *25*, 1157–1174. [CrossRef] [PubMed]
76. Case, D.A.; Cheatham, T.E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688. [CrossRef] [PubMed]
77. Jorgensen, W.L.; Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *J. Comput. Chem.* **1998**, *19*, 1179–1186. [CrossRef]
78. Andersen, H.C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983**, *52*, 24–34. [CrossRef]
79. Rastelli, G.; Degliesposti, G.; Del Rio, A.; Sgobba, M. Binding Estimation after Refinement, a New Automated Procedure for the Refinement and Rescoring of Docked Ligands in Virtual Screening. *Chem. Biol. Drug Des.* **2009**, *73*, 283–286. [CrossRef] [PubMed]
80. Miller, B.R., III; McGee, T.D., Jr.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA. py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321. [CrossRef] [PubMed]
81. Halekotte, J.; Witt, L.; Ianes, C.; Krüger, M.; Bührmann, M.; Rauh, D.; Pichlo, C.; Brunstein, E.; Luxenburger, A.; Baumann, U.; et al. Optimized 4,5-Diarylimidazoles as Potent/Selective Inhibitors of Protein Kinase CK1 δ and Their Structural Relation to p38 α MAPK. *Molecules* **2017**, *22*, 522. [CrossRef] [PubMed]
82. Fedorov, O.; Huber, K.; Eisenreich, A.; Filippakopoulos, P.; King, O.; Bullock, A.N.; Szklarczyk, D.; Jensen, L.J.; Fabbro, D.; Trappe, J.; et al. Specific CLK Inhibitors from a Novel Chemotype for Regulation of Alternative Splicing. *Chem. Biol.* **2011**, *18*, 67–76. [CrossRef] [PubMed]
83. Vidal, D.; Nonell-Canals, A. ADMET Models. Available online: <https://www.mindthebyte.com/> (accessed on 3 May 2017).
84. Pires, D.E.V.; Blundell, T.L.; Ascher, D.B. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J. Med. Chem.* **2015**, *58*, 4066–4072. [CrossRef] [PubMed]
85. Muehlbacher, M.; Spitzer, G.M.; Liedl, K.R.; Kornhuber, J. Qualitative prediction of blood–brain barrier permeability on a large and refined dataset. *J. Comput. Aided Mol. Des.* **2011**, *25*, 1095–1106. [CrossRef] [PubMed]

86. Zhu, X.-W.; Sedykh, A.; Zhu, H.; Liu, S.-S.; Tropsha, A. The Use of Pseudo-Equilibrium Constant Affords Improved QSAR Models of Human Plasma Protein Binding. *Pharm. Res.* **2013**, *30*, 1790–1798. [[CrossRef](#)] [[PubMed](#)]
87. Huuskonen, J. Estimation of Aqueous Solubility for a Diverse Set of Organic Compounds Based on Molecular Topology. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 773–777. [[CrossRef](#)] [[PubMed](#)]
88. Yuan, S.; Chan, H.C.S.; Hu, Z. Using PyMOL as a platform for computational drug design. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1298. [[CrossRef](#)]
89. Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: Fully automated protein–ligand interaction profiler. *Nucleic Acids Res.* **2015**, *43*, W443–W447. [[CrossRef](#)] [[PubMed](#)]
90. Prudhomme, M.; Rossignol, E.; Youssef, A.; Anizon, F.; Moreau, P.; Fabbro, D.; Cohen, P. Aminopyrimidylindoles structurally related to meridianins as kinase inhibitors. *Cancer Res.* **2014**, *68*, 1293.
91. Kingwell, K. Drug delivery: New targets for drug delivery across the BBB. *Nat. Rev. Drug Discov.* **2016**, *15*, 84–85. [[CrossRef](#)] [[PubMed](#)]
92. Banks, W.A. From blood–brain barrier to blood–brain interface: New opportunities for CNS drug delivery. *Nat. Rev. Drug Discov.* **2016**, *15*, 275–292. [[CrossRef](#)] [[PubMed](#)]
93. Saraiva, C.; Praça, C.; Ferreira, R.; Santos, T.; Ferreira, L.; Bernardino, L. Nanoparticle-mediated brain drug delivery: Overcoming blood–brain barrier to treat neurodegenerative diseases. *J. Control. Release* **2016**, *235*, 34–47. [[CrossRef](#)] [[PubMed](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Article

Kororamides, Convolutamines, and Indole Derivatives as Possible Tau and Dual-Specificity Kinase Inhibitors for Alzheimer's Disease: A Computational Study

Laura Llorach-Pares ^{1,2} , Alfons Nonell-Canals ², Conxita Avila ^{1,*} 
and Melchor Sanchez-Martinez ^{2,*} 

¹ Department of Evolutionary Biology, Ecology and Environmental Sciences, Faculty of Biology and Biodiversity Research Institute (IRBio), Universitat de Barcelona, 08028 Barcelona, Catalonia, Spain; laura@mindthebyte.com

² Mind the Byte S.L., 08007 Barcelona, Catalonia, Spain; alfons@mindthebyte.com

* Correspondence: conxita.avila@ub.edu (C.A.); melchor@mindthebyte.com (M.S.-M.); Tel.: +34-934020161 (C.A.); +34-936551112 (M.S.-M.)

Received: 30 August 2018; Accepted: 11 October 2018; Published: 16 October 2018



Abstract: Alzheimer's disease (AD) is becoming one of the most disturbing health and socioeconomic problems nowadays, as it is a neurodegenerative pathology with no treatment, which is expected to grow further due to population ageing. Actual treatments for AD produce only a modest amelioration of symptoms, although there is a constant ongoing research of new therapeutic strategies oriented to improve the amelioration of the symptoms, and even to completely cure the disease. A principal feature of AD is the presence of neurofibrillary tangles (NFT) induced by the aberrant phosphorylation of the microtubule-associated protein tau in the brains of affected individuals. Glycogen synthetase kinase-3 beta (GSK3 β), casein kinase 1 delta (CK1 δ), dual-specificity tyrosine phosphorylation regulated kinase 1A (DYRK1A) and dual-specificity kinase cdc2-like kinase 1 (CLK1) have been identified as the principal proteins involved in this process. Due to this, the inhibition of these kinases has been proposed as a plausible therapeutic strategy to fight AD. In this study, we tested in silico the inhibitory activity of different marine natural compounds, as well as newly-designed molecules from some of them, over the mentioned protein kinases, finding some new possible inhibitors with potential therapeutic application.

Keywords: meridianins; kororamide A–B; convolutamine I–J; indole scaffold; computer-aided drug design; Alzheimer's disease; GSK3 β ; CK1 δ ; DYRK1A; CLK1

1. Introduction

Constituting about 2% of all human genes, protein kinases are an important family of enzymes with a critical role in signal transduction pathway by modification of substrate activity. They are also responsible to control different aspects of cell functions by its phosphorylation activity, which plays a critical role in intracellular communication during development, and in the function of the nervous and immune systems [1]. Due to that, kinases are related with many diseases, such as Alzheimer's Disease (AD) or Amyotrophic Lateral Sclerosis (ALS), among others. AD, the neurodegenerative pathology that is considered to represent the most common type of dementia (60–80% of the total cases), is characterized by memory deterioration and modification of cognitive abilities. Alzheimer's pathologies are associated with the presence of senile plaques (SP), mainly composed by beta-amyloid (A β) peptides, and neurofibrillary tangles (NFT), that are intraneuronal aggregations principally

composed of abnormal phosphorylated tau protein. Tau is a soluble microtubule-binding protein and is hyperphosphorylated in AD. Tau phosphorylation is regulated by a balance between tau kinase and phosphatase activities. Anti-phosphorylation strategies (kinase inhibitors) aim to inhibit these processes of aggregation and the formation of NFT [2–4]. The abovementioned evidence may suggest that one of the key strategies is to prevent tau phosphorylation and, thus, combat AD, could be the inhibition of the protein kinases involved in the tau phosphorylation pathway [4].

Despite the catalytic subunits of many protein kinases are highly conserved, there are several differences between them that allow to classify protein kinases into subfamilies: (1) Protein kinases (EC 2.7.10); (2) serine-threonine protein kinases (EC 2.7.11); (3) dual-specificity kinases (those acting on Ser/Thr and Tyr residues) (EC 2.7.12); (4) protein-histidine kinases (EC 2.7.13); (5) protein-arginine kinases (EC 2.7.11.14); and (6) other protein kinases (EC 2.7.99), that can be also divided into sub-subfamilies, such as tau protein kinase (EC 2.7.11.26) and dual-specificity kinase (EC 2.7.12.1). The main relevant protein kinases involved in tau phosphorylation belong to the sub-subfamilies tau protein kinase and dual-specificity kinases. As tau protein kinases we find glycogen synthetase kinase-3 beta (GSK3 β) and casein kinase 1 delta (CK1 δ), while within dual-specificity kinases, we find dual-specificity tyrosine phosphorylation regulated kinase 1A (DYRK1A) and cdc2-like kinase 1 (CLK1). Each of them has different roles regarding AD pathology. For GSK3 β several authors suggest its link between A β and tau pathology, and in AD patients it has been co-localized with NFT. GSK3 β is suggested to phosphorylate and hyper-phosphorylate tau, while increasing the production of A β and mediating neuronal death. Phosphorylation of tau by GSK3 β occurs at 42 sites, where 29 of them are phosphorylated in AD brains. CK1 δ is part of the non-proline-directed protein kinase (non-PDPK) group inside the tau kinases and its levels are increased while is co-localized with NFT. CK1 δ has an important role on protein aggregation and regulates the microtubule dynamics through tau phosphorylation at 46 sites, 25 of them phosphorylated in AD brains. DYRK1A phosphorylates the amyloid precursor protein (APP) and tau proteins, thus increasing neuronal death and the formation of aggregates. DYRK1A induces tau phosphorylation at serine 202, threonine 212, and serine 404, sites that were found phosphorylated in AD brains. Finally, cdc2-like kinase 1 (CLK1), one of the four isoforms conforming an evolutionary conserved group of dual-specificity kinases, is related with AD by phosphorylating the serine residues in arginine-rich (SR) proteins [2,3,5–15].

The natural-product-inspired design plays an important role in chemical science, as historically natural products (NP) from diverse sources, such as plants or microbes, have been a rich source of compounds [16–18]. NP are optimized biologically active metabolites which can be used as a template to design drug-like compounds [16–18]. Evaluation of Food and Drug Administration (FDA) approved new molecular entities (NMEs) reveals that NP and their derived compounds represent over one-third of all NMEs [19], a percentage that is even higher regarding the active compounds in the central nervous system (CNS) domain [20]. AD is not an exception, and several drug candidates have been developed from natural sources against the different therapeutic targets identified to date [21–23]. In fact, few reasonable selective and potent GSK3 β , CK1 δ , DYRK1A, and CLK1 inhibitors have been described so far, most of them being marine natural products or derived molecules from them [5,24–36].

Recently, it has been shown that meridianins, indole alkaloids from the marine tunicate *Aplidium* from the Southern Ocean, could act as inhibitors of these four kinases, with possible inhibitors being derived from them [24,29,34,37]. In addition to that, kororamide A–B, two brominated alkaloids from the bryozoan *Amathia tortuosa* from Australia, showed a phenotypic signature on Parkinson's disease [38]. Their structure resembles that of meridianins and because of that we decided to study whether these compounds could also act as inhibitors of the four mentioned kinases, although, as far as we know, this relation has never been established before. Following with this, and having into account that marine indole alkaloid conform a large group of compounds with diverse biological activities that make them attractive starting points for pharmaceutical development [39–41], we have designed in this work several compounds starting from this well-known scaffold as a core element. Further, we modified the structural features observed in meridianins and kororamides, as well as with

the presence of halogen substituents (present also in both chemical species), which has been revealed as key player to increase activity over these four kinases [24,37,42,43].

To strengthen our initial assumption, we tested the indole scaffold and halogen substituents' effect on the inhibition of GSK3 β , CK1 δ , DYRK1A, and CLK1. To determine the importance of the indole scaffold for the inhibition of the four studied kinases we also screened the MarinLit database [44] to find other possible marine compounds that were similar to meridianin F and kororamide A (which were the best theoretical inhibitors of the four kinases), or at least incorporate the indole scaffold. Thereafter, we analysed their binding behaviour against them. Moreover, and because of the importance of the halogen substituents, we decided to investigate whether the halogen substituents are important with respect to the indole scaffold. To do that, we evaluated the inhibitory behaviour of convolutamine I–J, two halogenated heterocyclic compounds (that do not present an indole scaffold) extracted from the bryozoan *Amathia tortuosa*, and which are structurally and functionally related to kororamide A–B [38].

To sum up, with the general objective to help in the discovery of anti-AD drugs (protein inhibitor/s to reduce or alleviate AD symptoms), the concrete aim of this study is three-fold: (1) to validate if kororamide A–B and convolutamine I–J could act as novel inhibitors of the four studied kinases; (2) to test the indole scaffold importance on the kinases inhibition; and (3) to design new possible inhibitors of the four kinases starting from meridianin and kororamide indole scaffolds. To do so, a computational study targeting the adenosine triphosphate (ATP)-binding site of the aforementioned kinases has been carried out. Computer-aided drug design (CADD) techniques are widely used in (marine natural product) drug discovery, as they constitute an appropriate tool to rational design and developing new drug candidates, reducing the time and costs derived from their identification, characterization, and structure-optimization [45].

2. Results and Discussion

2.1. New Possible GSK3 β , CK1 δ , DYRK1A, and CLK1 ATP-Competitive Inhibitors

It is generally accepted that the ATP binding site of protein kinases, despite the fact that their catalytic domains are highly conserved, still remain the most used cavity in (rational) drug design over this family of proteins [46]. Protein kinases have two different lobes, the N-lobe that is mainly formed by β -sheets and the C-lobe formed by a helical structure. Between both lobes can be found the catalytic ATP cavity, which can be divided into five regions: glycine-rich region (GRR), hydrophobic pocket (HP), adenine region (AR), sugar pocket (SP), and the phosphate binding pocket (PBP) [46–48]. GRR and HP are located at the N-terminal lobe, while SP and PBP are placed at the C-terminal lobe. AR is in the middle of these regions, providing a link between them (see Figure 1).

All five regions are quite evolutionarily conserved between the kinases, but they are not identical [37]. GRG is a highly conserved region with a GxGxFG motif (Table 1). The same occurs with the HP, as all the four kinases have a VAIK motif, except DYRK1A with a Valine (V) residue instead of an Isoleucine (I). On the contrary, the AR does not seem to have any conserved motif, while SP can be identified by the PxNxL pattern. For the PBP, only the last aspartate residue (D) seems to be conserved along the four kinases.

As explained previously, the kinase ATP binding site is the most exploited cavity as far as inhibition is concerned. Several inhibitors have been reported in the past, some of them being marine natural products, such as meridianins [28,49]. Most of them can bind to all of these regions, with a different binding strength depending on their chemical structure. Interestingly, a common feature seems to be shared between the majority of them: the presence of an indole scaffold [5,25,26,30,31,33–35].

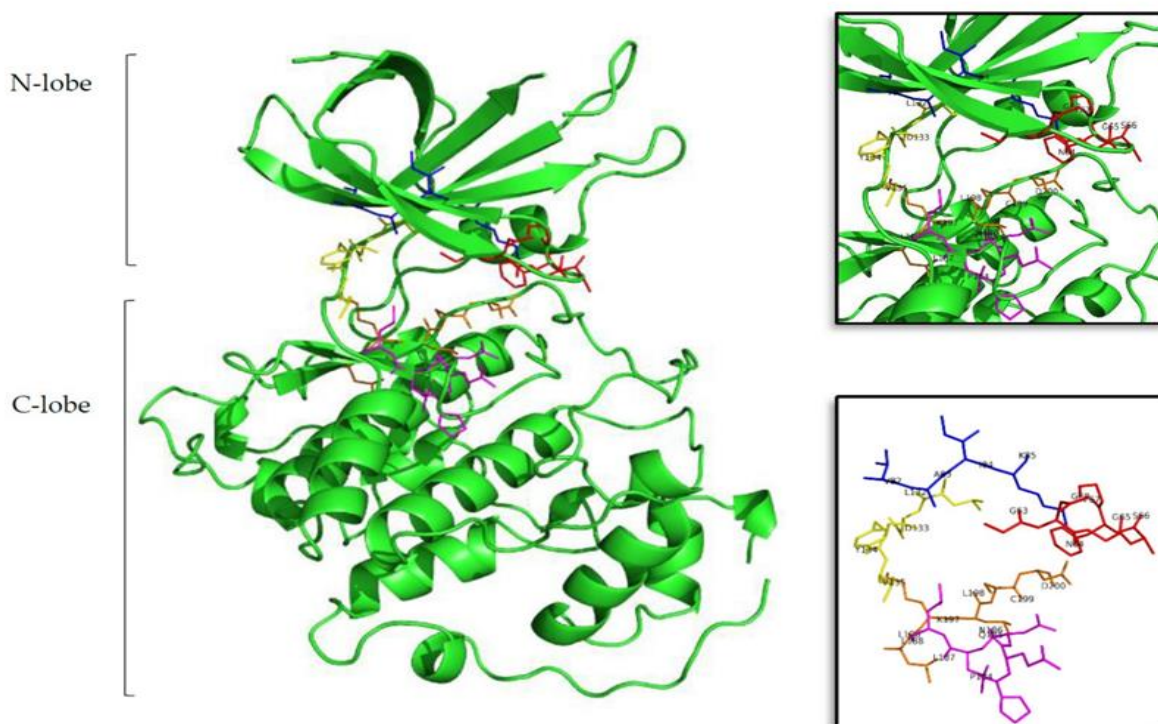


Figure 1. Structure of the tau protein kinase GSK3 β (Protein Data Bank ID (PDB) 3PUP). In the first, largest, image the two lobes can be seen in cartoon representation, and in sticks the residues that form the ATP cavity. In the top and bottom zoom images all the amino acid residues involved on the ATP binding site are shown. Residues in red represent the glycine-rich region (GRR), in blue the hydrophobic pocket (HP), in yellow the adenine region (AR), in lilac the sugar pocket (SP), and finally, in orange the phosphate binding pocket (PBP). Letters and numbers correspond to their position in the amino acid sequence and the PDB file numbering.

Table 1. Summary of the ATP binding site regions of GSK3 β , CK1 δ , DYRK1A, and CLK1. Five regions are found inside the ATP cavity and their respective residues are shown in a single letter code, as well as their sequence position that corresponds to each PDB file numbering.

	Glycine-Rich Region	Hydrophobic Pocket	Adenine Region	Sugar Pocket	Phosphate Binding Pocket
GSK3β	GNGSFG 63-68	VAIK 82-85	LDYV 132-135	PQNLL 184-188	LKLCD 196-200
CK1δ	GSGSFG 16-21	VAIK 35-38	MELL 82-85	PDNFL 131-135	VYIID 145-149
DYRK1A	GKGSFG 166-171	VAIK 184-187	FEML 238-241	PENIL 290-294	IKIVD 303-307
CLK1	GEGAFG 168-173	VAVK 188-191	FELL 241-244	PENIL 291-295	IKVVD 312-325

2.2. Kororamide A–B and Convolutamine I–J as Possible Kinase Inhibitors

Indole alkaloids are marine natural products that show specific biological activities, such as anti-inflammatory and serotonin antagonism [41]. Moreover, the therapeutic importance of this kind of indole scaffolds is well known, as demonstrated by clinical and preclinical studies showing pharmacological activities over neurodegenerative diseases, such as AD [41,50]. Within the group of compounds containing the indole moiety are meridianins, for instance. These molecules constitute a group of indole alkaloids consisting of an indole framework linked to an aminopyrimidine ring with a reported inhibitory activity over GSK3 β , CK1 δ , DYRK1A, and CLK1 [30,34,37]. Within the

list of indole-containing compounds, structurally similar to meridianins, different molecules can be found, among which are kororamides. Kororamide A and B are two tribrominated indole alkaloid compounds from the Southern Ocean bryozoan *Amanthia tortuosa*. These two marine molecules share a common halogenated indole scaffold with meridianins and, based on their chemical structural similarity, one could assume that kororamides could have an inhibitory activity similar to meridianins. In the same study where kororamide B was identified, three other compounds were also isolated, kororamide A and convolutamine I and J. The last two compounds do not present an indole scaffold, but they are halogenated heterocyclic compounds as other known kinase inhibitors [51–53] (Figure 2). To test this hypothesis, docking calculations and Molecular Dynamics (MD) simulations were carried out to evaluate if kororamide A–B and convolutamine I–J could behave as meridianins regarding GSK3 β , CK1 δ , DYRK1A, and CLK1 binding, thus indicating that they could be potential anti-AD therapeutic agents.

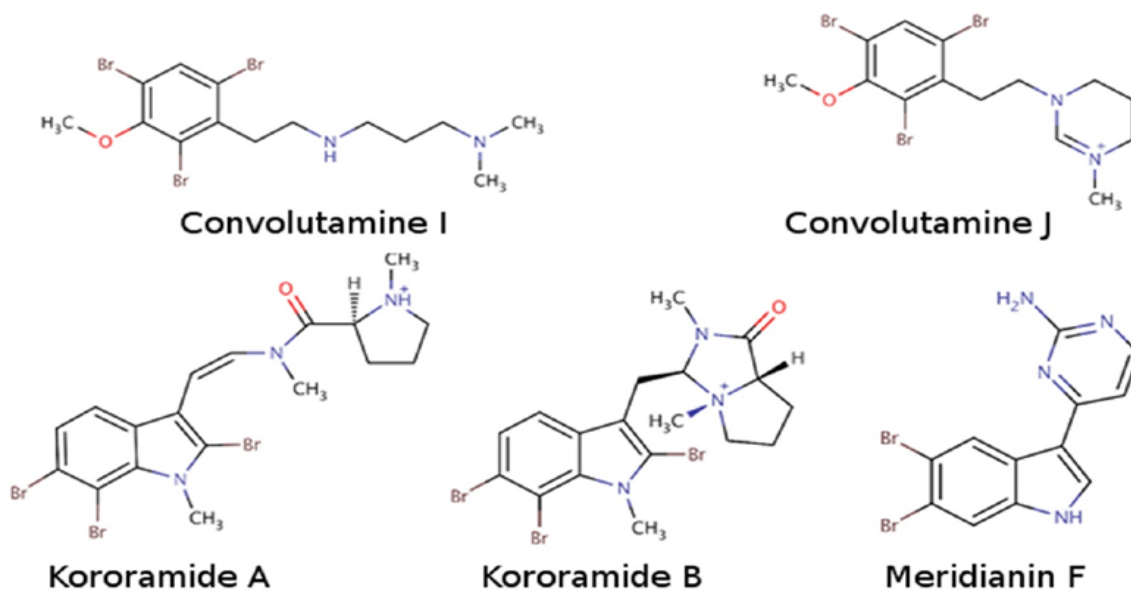


Figure 2. Chemical structures of convolutamine I, convolutamine J, kororamide A, kororamide B, and meridianin F.

In previous studies the presence of halogen atoms was considered important to achieve a good inhibitory activity over the four studied kinases [24,37]. In order to test whether the presence of a halogenated indole scaffold, or just the presence of aromatic cycle substituted with halogen atoms, enhances a higher binding affinity against GSK3 β , CK1 δ , DYRK1A, and CLK1, we analyse it by means of docking and MD simulations. Thereafter, we compared the obtained results (Table 2) with the values from kororamide A–B and convolutamine I–J with meridianin F, the most promising compound of the chemical family (meridianin A–F) [37].

Our results indicate that all the analysed compounds could bind to the ATP binding pocket of each of the mentioned kinases, thus theoretically acting as ATP competitive inhibitors (Figure 3). Binding energies obtained after docking and MD simulations (Table 2) show that convolutamine J and kororamide A tend to have higher energies than convolutamine I and kororamide B. To be more precise, kororamide A shows better energies when bound against GSK3 β , DYRK1A, and CLK1, while convolutamine J shows better energies over CK1 δ . Comparing the energies obtained between the four tribrominated metabolites found on the bryozoa *Amanthia tortuosa* and meridianin F, we observe that the last one has slightly better energies in all cases after MD. These energies do not allow us to discard any of the compounds as an ATP competitive inhibitor, and we can prioritize kororamide A and

convolutamine J over kororamide B, and especially over convolutamine I. Additionally, these results do not allow us to discriminate between which structural features influence most of the binding strength against the four studied kinases: the indole scaffold, the presence of halogen atoms, or the combination of both features.

Table 2. Summary of classical rigid docking and molecular mechanics/generalized born surface area (MM/GBSA) calculations of the two best models selected per meridianin F (F), convolutamine I (I) and J (J), and kororamide A (A) and B (B). To avoid false positives, each docking calculation was performed twice (R0/R1). All energies values are kcal/mol. For each target the first (left) column refers to the results of docking calculations while the second (right) column indicate the binding energy results obtained after MD calculations.

GSK3 β		CK1 δ		DYRK1A		CLK1					
Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)	Binding Energy (kcal/mol)				
R0/R1		R0/R1		R0/R1		R0/R1					
F	-7.9/-7.9	-35.18	F	-7.2/-7.3	-38.55	F	-8.0/-7.8	-39.99	F	-8.7/-8.7	-37.71
	-7.7/-7.9	-34.73		-7.1/-7.1	-38.93		-7.8/-7.7	-39.91		-8.5/-8.5	-37.61
I	-5.6/-5.6	-23.08	I	-5.0/-5.0	-3.19	I	-5.6/-5.6	-26.52	I	-5.8/-5.5	-33.23
	-6.3/-6.3	-18.38		-5.4/-5.4	-11.26		-4.8/-4.8	-11.02		-5.8/-5.8	-31.93
J	-6.7/-6.7	-31.58	J	-6.2/-6.2	-37.76	J	-7.4/-7.4	-31.35	J	-6.0/-6.0	-21.47
	-5.9/-5.9	-31.61		-5.8/-5.8	-28.91		-7.0/-7.0	-32.27		-4.6/-4.6	-24.37
A	-8.3/-8.3	-34.88	A	-8.0/-8.0	-35.48	A	-8.2/-8.2	-32.94	A	-6.7/-6.7	-37.46
	-8.1/-8.1	-31.02		-7.4/-7.4	-33.94		-6.7/-6.7	-14.61		-2.9/-2.9	-38.93
B	-9.1/-9.1	-31.80	B	-8.1/-8.1	-28.68	B	-7.7/-7.7	-23.83	B	-4.4/-4.4	-28.71
	-8.3/-8.3	-32.34		-6.6/-6.6	-35.53		-7.3/-7.3	-24.29		-4.0/-4.0	-22.96

With the aim of performing a deeper analysis of the inhibitory behaviour of these compounds, an interaction and binding mode analysis, of the best and prioritized compounds per target, was performed. On the ATP catalytic cavity of GSK3 β it is observed that key binders I62, F67, and V70, conforming to the GRR or placed nearby, and Y138 and L188 placed at the C-terminal lobe placed near the AR and inside SP, respectively, are involved on the kororamide stabilization. For CK1 δ it is observed that convolutamine J is stabilized by interacting with several key binders, like I23, which is placed near the GGR and A36, M82, and I148 placed at HP, AR, and PBP, respectively. Looking at DYRK1A ATP cavity, it is observed that kororamide A, at the N-terminal region, is interacting with I165 and V173, as other known inhibitors like meridianin F or the co-crystal 3RA, both placed near the GRR. In the same way, kororamide A is also stabilized by A185, which is found at the HP. At the C-terminal zone it is also interacting with E291 and the L294 conforming PENIL motif and V306 present at the PBP. Finally, kororamide A is also stabilized by L241 and D244, placed near the AR. Looking at the ATP cavity of CLK1, it can be seen that on the N-terminal domain, L167, F172, and V175 at the GRR, and K191 at the HP, that some of them are known key binders, and are interacting with kororamide A. Moreover, on the C-lobe, kororamide A is interacting with F241 coming from the FELL adenin motif, E292, and L295 (a known key binder), placed at the SP, and V324 found at the PBP.

The binding mode of the best compounds, as well as of the four brominated compounds studied, per target pointed out that they are performing key interactions, most of them previously described in other well-known inhibitors. This fact together with the obtained binding energies, reinforce their capacity to behave as inhibitors for the four analyzed kinases, in a similar way to meridianin F.

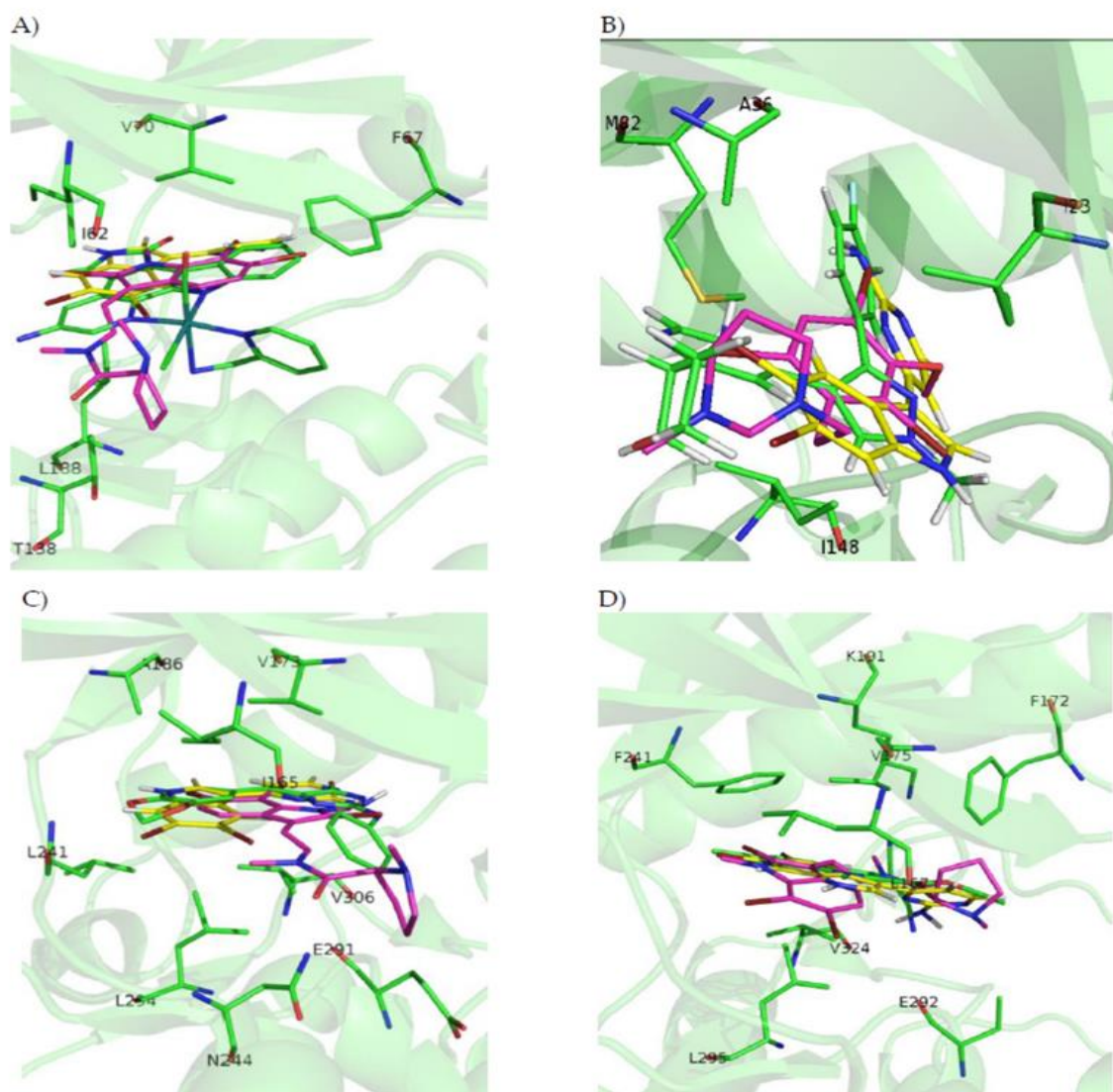


Figure 3. (A) ATP cavity site of GSK3 β (Protein Data Bank ID (PDB) 3PUP) with meridianin F (yellow), the co-crystallized OS1 inhibitor (green), and the best pose of kororamide A (magenta). (B) ATP cavity site of CK1 δ (Protein Data Bank ID (PDB) 4KBK) with meridianin F (yellow), the co-crystallized 1QG inhibitor (green), and the best pose of convolutamine J (magenta). (C) ATP cavity site of DYRK1A (Protein Data Bank ID (PDB) 4AZE) with meridianin F (yellow), the co-crystallized 3RA inhibitor (green), and the best pose of kororamide A (magenta). (D) ATP cavity site of CLK1 (Protein Data Bank ID (PDB) 2VAG) with meridianin F (yellow), the co-crystallized V25 inhibitor (green), and the best pose of kororamide A (magenta). Letters and numbers correspond to their position in the amino acid sequence and the PDB file numbering.

2.3. Marine Natural Products and Indole Scaffold Validation

With the aim of testing the importance of the indole scaffold as structural key feature on the kinases ATP inhibitors and assuming the well-known Structure Activity Relationship (SAR) principle (i.e., structurally similar compounds will have similar biological activities) a substructure search was performed over the MarinLit database, a dataset that includes revised compounds from marine natural products [44]. In that sense, similar compounds to meridianin F and kororamide A and the indole

2.4. Indole Derivatives

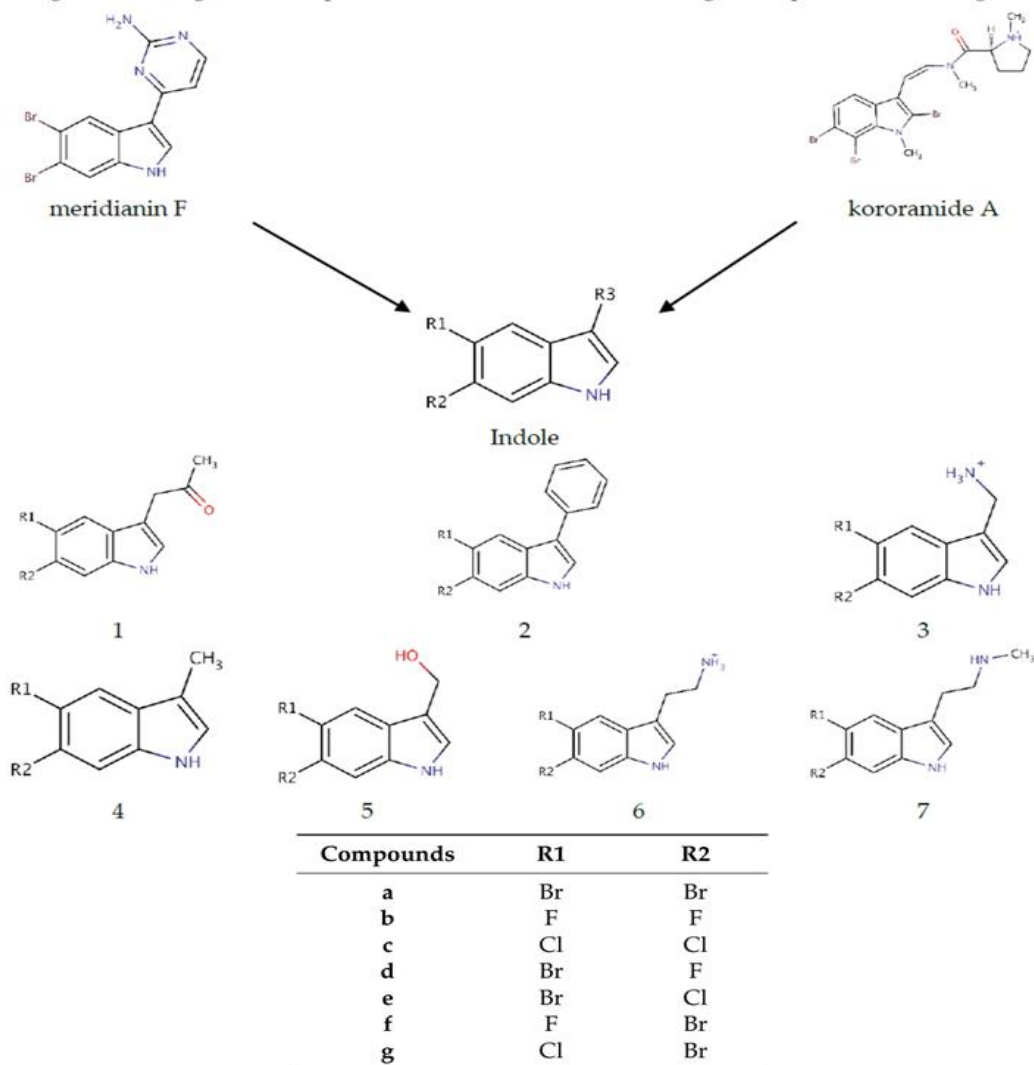
As mentioned above the SAR hypothesis, widely used in drug discovery, has the premise that structurally-similar molecules have similar biological activities and, thus, similar biological targets. Several known kinase inhibitors possess this moiety and some of them even present a halogenated version of it. In a previous work, we showed that meridianin F, which has a halogenated indole scaffold was the more active member of the family, highlighting the role of this moiety. Now, we have observed that kororamide A and B, given the similarity to meridianin following the SAR principle, could be possible inhibitors of these kinases. This fact is at least partially confirmed (further experiments are needed for a complete validation) because of the *in silico* obtained binding energies over GSK3 β , CK1 δ , DYRK1A, and CLK1, reported above. All these facts together, with the observed results in Table 3, made us hypothesize that starting from an halogenated indole moiety and following structural features extracted from meridianin F and kororamide A, we could design indole derivatives that could become kinase inhibitors. Concretely, the indole group was used as a template for the design of a series of seven analogue compounds with different fragments attached to the R3 position of the indole (compounds 1–7; Table 4) and substituted with different combinations of halogen atoms at positions R1 and R2 (a–g combinations; Table 4). Altogether, 49 compounds were designed.

Marine animals have demonstrated to be rich sources of halogenated metabolites and halogenated compounds have a wide range of biological activities [42]. Most halogenated drugs are fluorine (F), followed by chlorine (Cl) and bromine (Br). Contrastingly, for marine-derived molecules, rather than chlorine, the most prevalent halogen found is Br [57]. Halogenated molecules are interesting therapeutic opportunities and it is estimated that one quarter of the total number of final compounds synthesized have an insertion that involves halogens [58]. Halogenated ligands lead to more stable complexes than non-halogenated ligands, and this is important to explain molecular recognition or to planning a screening study [58,59]. Moreover, the capability of halogen atoms to improve oral absorption, lipophilicity, blood brain barrier (BBB) permeability, metabolic and chemical stability, or even potency is well known [58,60]. Therefore, the three mentioned halogen groups at R1 and R2 positions were introduced and evaluated per compound (1–7 + a–g) with the aim of designing the best possible kinase inhibitors (Table 4).

2.5. *In Silico* Binding and Binding Mode Analysis of Indole Derivatives

To analyse the feasibility of the designed compounds as kinase inhibitors by an *in silico* binding analysis, their binding mode and binding strength against GSK3 β , CK1 δ , DYRK1A, and CLK1 were analysed. To start with, docking experiments were performed. A total of 441 poses per target were obtained from the 49 compounds of the set. Thereafter, the binding behaviour of all the poses was analysed, showing that the most populated binding region is, as expected, the ATP cavity. With all these results in hand, best poses per target in terms of binding mode and binding energy were selected to perform short MD simulations, for post-processing docking results. For some derivatives none pose was considered for further studies, as the selection of best compounds was carried out considering not only the binding energy but also the binding mode of each molecule, after an interaction analysis study. The poses that did not present good interactions were discarded. Finally, 166 simulations were carried out, corresponding to diverse poses belonging to 45 compounds for GSK3 β , 45 for CLK1, 46 in the case of DYRK1A and 30 for CK1 δ . After MD simulations, the binding energies of the target-ligand complexes were estimated by molecular mechanics/generalized born surface area (MM/GBSA) calculations. Table 5 summarizes the binding energies of the best indole derivatives, obtained after MD, per compound (1–7) and target. The rest of the binding energies obtained per derivative and target are reported at Tables A1–A4, respectively.

Table 4. From meridianin F and kororamide A, the indole scaffold was selected to derive new compounds. More precisely, seven indole derivatives were designed (compound 1–7). The R3 position was fulfilled with diverse structural elements mainly inspired on meridianin and kororamide structures. Compound 1 with a ketone group, compound 2 with an aromatic ring, compound 3 with a methylamine, compound 4 with a methyl group, compound 5 with methanol, compound 6 with an ethylamine and compound 7 with an ethyl-methylamine. The R1 and R2 positions were completed with the permutation of Br, Cl and F halogen atoms (a–g) over both positions. At the end 49 indole analogue compounds were designed.



As a general result, we observe that all the evaluated compounds present better binding interaction energies against CK1 δ , DYRK1A and CLK1 than GSK3 β , as observed for meridianins [37]. Additionally, as a general trend, compound 1 and 2 always show better energies than the rest of derivatives, highlighting that the fragments introduced in the pyrrole ring (a ketone and an aromatic ring, respectively) of the indole scaffold at R3 position could have beneficial effects to achieve better inhibitory activities over the ATP-binding site of the four studied kinases. Finally, it must be remarked that the designed compounds that do not work against the kinases are different for each one of them, thus opening the door for exploiting these differences in the future to gain selectivity over the four analysed kinases.

Table 5. Summary of molecular mechanics/generalized born surface area (MM/GBSA) calculations of the best derived analogues over the four targets studied. Lowercase letters indicate the halogen substituent group (a–g), as described in Table 4.

	GSK3 β		CK1 δ		DYRK1A		CLK1	
		Binding Energy		Binding Energy		Binding Energy		Binding Energy
Compound 1	a	−30.3141	a	−35.4499	e	−32.8862	g	−30.3541
Compound 2	a	−31.2458	e	−37.8982	a	−37.8422	a	−34.1041
Compound 3	a	−13.8779	g	−28.7631	a	−15.2733	f	−20.4786
Compound 4	a	−27.6481	e	−28.6573	a	−30.7518	e	−28.3695
Compound 5	a	−27.6534	e	−28.5831	a	−31.2535	c	−29.4190
Compound 6	a	−18.5779	a	−26.4630	a	−18.9387	g	−30.7737
Compound 7	a	−18.8955	a	−18.4901	a	−20.8203	g	−25.4765

All energies values are kcal/mol. Maeridianin F results come from our previous publication [37].

2.5.1. GSK3 β

As said, the best docked complexes were selected to perform further analysis. For GSK3 β 75 poses were chosen and over them MD simulations were performed. From the total studied set, and with the aim of analysing the diverse derivatives, the best a–g combination for each of the 1–7 compounds per target was selected. Over the seven best compounds found after MD simulations in terms of binding energy, further analyses were performed, extracting some interesting features. Focusing on the halogen substituents, the best compounds are always those that contain two Br atoms at R1 and R2 position, reflecting the importance of Br substituents observed in previous studies [24,37].

A general pattern regarding the interactions performed by each of the seven best derived compounds at the catalytic ATP binding site was observed. In general, I62, V70, A83, V110, L132, D133, Y134, V135, Y138, and L188 are the most important amino acids for their stabilization over the ATP catalytic pocket (Figure 4). The NH indole group is essential to establish hydrogen bond interactions with the carboxylic acid group (deprotonated under biological conditions) of D133 and/or V135. AR, described by LDYV motif, accommodates the seven best compounds, all of them showing the same binding mode/pose, stabilized by hydrophobic contacts. The indole group is wrapped by N-terminal I62 and V70 residues found near the GRR, together with A83 placed at the HP and C-terminal residues V110 and L188 present at the SP. As the binding mode analysis reveals, all the compounds have the same binding mode, thus binding energy results and MD simulations were used with the aim of identifying some differential features among them. MD analysis reveals that the indole scaffold is maintained wrapped in the same position during all the simulation while the fragments introduced at R3 are more flexible. A binding energy analysis showed that compound 2a has a slightly better energy than compound 1a, although both could be considered good plausible options, as the binding energy differences are around 1 kcal/mol, which seems to point out compound 2a as the best possible inhibitor.

Looking at the literature, our results show that the binding mode displayed by most of the analysed compounds, specially by compound 2a, correlates with the binding mode of known inhibitors, and also that the residues involved on it are key binders [35,61].

2.5.2. CK1 δ

For CK1 δ , 97 docking poses were subjected to MD simulations. Thereafter, the seven best compounds, in terms of binding energy were selected to be further analysed. Differently to GSK3 β , there is not a common binding mode shared by the 7 derived analogues and there is not a specific location of the halogens in the ATP binding site, which can be inferred from the observed binding modes. Although a general pattern could not be observed, there are common features between the studied derivatives than can be highlighted. There is a common behaviour between compounds 1, 6, and 7, and compounds 2, 4, and 5 (Figure 5). For the first group the best halogen composition is

Br-Br (compound **1a**, compound **6a** and compound **7a**), whereas for the second group, the best halogen composition is e (Br-Cl), while for compound **3**, which behaves differently to the rest of the compounds, is g (Cl-Br). In all compounds a Br atom is present, which seems to indicate that this presence could be important to increase the binding strength. In general, with few exceptions, the worst binding energies are obtained when there is no Br atom present. This trend is also observed on the rest of kinases (Tables A1–A4). In addition, an accurate analysis of the most important residues involved on the seven compounds binding mode, was performed. This analysis reveals that despite each compound has a different binding pose, there are conserved interactions at the ATP catalytic cavity. According to that, the most important residues on the binding of the seven compounds to CLK1 δ are I23, A36, Y56, L84, I148, and D149. All seven derivatives are placed between the HP defined by A36 and the residue I23 that is placed near the GRR, both zones located at the N-terminal region and L84, I149, and D149 placed at the AR and PBP at the C-terminal domain. All the interactions observed between the analogues and the residues are mainly hydrophobic contacts. Binding energies reveal that compound **2e** (Br-Cl) seems to be a slightly better inhibitor than compound **1a**, although both can be considered good options as the energy differences are around 2 kcal/mol.

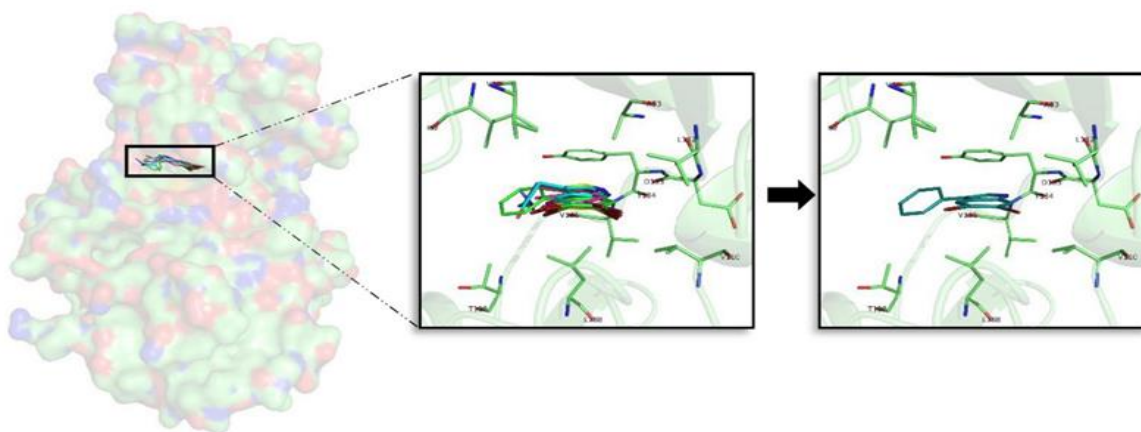


Figure 4. Superposition of the seven best compounds over GSK3 β (PDB code: 3PUP) ATP cavity. The active site amino acid residues involved in the binding of the best compounds and the binding position of each of them are enlarged. In the first enlarged panel the 7 top compounds are represented, whereas in the right panel only compound **2a** is shown.

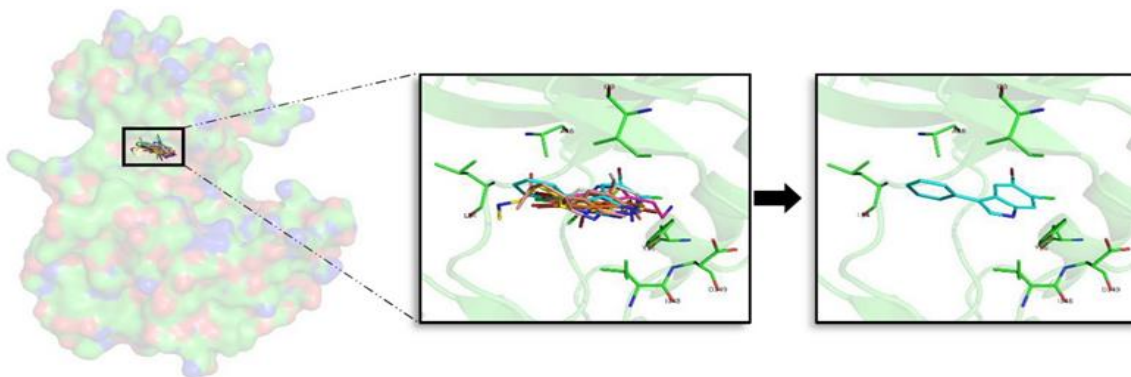


Figure 5. Superposition of the seven best compounds over CK1 δ (PDB code: 4KBK) ATP cavity. The active site amino acid residues involved in the binding of the best compounds and the binding position of each of them are enlarged. In the first enlarged panel the seven top compounds are represented, whereas in the right panel only compound **2e** is shown.

Different studies have been addressed to find novel and potent CK1 δ inhibitors in the last years. Looking at them, it is easy to observe that the interactions made by all of these molecules are aligned, validating it, with the binding mode of our proposed derivatives [27,32,33,36].

2.5.3. DYRK1A

For DYRK1A, 72 docking poses were selected for further analysis. MD simulations were performed over all of them, and thereafter the best compound per target, as for the rest of kinases, was selected. Despite the indole derivatives tested do not show a shared binding mode as GSK3 β , it is more conserved than for CK1 δ . All compounds, except compound 3 that is oriented right upside down and moreover shows the worst binding energy, shared the same placement at the ATP catalytic pocket (Figure 6). Analysing the halogen composition of the best compounds it is observed that Br-Br, at R1–R2 positions, is the most common substituent; only compound 1 has a different combination (Br-Cl). As a general conclusion, as with the other three kinases, the presence of at least one Br atom is important to have a good binding affinity.

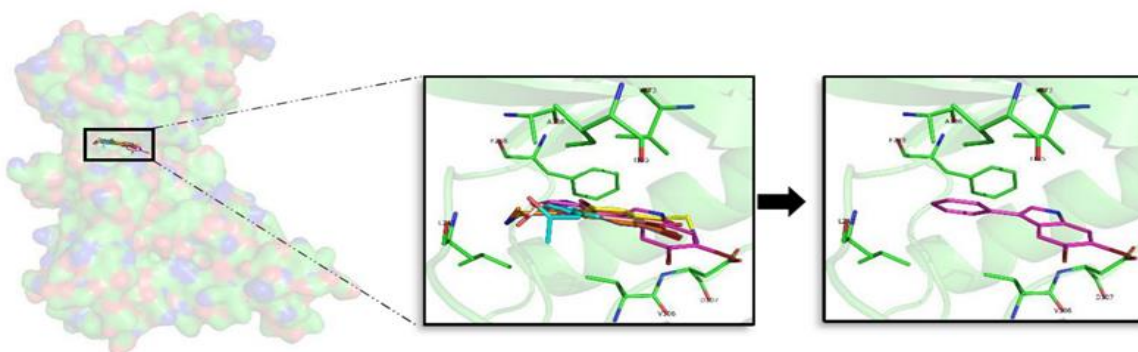


Figure 6. Superposition of the seven best compounds over DYRK1A (PDB code: 4AZE) ATP cavity. The active site amino acid residues involved in the binding of the best compounds and the binding position of each of them are enlarged. In the first enlarged panel the seven top compounds are represented, whereas in the right panel only compound 2a is shown.

Looking forward to extract common patterns from the binding modes of the top seven derivatives, it is clear that the seven compounds placed at the catalytic ATP cavity are interacting with I165 and V173, both residues delimitate the GRR, and A186 that is found at the HP, all of them located at the N-terminal region. The important AR formed by a FEML motif also participates on each of the seven bindings, F238 and L241 being the most important residues to stabilize the analysed derivatives. At the C-terminal region, V306 and D307, present at the PBP, are also key binders. Interaction analysis reveals that most of the interactions performed by the derived analogues were hydrophobic contacts. For DYRK1A, after analysing the MD obtained results, it is observed again that compound 2a is the best derivative in terms of binding energy and binding mode. Interestingly, the observed binding patterns are shared by most of the known inhibitors of this target, which could be found in the literature. Even more, all of them are proposed as ATP-competitive inhibitors like the derivatives we described here [25,26,34].

2.5.4. CLK1

For CLK1, 87 docking poses were selected for further analysis. All of them were subjected to MD simulations selecting then the best one per target. A first binding mode observation reveals that a common binding mode was found for compounds 1, 2, and 6 (Figure 7). These three compounds have the best binding energy, and this could point out the importance of R1, R2 and R3 substituents to gain inhibitory capacity. Compound 3, despite having a similar binding pose, does not show good

energies. The other compounds (4, 5, and 7) show slightly lower binding energies and a different binding mode, even between them. Focused on the halogen groups, in this target there is not a clear trend, as the seven best compounds show five different halogen substituents (a, c, e, f, and g). Despite this fact, not observed in the rest of studied kinases, a similar trend can be seen. Most of the seven top compounds have a Br atom, except compound 5c. Moreover, in agreement with the rest of compounds the seven top derivatives are mainly Br or Cl substituents on R1 or R2 position, with the exception of compound 3f. This seems to suggest that, as for the other targets, all of the analysed halogen substituents combinations could give good inhibitory results, but the presence of a Br is a key factor. In fact, for this target, as seen for the other kinases, the compound 2a is the best one in terms of binding energy.

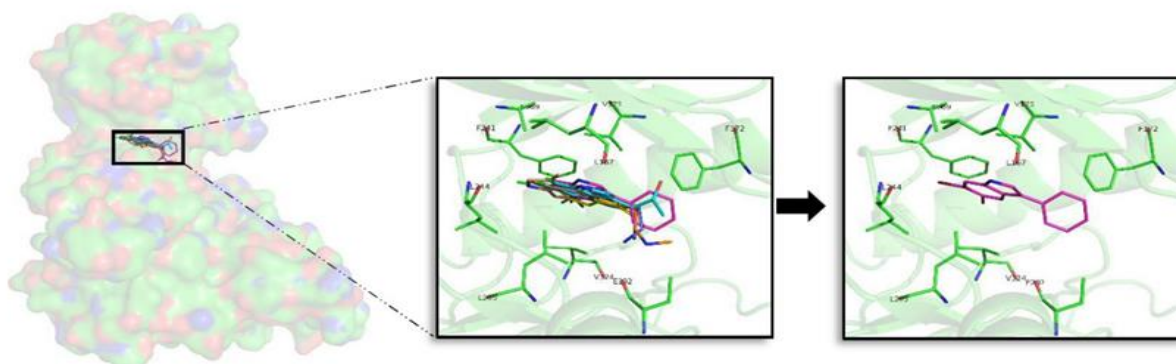


Figure 7. Superposition of the seven best compounds over CLK1 (PDB code: 2VAG) ATP cavity. The active site amino acid residues involved in the binding of the seven best compounds and the binding position of each of them are enlarged. In the first enlarged panel the seven top compounds are represented, whereas in the right panel only compound 2a is shown.

A detailed analysis of the displayed binding modes by each compound at the ATP cavity site, reveals interesting shared patterns. On the N-terminal domain L167, F172, and V175 can be found at the GRR, and A189 at the HP acting as key binders. Adenine motif FELL was also revealed important, in particular F241 and L244. On the C-terminal region, residues E292 and L295 at the SP and V324 placed at PBP are the most important amino acids to stabilize the derived compounds over CLK1. As compound 2a, the other best compounds tend to point their halogen groups between the AR and the HP, fact that facilitate residues as F175 placed at the GRR and E292 or L295 placed at the opposite SP, surround and fixed the indole scaffolds. Interestingly, the binding mode exposed here for the derived analogues in general, and also for the best compound, 2a, in particular, is validated by other inhibitors reported in the literature [30,31].

The *in silico* binding studies performed over the four kinases indicate that the derivatives coming from compound 2, 2a and 2e, located at the middle of the ATP binding cavity, seem to be the most plausible ATP competitive inhibitor. However, other derivatives, especially for compound 1 should not be discarded. In general, the presence of the benzene ring at position R3 could have a more positive influence on compounds stabilization at the catalytic site than other substituents. Looking at the literature, several inhibitors described for GSK3 β , CK1 δ , DYRK1A, and CLK1, as well as other members of the protein kinase family, have aromatic rings in the terminal positions. Moreover, the analysis of the effect of the halogen groups used as substituents at R1 and R2 positions pointed out that its presence can influence the binding strength of the complex (ligand-target). In general, if at least one of the substituents is a Br atom the binding energy is better. An interesting trend found here is that Br seems to be the “best” halogen followed by Cl and F, which, in general, gives worst binding energies. This finding is in line with what is observed in nature, since natural halogenated indole

alkaloids contain mostly bromine and chlorine, being the iodinated and fluorinated compounds less abundant [43].

2.6. Selectivity

One of the most important challenges on the design of novel kinase inhibitors is the lack of selectivity over the ATP binding site, which is critical in clinical effectiveness of most drugs [46,48]. Most kinase small-molecule inhibitors bind to the ATP catalytic cavity near the AR and wrapped by GRR and HP on the T-lobe and SP and PBP at the C-lobe. The herein performed study does not reveal any significant selectivity over the four kinases for any of the analysed compounds, which could be easily observed looking at the obtained binding modes and energies. However, analysing the residues involved on the binding and the regions occupied by the analogues, some interesting trends that could be exploited in the future can be observed. Interestingly, regarding the binding modes, the best binding energies were obtained on those compounds that are (partially) placed at the PBP. This region, that is very exposed to the solvent and is not usually exploited to gain binding affinity, can be useful to improve the inhibitors selectivity since it contains non-conserved amino acids [46].

Regarding the binding energy results per se, without having into account the binding mode, remarkable significant differences are not observed. The best compound for each target (**2a** and **2e** respectively) comes from the same scaffold, compound **2a** being the best theoretical inhibitor for three of the four targets. If we analyse the binding energies of these top compounds, compounds **2a** and **2e** over DYRK1A and CK1 δ , respectively, show a better interaction energy, around 6 kcal/mol of difference, respect to the binding energy of compound **2a** over GSK3 β and 3 kcal/mol over CLK1. However, although a slight preference could be inferred from this, the binding of these four compounds to all the four targets is possible with a reasonably good strength. In general, the main differences are observed between the derived compounds **2** (mainly) and **1**, which seems to have better energies than those molecules coming from analogues **3** to **6**, and especially with respect to the molecules coming from analogue compound **7** (Table 5). For GSK3 β the best compounds coming from derivatives from **2** and **1** (1 kcal/mol of difference between them) are displaying the best interaction energies, followed by those from analogues **4** and **5** (around 4 kcal/mol of difference to compound **2a**), and finally the worst compounds come from analogues **3**, **6**, and **7** with differences around 13 to 17 kcal/mol respect compound **2a**. In the case of CK1 δ , as for GSK3 β , the top ranked compounds from analogues **2** and **1** (1.5 kcal/mol of difference between them) have the best binding energies, followed by those from analogues **3–6** with differences around 9 to 11 kcal/mol with respect to compound **2e**, and finally compound **7a** with a difference of around 19 kcal/mol with respect to **2e**. For DYRK1A, the best from compound **2** is the top molecule in terms of interaction energy. Compound **1a** shows a difference of around 5 kcal/mol, whereas compounds **4** and **5** present differences between 6.5 and 7 kcal/mol, respectively, and compounds coming from scaffolds **3**, **6**, and **7** between 17 and 22.5 kcal/mol. In the case of CLK1, compound **2a** has the better binding energies, followed by those from analogues **1** and **6** (differences around 3 kcal/mol), molecules derived from compounds **4** and **5** (differences around 5 kcal/mol), and finally those from analogues **3** and **7**, that show differences around 8.5 to 13 kcal/mol respect to the binding energy

Looking to the **1–7** compounds per target, it can be observed that for compounds **4** and **5** the binding energy differences between the top a–g derivatives range between 2 and 4 kcal/mol between the four kinases. For compounds **1**, **2**, and **7** the differences are higher, ranging between 2.5 and 7 kcal/mol, depending on the compound and target. Finally, for compounds derived from analogues **3** and **6** the differences are even higher, ranging between 8 and 15, and 4 and 12 kcal/mol, respectively. In general, there is not any noticeable selectivity trend derived from the binding energy, although there are some features that could be further exploited. For instance, for DYRK1A and CLK1 an aromatic ring at R3 position is the best choice to gain activity over them, whereas for GSK3 β and CK1 δ a ketone group at this position could also work, enhancing a way to design selective compounds at least for some of the four kinases.

Exploring the effect of the halogen atoms over the binding strength, as said above, some general trends could be observed but again, its presence does not give any clearly marked or significant selectivity trend between targets. The presence of Br atoms seems to increase the binding strength more than the presence of Cl or F, being in general Cl “better” than F to get good energetic results. However, a possible selectivity feature could be observed due to compound **2e**. Docking energy results are similar for the four kinases, but it only performs good interactions for CK1 δ . This is the reason why MD simulation over this compound was only performed in complex with it, while for the other three kinases it was not selected. Compound **2a** gave good docking energies for all four targets but performed good interactions only with GSK3 β , DYRK1A and CLK1, so the fact of having a Br-Cl combination at R1-R2 plus an aromatic ring at R3 could be a sign of selectivity over CK1 δ , although it should be further explored, as other Cl combinations give good results for the other kinases (Tables A1–A4).

2.7. **2a** and **2e** Unbinding

To reinforce and validate the observed binding trends, as well as to find a differential feature that could help to enhance the selectivity of future derived compounds over the four kinases, steered molecular dynamics (SMD) simulations were performed. Since at the energy and binding mode level there are no significant differences, we intended to see if there was some type of selectivity derived from the protein structure that influences the facility/difficulty of unbinding of the most promising inhibitory compounds **2a** and **2e** per target (Figure 8).

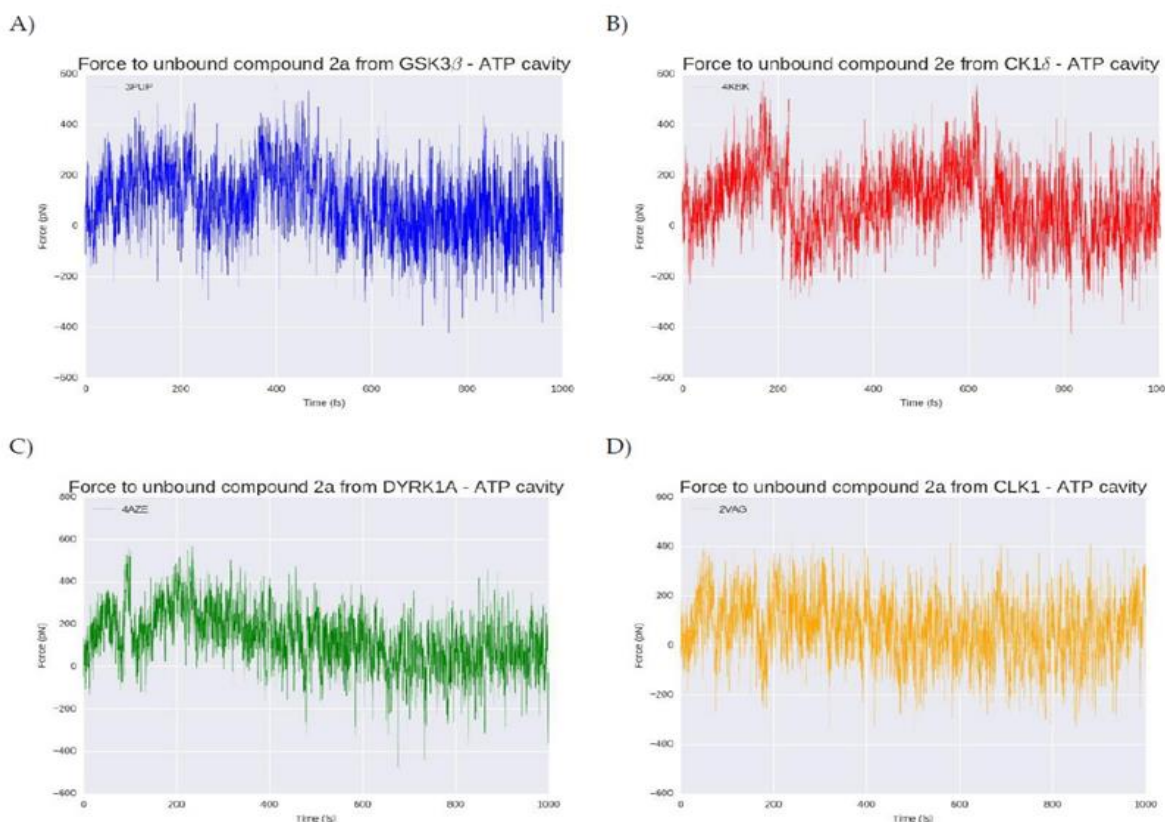


Figure 8. Exerted force in piconewtons (pN) needed to (A) remove compound **2a** (blue) from the GSK3 β ATP catalytic cavity, (B) remove compound **2e** (red) from the CK1 δ ATP catalytic cavity, (C) remove compound **2a** (green) from the DYRK1A ATP catalytic cavity and remove compound **2a** (orange) from CLK1 ATP catalytic cavity. The x axis represents the computational residence time in femtoseconds (fs).

At the beginning of each simulation ($t = 0$), the compound is in the bound state, placed inside the ATP cavity interacting with the residues previously described. After one nanosecond (ns), at the four kinases, the ligand is completely out of the cavity. In the case of GSK3 β (Figure 6A) a force of around 400 pN (piconewtons) is needed to extract compound **2a** from its catalytic cavity. The compound dissociation from the target take place at 200 femtoseconds (fs), moment where the force decrease approach zero pN which means that the compound is out of the cavity. For CK1 δ (Figure 7B) that hosts the best compound in terms of binding energy, **2e**, the necessary force to break the ligand-target complex is higher than for the GSK3 β complex, with forces that reach up to 600 pN. The ligand unbinding takes place at a similar time than for GSK3 β complex, around 200 ps although it takes slightly more time. The dissociation of compound **2a** from DYRK1A (Figure 7C) is done in two phases. A primary rupture force seems to occur before 100 fs, and immediately afterwards it could be observed the highest energy point (around 500 pN), corresponding to the second break. A visual inspection of the SMD confirmed that at this moment, the compound is still inside the ATP pocket. Over 200 fs, the force, after a progressive decrease, arrives to zero pN. This progressive decline correlates with the progressive loose of interactions during the way out of the compound from the catalytic cavity. For the CLK1-compound **2a** complex (Figure 7C) a similar situation is observed. A primary rupture around 100 fs, moment in that, as for DYRK1A, the compound is still placed at the ATP binding site and it is not until later, at 200 fs, when a sudden drop in the energy can be observed, indicating the completely loss of interactions and therefore, the leaving of the cavity. As a general trend, around 200 ps compounds **2a** and **2e** leave the catalytic pocket of the different kinases, requiring a different amount of force that is in line with the observed binding energy. Usually, the better the binding energy, the higher the force needed to break the complex and the longer the residence time. In that sense, the SMD results corroborate what has been seen so far: CK1 δ -**2e** and DYRK1A-compound **2a** complexes that have the higher binding energies also seem to have (slightly) longer residence times and require a higher force to take out their respective ligands from their catalytic pockets. There is not any feature that suggests a selectivity trend derived from the unbinding process that could not be extracted from the binding energy results. Compound **2a** is more selective (it binds stronger and requires a higher effort to remove it) for DYRK1A than for GSK3 β and CLK1 but could bind to all of them. Compound **2e** seems to be more difficult to unbind that compound **2a**, but this correlates with the higher binding energy it shows after MD.

2.8. Pharmacokinetic Properties of Kororamide A–B, Convolutamine I–J, and the Designed Derivatives

Due to the importance of pharmacokinetics (PK) and its impact on drug discovery process, convolutamine I–J, kororamide A–B and the whole set of 49 analogues compounds were analysed, studying their ADME/Tox features. The PK properties of the two best derived compounds **2a** and **2e** are summarized on Tables 6 and 7. The full set of derivates were also analysed and results can be found at Table A5 (absorption and distribution) and Table A6 (metabolism, excretion and toxicity).

Table 6. Summary of absorption and distribution properties of the two best compounds **2a** and **2e** found on the four studied kinases. BBB: blood brain barrier, PPB: protein-protein binding, VDss: steady state volume of distribution, CNS: central nervous system.

Compound	Mol Weight	LogS	Absorption			Distribution				
			P-Glycoprotein	Caco-2 Permeability	Intestinal Absorption	LogP	BBB	PPB	VDss	CNS Permeability
Compound 2a	351	−6.1	inactive	Moderate	90.067	4.1	0.477	High	0.234	−0.894
Compound 2e	290.1	−5.7	inactive	Moderate	91.036	3.8	0.508	High	0.076	−0.92

The first PK property analysed was molecular weight, and all compounds show values under 500 Dalton (Da). The higher molecular mass was found for compound **2a** with 351 Da, which is in good agreement with the sizes that a small therapeutic molecule that should cross the BBB should have.

Table 7. Summary of metabolism, excretion and toxicity properties of the two best compounds **2a** and **2e** found on the four studied kinases. CYP: cytochrome, OCT2: organic cation transporter 2, hERG: human ether-a-go-go gene, MRTD: maximum recommended tolerated dose.

Compound	Metabolism	Excretion	Toxicity			
	CYP450	OCT2 Substrate	hERG	MRTD	AMES Toxicity	Hepatotoxicity
Compound 2a	Yes	No	<4.0	0.673	Yes	No
Compound 2e	Yes	No	<4.0	0.641	Yes	No

2.8.1. Absorption Properties

Absorption describes the process by which drug candidates move from the point of administration to the blood. LogS descriptor confirmed good solubility in water and good bioavailability for each compound. The derivatives coming from compounds **2** show values that are between -5.1 and -6.1 , while for the rest of the derivatives, values are between -3 and -4 . Caco-2 permeability revealed medium to high values for all the compounds, except for kororamide A, which was low. The compounds that have a benzene at R3 position as well as the derivatives with F at R1 and R2 positions show moderate permeability and should be optimized in the future. Regarding P-glycoprotein (Pgp) binding, no compound was predicted to act over it. The interaction with Pgp has many pharmacological implications that could result in pharmaceutical advantages or contraindications. For instance, Pgp modulation has been suggested as a mechanism to improve CNS pharmacotherapy [62–65], but none of the derivatives here seem to have this ability. On the other hand, intestinal absorption values higher than 30% are considered well-absorbed compounds, and for the entire set obtained values are higher than 89%. All of these absorption results suggest good absorption properties for the 49 designed derivatives, plus kororamides and convolutamines.

2.8.2. Distribution Properties

Distribution describes the migration of a compound from the circulation to the extravascular system. LogP values lower than 5 indicate that the compounds have an appropriate hydrophobicity and permeability. In that sense, the derivatives coming from compound **2**, as well as convolutamine J and kororamide A have the highest values (≈ 4) while the rest of compounds are between 2 and 3. Opposite to LogP behaviour, plasma-protein binding (PPB) and steady state volume of distribution (VDss) are not showing as good tendencies for the best derivatives compounds coming from scaffold 1 and 2, and Convolutamine J. Most of the analysed molecules showed medium to high PBP values (except kororamide A and compounds **3b**, **5b**, **6b**, and **7b** with low PPB values) indicating that a high percentage of the administrated compound will be found attached to proteins, affecting its diffusion and its efficiency. As less bound a drug is to plasma proteins, the more efficient it is, as it can traverse cell membranes or diffuse. Regarding VDss, derivatives from scaffolds 3, 6, and 7 and convolutamine I–J plus kororamide A, have high VDss values (>0.45), while for the rest of compounds, its distribution is low to medium, in a close agreement with PBP results. BBB descriptors with results higher than >0.3 reveals good distribution to the brain, as they could pass the blood brain barrier. The highest values are found for convolutamine I, kororamide A, and the derivatives coming from scaffolds 2 and 4, as well as for compounds **5b** and **7b**. However, it should be considered that most of the compounds not predicted to cross BBB have values near the threshold. In addition to BBB, Central nervous system (CNS) permeability was measured. This measure seems to be more precise than BBB, as it is a more direct measurement [66]. Kororamides and convolutamines do not show good permeability values, whereas all the derived compounds showed good results (> -2) allowing us to consider that most of the designed compounds could penetrate the CNS, specially the compounds coming from scaffolds 2 and 4, among which there are the two best candidates **2a** and **2e**.

2.8.3. Metabolism Properties

Cytochrome (CYP) P450 is an important enzyme used to predict drug metabolism. Many drugs could be deactivated or activated by CYP450, as cytochrome P450 enzymes that can be inhibited or induced by drugs, resulting in clinically significant drug-drug interactions that can cause unanticipated adverse reactions or therapeutic failures. Our results revealed that all the analysed compounds, except kororamide B, are likely to be metabolised by CYP450, so their properties should be carefully analysed to design lead compounds from the herein-studied molecules [67,68].

2.8.4. Excretion Properties

Regarding excretion properties, describing the transport of drugs into the urine or bile, good results were obtained. It was found that only kororamide A and B tend to act as a substrate of the organic cation transporter 2 (OCT2 or Solute carrier family 22 member 2, SLC22A2), which means that, in general, and for the two best derivatives **2a** and **2e**, non-clearance problems and adverse interactions with co-administrated OCT2 inhibitors are expected.

2.8.5. Toxicity Properties

During drug development, safety is always the most important issue, including a variety of toxicities and adverse drug effects that should be evaluated in preclinical and clinical trial phases [69]. Between the measured properties, the inhibition of the potassium channels encoded by the human ether-a-go-go gene (hERG) is basic. Our results indicate that none of the compounds seem to be toxic due to hERG. In the same way, none of the designed derivatives is susceptible to be hepatotoxic. However, convolutamine I and J as well as kororamide A tend to be hepatotoxic. Looking at AMES toxicity, which predicts mutagenic and carcinogenic properties, our results revealed that the derivatives from compound 2, as the top derivatives **2a** and **2e**, and kororamide B are predicted to be toxic, while the rest of the set does not. Regarding the maximum recommended tolerated dose (MRTD), the four brominated alkaloids as well as compounds coming from scaffolds 3, 6 and 7 showed low values/doses, which is not the best scenario, whereas the rest of the compounds present good MRTD values.

The well-known Lipinski's rule of five, formulated in 1997 and that remains in force [70] was also used in combination with the different ADME/Tox properties described above with the aim of evaluate/determine druglikeness of the analysed compounds. To assess how druglike a substance is based on Lipinski's rules it is accepted that it should have (1) not more than five hydrogen bond donors, (2) ten hydrogen bond acceptors, (3) a molecular mass less than 500 Da, and (4) a LogP not greater than 5. Focusing on the two best compounds (**2a** and **2e**), both have one hydrogen donor and no acceptors. Additionally, as seen in Tables 6 and 7, the other Lipinski requirements are met. Thus, taking into consideration all the ADMET results described previously, these two compounds can be proposed as good hit candidates, having into account that some properties, such as the possible carcinogenesis and mutagenesis problems should be carefully addressed. In fact, absorption, distribution metabolism, excretion, and toxicity properties should be more or less improved for all the designed compounds, in a further hit to lead (H2L) optimization process. Toxicity should be removed, and compounds interaction with cytochrome P450 carefully analysed and, given the case, eliminated or modulated. Moreover, Caco-2 permeability could be increased, as well as their distribution properties should be improved, lowering the PPB and VDss, to be able to diffuse and penetrate into cells easily.

3. Materials and Methods

3.1. Computational Virtual Screening

It is well known that there is a correlation between (chemical) structure and (biological) activity, the structure activity relationship (SAR). This SAR is widely exploited in many aspects of the drug discovery pipeline, ranging from compound screening to lead optimization processes, at the

experimental and computational levels. Herein, we have performed a 2D virtual screening search over the *MarinLit* database using its substructure search functionality. Using as an input meridianin F and kororamide A (the two indole compounds that have shown a better binding strength against the four analysed kinases), as well as the indole scaffold alone, a similarity search was performed over *MarinLit*, obtaining a list of compounds having an indole scaffold in their structure and/or are structurally similar to meridianin F and/or kororamide A. The name and structure of the similar compounds can be found in Figure A1.

3.2. Structure Modelling

Convolutamine J, I, and kororamide A and B, were modelled from Dashti et al. [38]. Ligands were prepared to generate energetically-minimized three dimensional (3D) coordinates.

To perform computational work, obtaining good structures to start with is crucial, so prior to any calculation, good computational models should be constructed. The structures of the analysed targets were modelled from 3D crystal structures obtained from the Protein Data Bank (RCSB PDB) [71]. Those structures represented human targets and are the best structures in terms of sequence coverage, of the whole target, in general, and of the binding pocket of each target, in particular. Since all the four kinases biological assembly is in monomeric forms, GSK3 β and CK1 δ chain B and DYRK1A and CLK1 chain A were respectively selected to perform further studies. To do so, due to the fact that the four studied targets have 3D crystallographic structures, the ATP competitive inhibitor OS1 was co-crystallized with GSK3 β (PDB: 3PUP) [35], 1QG was co-crystallized with CK1 δ (PDB: 4KBK) [36], the crystal structures of DYRK1A was complexed with 3RA (PDB: 4AZE) [26] and, finally, V25 was co-crystallized with CLK1 (PDB: 2VAG) [31] and used as a template to perform rigid docking calculations using Itzamna (Mind the Byte.SL, Barcelona, Spain) [72].

3.3. Docking Calculations

Docking calculations can identify small molecules (ligands) that fit well into the putative binding pocket of a given protein (target) in an optimal way. Without any other specification, generally speaking, docking refers to classical (rigid) docking where only the movement of the ligand is allowed [73]. This kind of calculation allows to elucidate the binding mode, as well as the binding strength of the analysed molecules. Moreover, the molecules could be ranked according to their binding energy. However, this static model is far from real, because proteins are dynamic entities that, to carry out any biological function, need to move. Thus, this movement should be taken into account to obtain good predictions that could be compared with experiments [74,75]. A good option to take the protein movement into account is post-processing docking results by MD simulations [76,77]. MD simulations used to improve docking prediction as they allow observing the induced fit events or, in other words, the conformational adaptation of the target to the ligand, not only the ligand adaptation as happens with rigid docking experiments. Moreover, the stability of the docked complex could be analysed using this pipeline [74,77].

All docking calculations were performed using Itzamna software (Mind the Byte.SL, Barcelona, Spain) [72]. Docking studies were performed between kororamide A–B, convolutamine I–J, and the set of 49 derived compounds against GSK3 β , CK1 δ , DYRK1A, and CLK1. Two runs were carried out for each calculation to avoid false positives. As the used 3D crystal structures of the kinases were co-crystallized with a ligand, this cavity was employed to dock the analysed compounds for each of the four targets.

3.4. Molecular Dynamics Simulations

MD simulations are able to capture the dynamic nature of proteins and bimolecular systems, in general. Herein, short (1 ns) MD simulations were performed using the NAMD program, version 2.11, over the best ligand-target complexes (the top ranked compounds according to the docking binding energies) [78]. The Amber ff99SB-ILDN and the General Amber Force Field (GAFF) set of

parameters were employed for modelling receptors and ligands, respectively [79,80], as both forcefields have been extensively tested and used in protein-ligand complexes, giving satisfactory results in several studies [79–82]. An antechamber was employed to calculate the ligand GAFF parameters and the leap module of Amber Tools to obtain the parameters of the proteins [83,84]. Explicit solvent MD simulations were carried out using a time-step of 2 fs and a TIP3P water model imposing periodic boundary conditions via a cubic box [85]. The distance between the complex and the edge of the box was set to 10 Å. The particle-mesh Ewald method was used to calculate the electrostatic interactions. The temperature and the pressure were kept constant at 300 K and 1 atm, respectively, using Langevin dynamics and a Langevin piston barostat. Bond lengths to hydrogens were constrained with the SHAKE algorithm [86]. Before starting the production simulation, all position restraints were removed, the system was firstly submitted to an energy minimization, following by a solvent equilibration (using harmonic position restraints on the heavy atoms of the protein-ligand complex) and, finally, to a slow heating-up, from 0 to 300 K.

3.5. Molecular Mechanics/Generalized Born Surface Area

After performing MD simulations to estimate the ΔG binding free energy of ligand-target complexes, Molecular Mechanics Generalized Born Surface Area continuum solvation (MM/GBSA) reweighting techniques were employed [87]. These techniques outperform docking results because they are employed after MD, thus taking into account the dynamic behaviour of the protein-ligand complexes. However, it should be highlighted that although improve docking binding energy values, are far to be experimentally comparable. Herein, like in other studies, we applied MM/GBSA reweighting techniques over the generated MD trajectories for post-processing docking results [77,88,89]. The MD generated trajectories were employed as input of the MM/GBSA calculations that were realized using the MMPBSA Python algorithm contained within the Amber Tools suite [89].

3.6. Steered Molecular Dynamics

Steered molecular dynamics (SMD) is a simulation tool used to explore processes, which cannot usually be achieved by standard MD simulation, such as ligand-protein unbinding or certain protein conformational changes. Here, we have employed it to study ligand unbinding processes. In that sense, in SMD simulations, a time-dependent external force is applied to the ligand, from an internal atom of the protein, to facilitate its unbinding. For a given ligand bound to a target, it allows establishing a theoretical correlation between unbinding forces and residence time and, in turn, its inhibitory capacity; the larger the force and time needed to unbind a ligand from a receptor the higher its binding affinity [90–93].

SMD simulations were performed using NAMD version 2.11 [78]. Compounds **2a** and **2e** over GSK3 β , DYRK1A, CLK1, and CK1 δ , respectively, were performed. The last frame obtained from the postprocessing MD simulations was used as an input. A harmonic constraint force constant of 4 kcal/mol/Å with a constant velocity of 0.00002 Å/ns was applied. The time length for each simulation was 1 ns, using a time-step of 2 fs, which was enough to observe the entire ligand unbinding process. The rest of the parameters of the simulations were the same employed for MD simulations. The generated trajectory was finally analysed using visual molecular dynamics (VMD) to extract the exerted force (pN) per simulation frame [94].

3.7. Interaction Analysis

To analyse the key residues of the binding pocket involved in the ligand binding, we deeply analysed the obtained binding modes after docking and MD simulations, comparing the obtained results against “known binders” of each of the targets. The “known binders” are important residues for the interaction of reported substrates/inhibitors and were identified at the literature [25–27,30–36,61] as well as in an in-house, recently-constructed database. It was built by crossing ChEMBL and the

PDB [62], and it is composed of all PDB structures per UniProt ID with active compounds, plus the residues with which they interact [95,96].

3.8. ADME/Tox Properties Prediction

ADME/Tox estimation of the analysed compounds was performed using machine-learning (ML) models. These models were enclosed within the ADMETer software tool (Mind the Byte.SL, Barcelona, Spain) and the pkCSM webserver, respectively [66,97]. Concretely, molecular weight, MRTD, logS, logP, Pgp, caco-2 permeability, BBB penetration, PPB, VDss, CNS penetration, intestinal absorption, AMES toxicity, CYP450 metabolism, hepatotoxicity, hERG binding, and OCT2 binding were predicted. A Pgp models were generated by Random Forest against Poongavanam and co-workers (Pgp) and Sedykh et al. datasets [98,99]. The other models and the datasets against the models that were generated are further explained by Llorach-Pares et al. and Pires and co-workers, respectively [37,66].

3.9. Graphical Representations

Graphical representations of protein-ligand complexes were prepared using PyMOL version 1.7 and PLIP version 1.3.3 [100,101]. 2D ligand images were prepared using the RDKit [102] Python library and SMD plots using matplotlib [103] and seaborn [104] Python libraries.

4. Conclusions

Kororamide A–B and convolutamine I–J can act as tau (GSK3 β and CK1 δ) and dual-specificity (DYRK1A and CLK1) protein kinase inhibitors. Kororamide A–B are brominated indole alkaloids structurally very similar to meridianins. Only taking this fact into account and following the SAR principle could a kororamides kinase inhibitory effect be hypothesized. Therefore, the *in silico* binding results we obtained were expected. These results corroborate the idea of that kororamides could be kinase inhibitors with a therapeutic role in AD. Convolutamine I–J, which are not structurally similar to meridianins or kororamides, but are brominated heterocyclic compounds like other known kinase inhibitors, have also shown a plausible inhibitory capacity over GSK3 β , CK1 δ , DYRK1A, and CLK1. Altogether, the results highlight the role of the indole scaffold and the halogen substituents on these kinase inhibitions, being common features among all the compounds.

However, as happened with several other compounds acting over kinases, their main problem is the selectivity. These compounds seem to be somehow selective for one of the kinases, and it is clear which kinase is the preferred one to bind and which one is undesired but, in general, the obtained energy differences are not enough to consider that these compounds are selective. Moreover, the four brominated alkaloids should be optimized according to their ADMET properties. They have moderated good absorption properties, but caco-2 permeability could be increased, especially for kororamide A, as well as the distribution properties. Additionally, the four compounds show a tendency to have toxicity problems that should be carefully revised, in the same way as compounds interacted with cytochrome P450, although kororamide B does not show this cytochrome interaction.

Through the inclusion of convolutamine into the analysis (as they are brominated but not indole compounds), as well as the exploration of some indole-containing compounds from the MarinLit database, we intended to disentangle whether the indole or the halogen substituents presence is the most important feature to gain activity over the four kinases studied. However, the main conclusion extracted is that, individually, both are equally important, and probably the best way to profit from both features is combining them into halogenated indole scaffolds.

Natural products possess very large therapeutic potential, as reported here and in the related literature. Within natural products, those of unexplored marine origin are of great interest in the discovery of novel chemical structures, since they harbour most of the biodiversity of the world [40,105]. Life started in the oceans and many organisms live only there. Due to this, they should be successfully exploited in the future using sustainability criteria and respecting biodiversity. All this makes computational CADD contributions very relevant, since no biological sample is needed to perform

an in silico analysis [106–109]. Taking all of these facts into account, and profiting from the scaffolds showed by meridianins and kororamides (examples of the importance of halogenated indole scaffolds to gain kinase inhibitory activity), we designed 49 marine natural product derivatives. Concretely, we performed a detailed computational study for the development of specific tau (GSK3 β and CK1 δ) and dual-specificity (DYRK1A and CLK1) protein kinase inhibitors, starting from marine natural products, meridianin F and kororamide A, until achieving the rational design of indole scaffolds derivatives as possible ATP-competitive kinase inhibitors for the treatment of AD. We illustrated how the indole derivate compounds derived from scaffold 2 (an indole with an aromatic ring at R3 position and halogen substituents at R1 and R2), in general, and compounds **2a** and **2e**, in particular, could be proposed as good hit compounds to start an H2L optimization process. Altogether, it could be concluded that kororamides, especially A, convolutamines, especially J, and compounds **2a** and **2e** could be possible ATP-competitive inhibitors against GSK3 β , CK1 δ , DYRK1A, and CLK1. These results come from in silico experiments and further in vitro and in vivo studies are required. Our results constitute a promising starting point for the development of novel anti-AD drugs.

Author Contributions: M.S.-M., L.L.-P. and C.A. conceived the study and designed the experiments. L.L.-P. carried out the experiments, whereas A.N.-C., C.A. and M.S.-M. supervised them. All the authors analysed and discussed the results as well as wrote the manuscript.

Funding: This research was partially funded by an Industrial Doctorate grant from the Generalitat of Catalonia to L.L.-P. (DI 2016-051).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

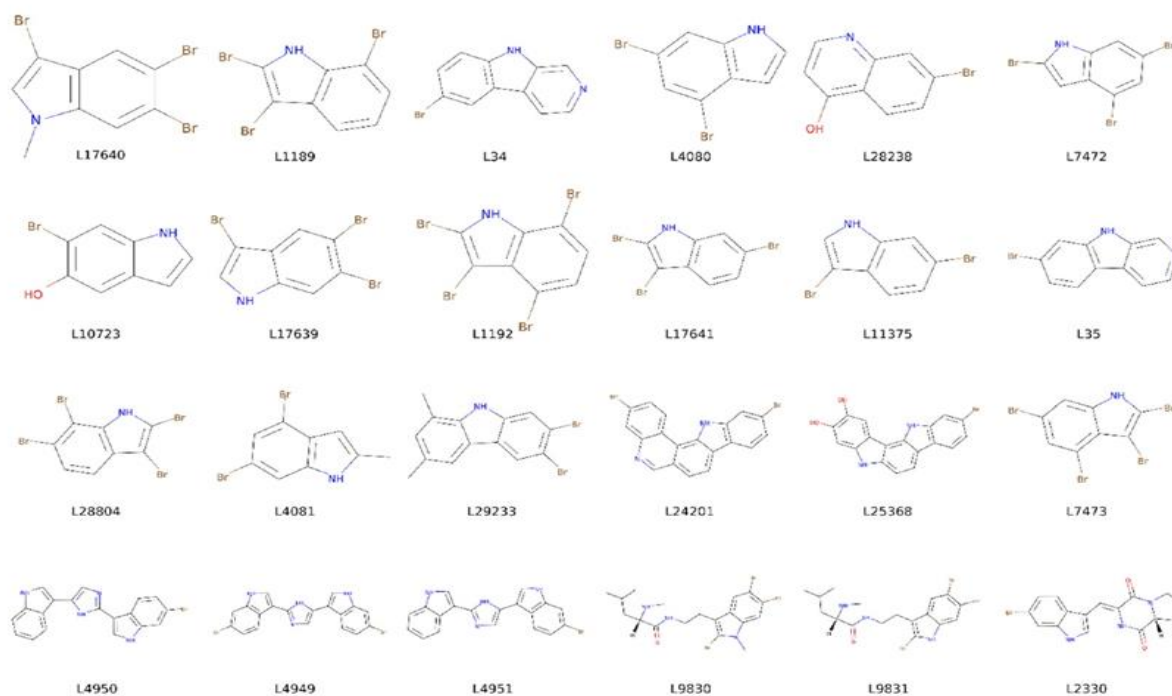


Figure A1. Structures of the 24 compounds found at MarinLit database after a similarity based substructure search using meridianin F, kororamide A, and the indole scaffold as a seed.

Table A1. Summary of classical rigid docking calculations of the derived analogues compound set over the GSK3 β and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the best posed analogue compounds. Lowercase letters represent the employed halogen group (a–g).

				GSK3 β						
		1	2	3	4	5	6	7		
Binding Energy	R0/R1	a	−6.9/−6.9	−8.1/−8.1	−5.9/−5.9	−6.5/−6.5	−6.3/−6.3	−6.2/−6.2	−6.6/−6.6	
MM/GBSA			−30.3141	−31.2458	−13.8779	−27.6481	−27.6534	−18.5779	−18.8955	
Binding Energy	R0/R1	b	−6.8/−6.8	−7.9/−7.9	−5.1/−5.1	−6.6/−6.6	−5.6/−5.6	−5.8/−5.8	−5.9/−5.9	
MM/GBSA			−22.0902	−23.9910	−7.0321	−17.6371	−19.3959	−9.2248	−20.6117	
Binding Energy	R0/R1	c	−6.8/−6.8	−8.1/−8.1	−5.8/−5.8	−6.3/−6.3	−6.3/−6.3	−6.2/−6.2	−6.7/−6.7	
MM/GBSA			−26.1345	−28.4927	−10.3167	−24.8857	−23.8207	−14.7674	−17.0307	
Binding Energy	R0/R1	d	−7/−7	−8.1/−8.1	−5.2/−5.2	−6.8/−6.8	−6.5/−6.5	−6.1/−6.1	−6.7/−6.7	
MM/GBSA			−26.2805	−29.6158	−12.1225	−22.7717	−23.6754	−14.4347	−13.6539	
Binding Energy	R0/R1	e	−7/−7	−8.1/−8.1	−6/−6	−6.7/−6.7	−5.8/−5.8	−6.1/−6.1	−6.7/−6.7	
MM/GBSA			−27.2898	−19.5564	−25.4497	−26.4593	−17.3314	−18.9577		
Binding Energy	R0/R1	f	−6/−6	−8.2/−8.2	−5.8/−5.8	−6.3/−6.3	−6.2/−6.2	−6/−6	−6.4/−6.4	
MM/GBSA			−26.4517	−6.2475	−23.7315	−19.8909	−13.3104			
Binding Energy	R0/R1	g	−6.2/−6.2	−8.2/−8.2	−5.8/−5.8	−6.5/−6.5	−6.2/−6.2	−6.3/−6.3	−5.9/−5.9	
MM/GBSA			−28.2864	−14.2501	−24.4026	−24.7124	−16.8986	−20.7272		

To avoid false positives, each docking calculation was performed twice (R0/R1). All energies values are kcal/mol.

Table A2. Summary of classical rigid docking calculations of the derived analogues compound set over the CK1 δ and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the best posed analogue compounds. Lowercase letters represent the employed halogen group (a–g).

				CK1 δ						
		1	2	3	4	5	6	7		
Binding Energy	R0/R1	a	−5.1/−5.1	−7.7/−7.7	−5.2/−5.2	−5.3/−5.3	−5.6/−5.6	−5.6/−5.6	−5.3/−5.3	
MM/GBSA			−35.4499	−3.7149	−26.5327	−26.4630	−18.4901			
Binding Energy	R0/R1	b	−5.8/−5.8	−7.7/−7.7	−5.5/−5.5	−5.3/−5.3	−5.8/−5.8	−5.6/−5.6	−5.2/−5.2	
MM/GBSA			−24.0479	−30.2266	−21.2435	−6.4142	−11.6429			
Binding Energy	R0/R1	c	−5.6/−5.6	−7.6/−7.6	−4.8/−4.8	−5.7/−5.7	−5.2/−5.2	−5.1/−5.1	−5.6/−5.6	
MM/GBSA			−29.9803	−19.4546	−22.8644	−26.4159	−12.0871			
Binding Energy	R0/R1	d	−5.9/−5.9	−7.5/−7.5	−4.7/−4.7	−5.5/−5.5	−5.9/−5.9	−5.1/−5.1	−5.5/−5.5	
MM/GBSA			−19.8892	−25.5694						
Binding Energy	R0/R1	e	−6.1/−6.1	−7.5/−7.5	−5.4/−5.4	−5.3/−5.3	−5.1/−5.1	−5.3/−5.3	−5.3/−5.3	
MM/GBSA			−37.8982	−28.6573	−28.5831	−16.2323				
Binding Energy	R0/R1	f	−6.2/−6.2	−7.5/−7.5	−5.4/−5.4	−5.2/−5.2	−5.1/−5.1	−5.5/−5.5	−5.4/−5.4	
MM/GBSA			−34.6944	−22.6616	−26.3915	−13.6562	−15.4050			
Binding Energy	R0/R1	g	−6.1/−6.1	−7.3/−7.3	−4.8/−4.8	−5.2/−5.2	−5.1/−5.1	−5.6/−5.6	−5.5/−5.5	
MM/GBSA			−33.2393	−28.7631	−26.0731	−28.0238				

To avoid false positives, each docking calculation was performed twice (R0/R1). All energies values are kcal/mol.

Table A3. Summary of classical rigid docking calculations of the derived analogues compound set over the DYRK1A and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the best posed analogue compounds. Lowercase letters represent the employed halogen group (a–g).

				DYRK1A						
		1	2	3	4	5	6	7		
Binding Energy	R0/R1	a	−6.2/−6.2	−8.6/−8.6	−5.9/−5.9	−6.8/−6.8	−6.7/−6.7	−5.7/−5.7	−6/−6	
MM/GBSA			−37.8422	−15.2733	−30.7518	−31.2535	−18.9387	−20.8203		
Binding Energy	R0/R1	b	−6.5/−6.5	−8.5/−8.5	−7/−7	−7.2/−7.2	−6.6/−6.6	−7/−7	−7.3/−7.3	
MM/GBSA			−23.7829	−28.0642	−8.4887	−19.4730	−21.8802	−10.2503	−11.8981	
Binding Energy	R0/R1	c	−6.3/−6.3	−8.6/−8.6	−5.9/−5.9	−7.2/−7.2	−6.5/−6.5	−6.9/−6.9	−7.3/−7.3	
MM/GBSA			−30.2004	−34.1231	−12.1852	−26.5473	−28.3000	−18.7332	−13.2158	
Binding Energy	R0/R1	d	−5.9/−5.9	−8.6/−8.6	−6.7/−6.7	−6.6/−6.6	−6.3/−6.3	−6.5/−6.5	−6.6/−6.6	
MM/GBSA			−30.8597	−36.1125	−10.3667	−26.3748	−26.9602	−17.1653	−16.4510	
Binding Energy	R0/R1	e	−6.4/−6.4	−8.6/−8.6	−6.5/−6.5	−6.6/−6.6	−6.4/−6.4	−6.6/−6.6	−6.4/−6.4	
MM/GBSA			−32.8862	−13.5197	−28.9038	−29.6780			−20.0635	
Binding Energy	R0/R1	f	−6.2/−6.2	−8.6/−8.6	−5.8/−5.8	−6.8/−6.8	−6.3/−6.3	−6.4/−6.4	−6.9/−6.9	
MM/GBSA			−28.9419	−33.0823	−14.5027	−25.4579	−28.4081	−15.8583	−17.0209	
Binding Energy	R0/R1	g	−6.7/−6.7	−8.7/−8.7	−5.8/−5.8	−6.8/−6.8	−6.7/−6.7	−6.1/−6.1	−6.9/−6.9	
MM/GBSA			−30.3186	−35.5805	−13.4928	−27.1510	−29.6010	−18.0146	−17.3973	

To avoid false positives, each docking calculation was performed twice (R0/R1). All energies values are kcal/mol.

Table A4. Summary of classical rigid docking calculations of the derived analogues compound set over the CLK1 and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations of the best posed analogue compounds. Lowercase letters represent the employed halogen group (a–g).

			CLK1						
			1	2	3	4	5	6	7
Binding Energy	R0/R1	a	−7.5/−7.5	−8.8/−8.8	−6.4/−6.4	−6.8/−6.8	−6.2/−6.2	−6.4/−6.4	−6.9/−6.9
MM/GBSA			−29.8546	−34.1041	−15.2943	−26.7584	−25.8120	−27.7489	−24.0832
Binding Energy	R0/R1	b	−7.1/−7.1	−8.4/−8.4	−7/−7	−5.6/−5.6	−7.5/−7.5	−6.6/−6.6	−7.2/−7.2
MM/GBSA			−25.9089	−27.3711	−13.7919	−26.2326	−22.1539	−16.6984	−20.9149
Binding Energy	R0/R1	c	−7.6/−7.6	−9.1/−9.1	−7/−7	−7.3/−7.3	−6.4/−6.4	−6.6/−6.6	−6.4/−6.4
MM/GBSA			−34.0221	−16.2165	−24.6708	−29.4190	−25.8368	−20.9436	
Binding Energy	R0/R1	d	−7.6/−7.6	−8.4/−8.4	−6.8/−6.8	−7.8/−7.8	−6.3/−6.3	−6.9/−6.9	−7/−7
MM/GBSA			−26.9398	−30.7361	−25.6581	−25.1797	−19.3712	−17.5727	
Binding Energy	R0/R1	e	−6.8/−6.8	−8.9/−8.9	−6.9/−6.9	−7.5/−7.5	−5.9/−5.9	−7.1/−7.1	−7/−7
MM/GBSA			−30.0891	−16.2097	−28.3695	−28.0697	−27.0478	−17.4985	
Binding Energy	R0/R1	f	−7.5/−7.5	−8.6/−8.6	−6.7/−6.7	−7.5/−7.5	−6.2/−6.2	−6.4/−6.4	−6.7/−6.7
MM/GBSA			−28.1471	−20.4786	−23.9274	−27.3596	−15.4231	−21.2829	
Binding Energy	R0/R1	g	−7/−7	−8.9/−8.9	−6/−6	−7.4/−7.4	−6.8/−6.8	−6.4/−6.4	−6.8/−6.8
MM/GBSA			−30.3541	−33.9082	−16.3122	−25.0002	−30.7737	−25.4765	

To avoid false positives, each docking calculation was performed twice (R0/R1). All energies values are kcal/mol.

Table A5. Summary of absorption and distribution properties of the entire set of 49 derived compounds and the four brominated alkaloids convolutamine I (I) and J (J), kororamide A (A) and B (B).

Compound	Mol Weight	LogS	Absorption			Distribution				
			P-Glycoprotein	Caco-2 permeability	Intestinal Absorption	LogP	BBB	PPB	VDss	CNS Permeability
Compound 1a	331	−4.8	inactive	High	92.328	2.8	0.298	High	0.231	−1.832
Compound 1b	209.2	−3.4	inactive	Moderate	94.522	2.6	0.186	Medium	0.131	−1.888
Compound 1c	242.1	−4.3	inactive	High	92.462	2.9	0.279	High	0.37	−1.832
Compound 1d	270.1	−4	inactive	High	93.463	2.6	0.152	High	0.249	−1.866
Compound 1e	286.6	−4.5	inactive	High	92.395	2.8	0.278	High	0.385	−1.832
Compound 1f	270.1	−4	inactive	High	93.49	2.6	0.152	High	0.256	−1.87
Compound 1g	286.6	−4.5	inactive	High	92.395	2.8	0.278	High	0.385	−1.832
Compound 2a	351	−6.1	inactive	Moderate	90.067	4.1	0.477	High	0.234	−0.894
Compound 2b	229.2	−5.1	inactive	Moderate	92.006	3.4	0.539	High	−0.081	−0.946
Compound 2c	262.1	−5.9	inactive	Moderate	90.201	4.6	0.482	High	0.197	−0.894
Compound 2d	306.6	−6.1	inactive	Moderate	90.134	4.4	0.48	High	0.215	−0.894
Compound 2e	290.1	−5.7	inactive	Moderate	91.036	3.8	0.508	High	0.076	−0.92
Compound 2f	290.1	−5.7	inactive	Moderate	91.721	3.8	0.708	High	0.051	−1.339
Compound 2g	306.6	−6.1	inactive	Moderate	90.819	4.4	0.68	High	0.196	−1.313
Compound 3a	304	−4.7	inactive	High	89.848	2.8	0.227	High	0.95	−1.961
Compound 3b	182.2	−3.4	inactive	Moderate	91.82	2.7	0.375	Low	0.764	−2.017
Compound 3c	215.1	−4.2	inactive	High	89.982	2.9	0.23	High	0.919	−1.961
Compound 3d	243.1	−4	inactive	High	90.899	2.7	0.222	Medium	0.868	−1.999
Compound 3e	259.5	−4.4	inactive	High	89.915	2.8	0.228	High	0.934	−1.961
Compound 3f	243.1	−4	inactive	High	90.872	2.7	0.222	Medium	0.845	−1.995
Compound 3g	259.5	−4.4	inactive	High	89.915	2.8	0.228	High	0.934	−1.961
Compound 4a	289	−4.9	inactive	High	91.487	3.3	0.351	High	0.432	−1.66
Compound 4b	167.2	−3.4	inactive	Moderate	93.459	3.2	0.437	Medium	0.248	−1.715
Compound 4c	200.1	−4.5	inactive	High	91.621	3.6	0.357	High	0.401	−1.66
Compound 4d	228.1	−4.1	inactive	High	92.538	3.2	0.382	Medium	0.344	−1.697
Compound 4e	244.5	−4.7	inactive	High	91.554	3.5	0.354	High	0.416	−1.66
Compound 4f	228.1	−4.1	inactive	High	92.511	3.2	0.382	Medium	0.32	−1.693
Compound 4g	244.5	−4.7	inactive	High	91.554	3.5	0.354	High	0.416	−1.66
Compound 5a	305	−4.4	inactive	High	89.763	3	0.284	High	0.253	−1.98
Compound 5b	183.2	−2.9	inactive	Moderate	91.734	2.6	0.432	Low	0.086	−2.036
Compound 5c	216.1	−3.6	inactive	High	89.897	2.7	0.287	High	0.223	−1.98
Compound 5d	244.1	−3.5	inactive	High	90.814	2.7	0.279	Medium	0.169	−2.017
Compound 5e	260.5	−3.9	inactive	High	89.83	2.9	0.286	High	0.238	−1.98
Compound 5f	244.1	−3.5	inactive	High	90.786	2.7	0.279	Medium	0.143	−2.014
Compound 5g	260.5	−3.9	inactive	High	89.83	2.9	0.286	High	0.238	−1.98
Compound 6a	318	−4.7	inactive	High	90.757	2.7	0.146	High	1.061	−1.917
Compound 6b	196.2	−3.2	inactive	Moderate	92.728	2.7	0.293	Low	0.867	−1.973
Compound 6c	229.1	−4.3	inactive	High	90.891	3	0.148	Medium	1.031	−1.917
Compound 6d	257.1	−3.9	inactive	High	91.78	2.6	0.14	Medium	0.956	−1.951
Compound 6e	273.6	−4.4	inactive	High	90.824	2.8	0.147	Medium	1.046	−1.917
Compound 6f	257.1	−3.9	inactive	High	91.808	2.6	0.14	Medium	0.978	−1.954
Compound 6g	273.6	−4.4	inactive	High	90.824	2.8	0.147	Medium	1.046	−1.917

Table A5. Cont.

Compound	Absorption					Distribution				
	Mol Weight	LogS	P-Glycoprotein	Caco-2 permeability	Intestinal Absorption	LogP	BBB	PPB	VDss	CNS Permeability
Compound 7a	332	-4.9	inactive	High	92.246	3	0.191	High	1.141	-1.487
Compound 7b	210.2	-3.4	inactive	Moderate	94.218	2.9	0.349	Low	0.994	-1.543
Compound 7c	243.1	-4.6	inactive	High	92.38	3.3	0.225	Medium	1.11	-1.487
Compound 7d	271.1	-4.1	inactive	High	93.297	2.9	0.258	Medium	1.084	-1.525
Compound 7e	287.6	-4.7	inactive	High	92.313	3.2	0.208	High	1.125	-1.487
Compound 7f	271.1	-4.1	inactive	High	93.27	2.9	0.258	Medium	1.055	-1.521
Compound 7g	287.6	-4.7	inactive	High	92.313	3.2	0.208	High	1.125	-1.487
J	470	-4.4	inactive	Moderate	90.483	4.4	0.386	High	0.868	-2.215
I	473	-4.3	inactive	Moderate	91.515	3.9	0.193	High	1.474	-2.024
A	534.1	-4.3	inactive	Low	90.979	4.6	0.316	Low	1.112	-2.449
B	535.1	-3.9	inactive	Moderate	100	3.4	0.184	High	0.002	-2.93

BBB: blood brain barrier, PPB: protein-protein binding, VDss: steady state volume of distribution, CNS: central nervous system.

Table A6. Summary of metabolism, excretion and toxicity properties of the entire set of 49 derived compounds and the four brominated alkaloids convolutamine I (I) and J (J), kororamide A (A) and B (B).

Compound	Metabolism	Excretion	Toxicity			
	CYP450	OCT2 Substrate	hERG	MRTD	AMES Toxicity	Hepatotoxicity
Compound 1a	Yes	No	<4.0	0.482	No	No
Compound 1b	Yes	No	<4.0	0.666	No	No
Compound 1c	Yes	No	<4.0	0.503	No	No
Compound 1d	Yes	No	<4.0	0.45	No	No
Compound 1e	Yes	No	<4.0	0.492	No	No
Compound 1f	Yes	No	<4.0	0.574	No	No
Compound 1g	Yes	No	<4.0	0.492	No	No
Compound 2a	Yes	No	<4.0	0.673	Yes	No
Compound 2b	Yes	No	<4.0	0.608	Yes	No
Compound 2c	Yes	No	<4.0	0.671	Yes	No
Compound 2d	Yes	No	<4.0	0.672	Yes	No
Compound 2e	Yes	No	<4.0	0.641	Yes	No
Compound 2f	Yes	No	<4.0	0.585	Yes	No
Compound 2g	Yes	No	<4.0	0.616	Yes	No
Compound 3a	Yes	No	<4.0	0.381	No	No
Compound 3b	Yes	No	<4.0	0.512	No	No
Compound 3c	Yes	No	<4.0	0.402	No	No
Compound 3d	Yes	No	<4.0	0.455	No	No
Compound 3e	Yes	No	<4.0	0.391	No	No
Compound 3f	Yes	No	<4.0	0.303	No	No
Compound 3g	Yes	No	<4.0	0.391	No	No
Compound 4a	Yes	No	<4.0	0.525	No	No
Compound 4b	Yes	No	<4.0	0.716	No	No
Compound 4c	Yes	No	<4.0	0.544	No	No
Compound 4d	Yes	No	<4.0	0.625	No	No
Compound 4e	Yes	No	<4.0	0.534	No	No
Compound 4f	Yes	No	<4.0	0.471	No	No
Compound 4g	Yes	No	<4.0	0.534	No	No
Compound 5a	Yes	No	<4.0	0.55	No	No
Compound 5b	Yes	No	<4.0	0.678	No	No
Compound 5c	Yes	No	<4.0	0.572	No	No
Compound 5d	Yes	No	<4.0	0.627	No	No
Compound 5e	Yes	No	<4.0	0.561	No	No
Compound 5f	Yes	No	<4.0	0.47	No	No
Compound 5g	Yes	No	<4.0	0.561	No	No
Compound 6a	Yes	No	<4.0	0.376	No	No
Compound 6b	Yes	No	<4.0	0.502	No	No
Compound 6c	Yes	No	<4.0	0.397	No	No
Compound 6d	Yes	No	<4.0	0.293	No	No
Compound 6e	Yes	No	<4.0	0.387	No	No
Compound 6f	Yes	No	<4.0	0.441	No	No

Table A6. Cont.

Compound	Metabolism	Excretion	Toxicity			
	CYP450	OCT2 Substrate	hERG	MRTD	AMES Toxicity	Hepatotoxicity
Compound 6g	Yes	No	<4.0	0.387	No	No
Compound 7a	Yes	No	<4.0	0.2	No	No
Compound 7b	Yes	No	<4.0	0.311	No	No
Compound 7c	Yes	No	<4.0	0.219	No	No
Compound 7d	Yes	No	<4.0	0.259	No	No
Compound 7e	Yes	No	<4.0	0.209	No	No
Compound 7f	Yes	No	<4.0	0.117	No	No
Compound 7g	Yes	No	<4.0	0.209	No	No
I	Yes	Yes	<4.0	0.029	No	Yes
J	Yes	Yes	<4.0	−0.814	No	Yes
A	Yes	No	<4.0	−0.599	No	Yes
B	No	No	<4.0	0.405	Yes	No

CYP: cytochrome, OCT2: organic cation transporter 2, hERG: human ether-a-go-go gene, MRTD: maximum recommended tolerated dose.

References

- Manning, G.; Whyte, D.B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* **2002**, *298*, 1912–1934. [[CrossRef](#)] [[PubMed](#)]
- Kolarova, M.; Garcia-Sierra, F.; Bartos, A.; Ricny, J.; Ripova, D.; Ripova, D. Structure and pathology of tau protein in Alzheimer disease. *Int. J. Alzheimers Dis.* **2012**, *2012*, 1–13. [[CrossRef](#)] [[PubMed](#)]
- Martin, L.; Latypova, X.; Wilson, C.M.; Magnaudeix, A.; Perrin, M.-L.; Yardin, C.; Terro, F. Tau protein kinases: Involvement in Alzheimer's disease. *Ageing Res. Rev.* **2013**, *12*, 289–309. [[CrossRef](#)] [[PubMed](#)]
- Citron, M. Alzheimer's disease: Strategies for disease modification. *Nat. Rev. Drug Discov.* **2010**, *9*, 387–398. [[CrossRef](#)] [[PubMed](#)]
- Tell, V.; Hilgeroth, A. Recent developments of protein kinase inhibitors as potential AD therapeutics. *Front Cell Neurosci.* **2013**, *7*. [[CrossRef](#)] [[PubMed](#)]
- Dolan, P.J.; Johnson, G.V.W. The role of tau kinases in Alzheimer's disease. *Curr. Opin. Drug Discov. Devel.* **2010**, *13*, 595–603. [[PubMed](#)]
- Stotani, S.; Giordanetto, F.; Medda, F. DYRK1A inhibition as potential treatment for Alzheimer's disease. *Future Med. Chem.* **2016**, *8*, 681–696. [[CrossRef](#)] [[PubMed](#)]
- Branca, C.; Shaw, D.M.; Belfiore, R.; Gokhale, V.; Shaw, A.Y.; Foley, C.; Smith, B.; Hulme, C.; Dunckley, T.; Meechoovet, B.; et al. Dyrk1 inhibition improves Alzheimer's disease-like pathology. *Aging Cell* **2017**, *16*, 1146–1154. [[CrossRef](#)] [[PubMed](#)]
- Li, G.; Yin, H.; Kuret, J. Casein kinase 1 delta phosphorylates tau and disrupts its binding to microtubules. *J. Biol. Chem.* **2004**, *279*, 15938–15945. [[CrossRef](#)] [[PubMed](#)]
- Hooper, C.; Killick, R.; Lovestone, S. The GSK3 hypothesis of Alzheimer's disease. *J. Neurochem.* **2008**, *104*, 1433–1439. [[CrossRef](#)] [[PubMed](#)]
- Llorens-Martín, M.; Jurado, J.; Hernández, F.; Avila, J. GSK-3 β , a pivotal kinase in Alzheimer disease. *Front. Mol. Neurosci.* **2014**, *7*, 46. [[CrossRef](#)] [[PubMed](#)]
- Hernández, F.; Gómez de Barreda, E.; Fuster-Matanzo, A.; Lucas, J.J.; Avila, J. GSK3: A possible link between beta amyloid peptide and tau protein. *Exp. Neurol.* **2010**, *223*, 322–325. [[CrossRef](#)] [[PubMed](#)]
- Hernandez, F.; Lucas, J.J.; Avila, J. GSK3 and Tau: Two Convergence Points in Alzheimer's Disease. *J. Alzheimer's Dis.* **2012**, *33*, S141–S144. [[CrossRef](#)] [[PubMed](#)]
- Ishizawa, T.; Sahara, N.; Ishiguro, K.; Kersh, J.; McGowan, E.; Lewis, J.; Hutton, M.; Dickson, D.W.; Yen, S.-H. Co-localization of glycogen synthase kinase-3 with neurofibrillary tangles and granulovacuolar degeneration in transgenic mice. *Am. J. Pathol.* **2003**, *163*, 1057–1067. [[CrossRef](#)]
- Ryoo, S.-R.; Kyeong Jeong, H.; Radnaabazar, C.; Yoo, J.-J.; Cho, H.-J.; Lee, H.-W.; Kim, I.-S.; Cheon, Y.-H.; Ahn, Y.S.; Chung, S.-H.; et al. DYRK1A-mediated hyperphosphorylation of tau a functional link between down syndrome and Alzheimer disease. *J. Biol. Chem.* **2007**, *282*, 34850–34857. [[CrossRef](#)] [[PubMed](#)]
- Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on natural products for drug design. *Nat. Chem.* **2016**, *8*, 531–541. [[CrossRef](#)] [[PubMed](#)]

17. Li, J.W.; Vederas, J.C. Drug discovery and natural products: End of an era or an endless frontier? *Science* **2009**, *325*, 161–165. [[CrossRef](#)] [[PubMed](#)]
18. Harvey, A.L.; Edrada-Ebel, R.; Quinn, R.J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* **2015**, *14*, 111–129. [[CrossRef](#)] [[PubMed](#)]
19. Patridge, E.; Gareiss, P.; Kinch, M.S.; Hoyer, D. An analysis of FDA-approved drugs: Natural products and their derivatives. *Drug Discov. Today* **2016**, *21*, 204–207. [[CrossRef](#)] [[PubMed](#)]
20. Bharate, S.S.; Mignani, S.; Vishwakarma, R.A. Why are the majority of active compounds in the CNS domain natural products? A critical analysis. *J. Med. Chem.* **2018**. [[CrossRef](#)] [[PubMed](#)]
21. Russo, P.; Kisialiou, A.; Lamonaca, P.; Moroni, R.; Prinzi, G.; Fini, M. New drugs from marine organisms in Alzheimer's disease. *Mar. Drugs* **2015**, *14*, 5. [[CrossRef](#)] [[PubMed](#)]
22. Ansari, N.; Khodaghali, F. Natural products as promising drug candidates for the treatment of Alzheimer's disease: Molecular mechanism aspect. *Curr. Neuropharmacol.* **2013**, *11*, 414–429. [[CrossRef](#)] [[PubMed](#)]
23. Dey, A.; Bhattacharya, R.; Mukherjee, A.; Pandey, D.K. Natural products against Alzheimer's disease: Pharmacotherapeutics and biotechnological interventions. *Biotechnol. Adv.* **2017**, *35*, 178–216. [[CrossRef](#)] [[PubMed](#)]
24. Bharate, S.B.; Yadav, R.R.; Battula, S.; Vishwakarma, R.A. Meridianins: Marine-derived potent kinase inhibitors. *Mini-Rev. Med. Chem.* **2012**, *12*, 618–631. [[CrossRef](#)] [[PubMed](#)]
25. Meine, R.; Becker, W.; Falke, H.; Preu, L.; Loaëc, N.; Meijer, L.; Kunick, C. Indole-3-carbonitriles as DYRK1A inhibitors by fragment-based drug design. *Molecules* **2018**, *23*, 64. [[CrossRef](#)] [[PubMed](#)]
26. Tahtouh, T.; Elkins, J.M.; Filippakopoulos, P.; Soundararajan, M.; Burgy, G.; Durieu, E.; Cochet, C.; Schmid, R.S.; Lo, D.C.; Delhommel, F.; et al. Selectivity, cocrystal structures, and neuroprotective properties of leucettines, a family of protein kinase inhibitors derived from the marine sponge alkaloid leucettamine B. *J. Med. Chem.* **2012**, *55*, 9312–9330. [[CrossRef](#)] [[PubMed](#)]
27. Salado, I.G.; Redondo, M.; Bello, M.L.; Perez, C.; Liachko, N.F.; Kraemer, B.C.; Miguel, L.; Lecourtois, M.; Gil, C.; Martinez, A.; et al. Protein kinase CK-1 inhibitors as new potential drugs for amyotrophic lateral sclerosis. *J. Med. Chem.* **2014**, *57*, 2755–2772. [[CrossRef](#)] [[PubMed](#)]
28. Eldar-Finkelman, H.; Martinez, A. GSK-3 inhibitors: Preclinical and clinical focus on CNS. *Front. Mol. Neurosci.* **2011**, *4*, 32. [[CrossRef](#)] [[PubMed](#)]
29. Giraud, F.; Alves, G.; Debiton, E.; Nauton, L.; Thery, V.; Durieu, E.; Ferandin, Y.; Lozach, O.; Meijer, L.; Anizon, F.; et al. Synthesis, protein kinase inhibitory potencies, and in vitro antiproliferative activities of meridianin derivatives. *J. Med. Chem.* **2011**, *54*, 4474–4489. [[CrossRef](#)] [[PubMed](#)]
30. Jain, P.; Karthikeyan, C.; Moorthy, N.S.; Waiker, D.; Jain, A.; Trivedi, P. Human CDC2-like kinase 1 (CLK1): A novel target for Alzheimer's disease. *Curr. Drug Targets* **2014**, *15*, 539–550. [[CrossRef](#)] [[PubMed](#)]
31. Fedorov, O.; Huber, K.; Eisenreich, A.; Filippakopoulos, P.; King, O.; Bullock, A.N.; Szklarczyk, D.; Jensen, L.J.; Fabbro, D.; Trappe, J.; et al. Specific CLK inhibitors from a novel chemotype for regulation of alternative splicing. *Chem. Biol.* **2011**, *18*, 67–76. [[CrossRef](#)] [[PubMed](#)]
32. Halekotte, J.; Witt, L.; Ianes, C.; Krüger, M.; Bührmann, M.; Rauh, D.; Pichlo, C.; Brunstein, E.; Luxenburger, A.; Baumann, U.; et al. Optimized 4,5-diarylimidazoles as potent/selective inhibitors of protein kinase CK1 δ and their structural relation to p38 α MAPK. *Molecules* **2017**, *22*, 522. [[CrossRef](#)] [[PubMed](#)]
33. Cozza, G.; Gianoncelli, A.; Montopoli, M.; Caparrotta, L.; Venerando, A.; Meggio, F.; Pinna, L.A.; Zagotto, G.; Moro, S. Identification of novel protein kinase CK1 delta (CK1 δ) inhibitors through structure-based virtual screening. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5672–5675. [[CrossRef](#)] [[PubMed](#)]
34. Yadav, R.R.; Sharma, S.; Joshi, P.; Wani, A.; Vishwakarma, R.A.; Kumar, A.; Bharate, S.B. Meridianin derivatives as potent Dyrk1A inhibitors and neuroprotective agents. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 2948–2952. [[CrossRef](#)] [[PubMed](#)]
35. Feng, L.; Geisselbrecht, Y.; Blanck, S.; Wilbuer, A.; Atilla-Gokcumen, G.E.; Filippakopoulos, P.; Kräling, K.; Celik, M.A.; Harms, K.; Maksimoska, J.; et al. Structurally sophisticated octahedral metal complexes as highly selective protein kinase inhibitors. *J. Am. Chem. Soc.* **2011**, *133*, 5976–5986. [[CrossRef](#)] [[PubMed](#)]
36. Mente, S.; Arnold, E.; Butler, T.; Chakrapani, S.; Chandrasekaran, R.; Cherry, K.; Dirico, K.; Doran, A.; Fisher, K.; Galatsis, P.; et al. Ligand-protein interactions of selective casein kinase 1 delta inhibitors. *J. Med. Chem.* **2013**. [[CrossRef](#)] [[PubMed](#)]

37. Llorach-Pares, L.; Nonell-Canals, A.; Sanchez-Martinez, M.; Avila, C. Computer-aided drug design applied to marine drug discovery: Meridianins as Alzheimer's disease therapeutic agents. *Mar. Drugs* **2017**, *15*, 366. [CrossRef] [PubMed]
38. Dashti, Y.; Vial, M.-L.; Wood, S.A.; Mellick, G.D.; Roullier, C.; Quinn, R.J. Kororamide B, a brominated alkaloid from the bryozoan *Amathia tortuosa* and its effects on Parkinson's disease cells. *Tetrahedron* **2015**, *71*, 7879–7884. [CrossRef]
39. Netz, N.; Opatz, T. Marine indole alkaloids. *Mar. Drugs* **2015**, *13*, 4814–4914. [CrossRef] [PubMed]
40. Blunt, J.W.; Copp, B.R.; Keyzers, R.A.; Munro, M.H.G.; Prinsep, M.R. Marine natural products. *Nat. Prod. Rep.* **2017**, *34*, 235–294. [CrossRef] [PubMed]
41. Gul, W.; Hamann, M.T. Indole alkaloid marine natural products: An established source of cancer drug leads with considerable promise for the control of parasitic, neurological and other diseases. *Life Sci.* **2005**, *78*, 442–453. [CrossRef] [PubMed]
42. Gribble, G.W. Biological activity of recently discovered halogenated marine natural products. *Mar. Drugs* **2015**, *13*, 4044–4136. [CrossRef] [PubMed]
43. Pauletti, P.M.; Cintra, L.S.; Braguine, C.G.; da Silva Filho, A.A.; Silva, M.L.A.E.; Cunha, W.R.; Januário, A.H. Halogenated indole alkaloids from marine invertebrates. *Mar. Drugs* **2010**, *8*, 1526–1549. [CrossRef] [PubMed]
44. Munro, M.H.G.; Blunt, J.W. MarinLit: A database of the marine natural products literature. Available online: <http://pubs.rsc.org/marinlit/> (accessed on 16 October 2018).
45. Acharya, C.; Coop, A.; Polli, J.E.; Mackerell, A.D., Jr. Recent advances in ligand-based drug design: Relevance and utility of the conformationally sampled pharmacophore approach. *Curr. Comput. Aid. Drug Des.* **2011**, *7*, 10–22. [CrossRef]
46. Traxler, P.; Furet, P. Strategies toward the design of novel and selective protein tyrosine kinase inhibitors. *Pharmacol. Ther.* **1999**, *82*, 195–206. [CrossRef]
47. McGregor, M.J. A pharmacophore map of small molecule protein kinase inhibitors. *J. Chem. Inf. Model.* **2007**, *47*, 2374–2382. [CrossRef] [PubMed]
48. Huang, D.; Zhou, T.; Lafleur, K.; Nevado, C.; Caffisch, A. Kinase selectivity potential for inhibitors targeting the ATP binding site: A network analysis. *Bioinformatics* **2010**, *26*, 198–204. [CrossRef] [PubMed]
49. Jarhad, D.B.; Mashelkar, K.K.; Kim, H.-R.; Noh, M.; Jeong, L.S. Dual-specificity tyrosine phosphorylation-regulated kinase 1A (DYRK1A) inhibitors as potential therapeutics. *J. Med. Chem.* **2018**. [CrossRef] [PubMed]
50. Klein-Junior, L.; Santos Passos, C.; Moraes, A.; Wakui, V.; Konrath, E.; Nurisso, A.; Carrupt, P.-A.; Alves de Oliveira, C.; Kato, L.; Henriques, A. Indole alkaloids and semisynthetic indole derivatives as multifunctional scaffolds aiming the inhibition of enzymes related to neurodegenerative diseases—A focus on psychotria L. Genus. *Curr. Top. Med. Chem.* **2014**, *14*, 1056–1075. [CrossRef]
51. Ren, P.; Liu, Y.; Li, L.; Chan, K.; Wilson, T.E. Heterocyclic Kinase Inhibitors. US Patent Application No. US20130095100A1, 18 April 2013. U.S. Patent Application No. US12331431, 9 December 2008.
52. Almstetter, M.T.; Trembl, A.; Koestler, R.; Yehia, N. Heterocyclic Compounds as Kinase Inhibitors. EP2699572B1, 10 August 2016.
53. Castro, A.C.; Chan, K.; Evans, C.A.; Janardanannair, S.; Lescarbeau, A.; Li, L.; Liu, T.; Liu, Y.; Ren, P.; Snyder, D.A.; et al. Heterocyclic Compounds and Uses Thereof. WO201315487A1, 17 October 2013.
54. An, W.F.; Germain, A.R.; Bishop, J.A.; Nag, P.P.; Metkar, S.; Ketterman, J.; Walk, M.; Weiwer, M.; Liu, X.; Patnaik, D.; et al. Discovery of potent and highly selective inhibitors of GSK3b. In *Probe Reports from the NIH Molecular Libraries Program*; NCBI: Bethesda, MD, USA, 2010.
55. Benek, O.; Hroch, L.; Aitken, L.; Gunn-Moore, F.; Vinklarova, L.; Kuca, K.; Perez, D.I.; Perez, C.; Martinez, A.; Fisar, Z.; et al. 1-(Benzo[d]thiazol-2-yl)-3-phenylureas as dual inhibitors of casein kinase 1 and ABAD enzymes for treatment of neurodegenerative disorders. *J. Enzyme Inhib. Med. Chem.* **2018**, *33*, 665–670. [CrossRef] [PubMed]
56. Coombs, T.C.; Tanega, C.; Shen, M.; Wang, J.L.; Auld, D.S.; Gerritz, S.W.; Schoenen, F.J.; Thomas, C.J.; Aubé, J. Small-molecule pyrimidine inhibitors of the cdc2-like (Clk) and dual specificity tyrosine phosphorylation-regulated (Dyrk) kinases: Development of chemical probe ML315. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 3654–3661. [CrossRef] [PubMed]
57. Fenical, W. Natural halogenated organics. *Elsevier Oceanogr. Ser.* **1981**, *31*, 375–393. [CrossRef]

58. Hernandez, M.; Cavalcanti, S.M.; Moreira, D.R.; de Azevedo Junior, W.; Leite, A.C. Halogen atoms in the modern medicinal chemistry: Hints for the drug design. *Curr. Drug Targets* **2010**, *11*, 303–314. [CrossRef] [PubMed]
59. Sirimulla, S.; Bailey, J.B.; Vegesna, R.; Narayan, M. Halogen interactions in protein–ligand complexes: Implications of halogen bonding for rational drug design. *J. Chem. Inf. Model.* **2013**, *53*, 2781–2791. [CrossRef] [PubMed]
60. Metrangolo, P.; Resnati, G. Chemistry. Halogen versus hydrogen. *Science* **2008**, *321*, 918–919. [CrossRef] [PubMed]
61. Kramer, T.; Schmidt, B.; Lo Monte, F. Small-molecule inhibitors of GSK-3: Structural insights and Their application to Alzheimer’s disease models. *Int. J. Alzheimers Dis.* **2012**, *2012*, 381029. [CrossRef] [PubMed]
62. Wei, W.; Bodles-Brakhop, A.M.; Barger, S.W. A role for P-glycoprotein in clearance of Alzheimer amyloid β -peptide from the brain HHS public access. *Curr. Alzheimer Res.* **2016**, *13*, 615–620.
63. Miller, D.S.; Bauer, B.; Hartz, A.M.S. Modulation of P-glycoprotein at the blood-brain barrier: Opportunities to improve central nervous system pharmacotherapy. *Pharmacol. Rev.* **2008**, *60*, 196–209. [CrossRef] [PubMed]
64. Chang, K.L.; Pee, H.N.; Yang, S.; Ho, P.C. Influence of drug transporters and stereoselectivity on the brain penetration of pioglitazone as a potential medicine against Alzheimer’s disease. *Sci. Rep.* **2015**, *5*, 9000. [CrossRef] [PubMed]
65. Cirrito, J.R.; Deane, R.; Fagan, A.M.; Spinner, M.L.; Parsadanian, M.; Finn, M.B.; Jiang, H.; Prior, J.L.; Sagare, A.; Bales, K.R.; et al. P-glycoprotein deficiency at the blood-brain barrier increases amyloid-deposition in an Alzheimer disease mouse model. *J. Clin. Invest.* **2005**, *115*, 3285–3290. [CrossRef] [PubMed]
66. Pires, D.E.V.; Blundell, T.L.; Ascher, D.B. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J. Med. Chem.* **2015**, *58*, 4066–4072. [CrossRef] [PubMed]
67. Lynch, T.; Price, A. The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. *Am. Fam. Physician* **2007**, *76*, 391–396. [CrossRef] [PubMed]
68. Bibi, Z. Role of cytochrome P450 in drug interactions. *Nutr. Metab.* **2008**, *5*, 27. [CrossRef] [PubMed]
69. Yang, H.; Sun, L.; Li, W.; Liu, G.; Tang, Y. In silico prediction of chemical toxicity for drug design using machine learning methods and structural alerts. *Front. Chem.* **2018**, *6*, 1–12. [CrossRef]
70. Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2012**, *64*, 4–17. [CrossRef]
71. Berman, H.; Henrick, K.; Nakamura, H.; Markley, J.L. The worldwide protein data bank (wwPDB): Ensuring a single, uniform archive of PDB data. *Nucleic Acids Res.* **2007**, *35*, 2006–2008. [CrossRef] [PubMed]
72. Asdsss Felix, E.; Santamaria-Navarro, E.; Sanchez-Martinez, M.; Nonell-Canals, A. Itzamna. Available online: <https://www.mindthebyte.com/> (accessed on 15 October 2018).
73. Yu, W.; Mackerell, A.D. Computer-aided drug design methods. *Methods Mol. Biol.* **2017**, *1520*, 93–94. [CrossRef]
74. Ganesan, A.; Coote, M.L.; Barakat, K. Molecular dynamics-driven drug discovery: Leaping forward with confidence. *Drug Discov. Today* **2017**, *22*. [CrossRef] [PubMed]
75. Lill, M.A. Efficient incorporation of protein flexibility and dynamics into molecular docking simulations. *Biochemistry* **2011**, *50*, 6157–6169. [CrossRef] [PubMed]
76. Alonso, H.; Bliznyuk, A.A.; Gready, J.E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* **2006**, *26*, 531–568. [CrossRef] [PubMed]
77. de Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. The role of molecular dynamics and related methods in drug discovery. *J. Med. Chem.* **2016**. [CrossRef] [PubMed]
78. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [CrossRef] [PubMed]
79. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950–1958. [CrossRef] [PubMed]
80. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [CrossRef] [PubMed]

81. Martín-García, F.; Papaleo, E.; Gomez-Puertas, P.; Boomsma, W.; Lindorff-Larsen, K. Comparing molecular dynamics force fields in the essential subspace. *PLoS ONE* **2015**, *10*, e0121114. [[CrossRef](#)] [[PubMed](#)]
82. Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M.P.; Dror, R.O.; Shaw, D.E. Systematic validation of protein force fields against experimental data. *PLoS ONE* **2012**, *7*, 1–6. [[CrossRef](#)] [[PubMed](#)]
83. Wang, J.; Wang, W.; Kollman, P.A.; Case, D.A. Antechamber, an accessory software package for molecular mechanical calculations. *J. Chem. Inf. Comp. Sci.* **2005**, *25*, 25–1174.
84. Case, D.A.; Cheatham, T.E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688. [[CrossRef](#)] [[PubMed](#)]
85. Jorgensen, W.L.; Jenson, C. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. *J. Comput. Chem.* **1998**, *19*, 1179–1186. [[CrossRef](#)]
86. Andersen, H.C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983**, *52*, 24–34. [[CrossRef](#)]
87. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, *0441*, 1–13. [[CrossRef](#)] [[PubMed](#)]
88. Rastelli, G.; Degliesposti, G.; Del Rio, A.; Sgobba, M. Binding estimation after refinement, a new automated procedure for the refinement and rescoring of docked ligands in virtual screening. *Chem. Biol. Drug Des.* **2009**, *73*, 283–286. [[CrossRef](#)] [[PubMed](#)]
89. Miller, B.R., III; McGee, T.D., Jr.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321. [[CrossRef](#)] [[PubMed](#)]
90. Mai, S.L.; Binh, K.M. Steered molecular dynamics-A promising tool for drug design. *Curr. Bioinform.* **2012**, *7*, 342–351. [[CrossRef](#)]
91. Patel, J.S.; Berteotti, A.; Ronsisvalle, S.; Rocchia, W.; Cavalli, A. Steered molecular dynamics simulations for studying protein–Ligand interaction in cyclin-dependent kinase 5. *J. Chem. Inf. Model.* **2014**, *54*, 470–480. [[CrossRef](#)] [[PubMed](#)]
92. Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wriggers, W.; Schulten, K. Steered molecular dynamics. In *Computational Molecular Dynamics: Challenges, Methods, Ideas*; Deuffhard, P., Hermans, J., Eds.; Springer: Berlin/Heidelberg, Germany, 1999; pp. 39–65.
93. Isralewitz, B.; Gao, M.; Schulten, K. Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.* **2001**, *11*, 224–230. [[CrossRef](#)]
94. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [[CrossRef](#)]
95. Bento, A.P.; Gaulton, A.; Hersey, A.; Bellis, L.J.; Chambers, J.; Davies, M.; Krüger, F.A.; Light, Y.; Mak, L.; McGlinchey, S.; et al. The ChEMBL bioactivity database: An update. *Nucleic Acids Res.* **2014**, *42*, 1083–1090. [[CrossRef](#)] [[PubMed](#)]
96. The UniProt consortium update on activities at the universal protein resource (UniProt) in 2013. *Nucleic Acids Res.* **2013**, *41*, D43–D47. [[CrossRef](#)]
97. Vidal, D.; Nonell-Canals, A. ADMETer. Available online: <https://www.mindthebyte.com/> (accessed on 15 October 2018).
98. Poongavanam, V.; Haider, N.; Ecker, G.F. Fingerprint-based in silico models for the prediction of P-glycoprotein substrates and inhibitors. *Bioorg. Med. Chem.* **2012**, *20*, 5388–5395. [[CrossRef](#)] [[PubMed](#)]
99. Sedykh, A.; Fourches, D.; Duan, J.; Hucke, O.; Garneau, M.; Zhu, H.; Bonneau, P.; Tropsha, A. Human intestinal transporter database: QSAR modeling and virtual profiling of drug uptake, efflux and interactions. *Pharm. Res.* **2013**, *30*, 996–1007. [[CrossRef](#)] [[PubMed](#)]
100. Yuan, S.; Chan, H.C.S.; Hu, Z. Using PyMOL as a platform for computational drug design. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, e1298. [[CrossRef](#)]
101. Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: Fully automated protein–ligand interaction profiler. *Nucleic Acids Res.* **2015**, *43*, W443–W447. [[CrossRef](#)] [[PubMed](#)]
102. RDKit: Open-source cheminformatics. Available online: <http://rdkit.org/> (accessed on 15 October 2018).
103. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [[CrossRef](#)]
104. Waskom, M.; Botvinnik, O.; O’Kane, D.; Hobson, P.; Ostblom, J.; Lukauskas, S.; Gempert, D.C.; Augspurger, T.; Halchenko, Y.; Cole, J.B.; et al. Mwaskom/Seaborn: V0.9.0 (July 2018). Zenodo. Available online: <http://doi.org/10.5281/zenodo.1313201> (accessed on 15 October 2018).

105. Montaser, R.; Luesch, H. Marine natural products: A new wave of drugs? *Future Med. Chem.* **2011**, *3*, 1475–1489. [[CrossRef](#)] [[PubMed](#)]
106. Kiuru, P.; D’Auria, M.; Muller, C.; Tammela, P.; Vuorela, H.; Yli-Kauhaluoma, J. Exploring marine resources for bioactive compounds. *Planta Med.* **2014**, *80*, 1234–1246. [[CrossRef](#)] [[PubMed](#)]
107. Grosso, C.; Valentão, P.; Ferreres, F.; Andrade, P.B. Review: Bioactive marine drugs and marine biomaterials for brain diseases. *Mar. Drugs* **2014**, *12*, 2539–2589. [[CrossRef](#)] [[PubMed](#)]
108. Martins, A.; Vieira, H.; Gaspar, H.; Santos, S. Marketed marine natural products in the pharmaceutical and cosmeceutical industries: Tips for success. *Mar. Drugs* **2014**, *12*, 1066–1101. [[CrossRef](#)] [[PubMed](#)]
109. Molinski, T.F.; Dalisay, D.S.; Lievens, S.L.; Saludes, J.P. Drug development from marine natural products. *Nat. Rev. Drug Discov.* **2009**, *8*, 69–85. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Appendix II

Other publications not related to this thesis. The first page is included below in their original format



Metformin directly targets the H3K27me3 demethylase KDM6A/UTX

Elisabet Cuyàs^{1,2,*} | Sara Verdura^{1,2,*} | Laura Llorach-Pares^{3,*} | Salvador Fernández-Arroyo⁴ | Fedra Luciano-Mateo⁴ | Noemí Cabré⁴ | Jan Stursa^{5,6} | Lukas Werner^{5,6} | Begoña Martín-Castillo^{7,†} | Benoit Viollet^{8,9,10} | Jiri Neuzil^{6,11} | Jorge Joven⁴ | Alfons Nonell-Canals³ | Melchor Sanchez-Martinez³ | Javier A. Menendez^{1,2,†}

¹ProCURE (Program Against Cancer Therapeutic Resistance), Metabolism & Cancer Group, Catalan Institute of Oncology, Girona, Catalonia, Spain

²Girona Biomedical Research Institute (IDIBGI), Girona, Spain

³Mind the Byte, Barcelona, Spain

⁴Unitat de Recerca Biomèdica, Hospital Universitari de Sant Joan, IISPV, Rovira i Virgili University, Reus, Spain

⁵Institute of Chemical Technology, Prague, Czech Republic

⁶Institute of Biotechnology, Czech Academy of Sciences, Prague-West, Czech Republic

⁷Unit of Clinical Research, Catalan Institute of Oncology, Girona, Spain

⁸INSERM U1016, Institut Cochin, Paris, France

⁹CNRS UMR 8104, Paris, France

¹⁰Université Paris Descartes, Sorbonne Paris Cité, Paris, France

¹¹School of Medical Science, Menzies Health Institute Queensland, Griffith University, Southport, Queensland, Australia

Correspondence

Javier A. Menendez, Girona Biomedical Research Institute (IDIBGI), Edifici M2, Parc Hospitalari Martí i Julià, E-17190 Salt, Girona, Spain.
Email: jmenendez@iconcologia.net or jmenendez@idibgi.org

Funding information

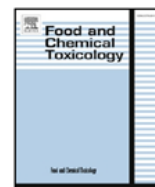
This work was supported by grants from the Ministerio de Ciencia e Innovación (Grant SAF2016-80639-P to J. A. Menendez), Plan Nacional de I+D+i, Spain, the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) (Grant 2014 SGR229 to J. A. Menendez), Departament d'Economia i Coneixement, Catalonia, Spain, and the Czech Health Council Foundation (Grant 16-31604A to J. Neuzil). Elisabet Cuyàs is supported by the Sara Borrell postdoctoral contract (CD15/00033) from the Ministerio de Sanidad y Consumo, Fondo de

Summary

Metformin, the first drug chosen to be tested in a clinical trial aimed to target the biology of aging per se, has been clinically exploited for decades in the absence of a complete understanding of its therapeutic targets or chemical determinants. We here outline a systematic chemoinformatics approach to computationally predict biomolecular targets of metformin. Using several structure- and ligand-based software tools and reference databases containing 1,300,000 chemical compounds and more than 9,000 binding sites protein cavities, we identified 41 putative metformin targets including several epigenetic modifiers such as the member of the H3K27me3-specific demethylase subfamily, KDM6A/UTX. AlphaScreen and AlphaLISA assays confirmed the ability of metformin to inhibit the demethylation activity of purified KDM6A/UTX enzyme. Structural studies revealed that metformin might occupy the same set of residues involved in H3K27me3 binding and demethylation within the catalytic pocket of KDM6A/UTX. Millimolar metformin augmented global levels of H3K27me3

[†]On behalf of the METTEN study group (EudraClinicalTrial Number 2011-000490-30).

*These authors contributed equally to this work.



Silibinin is a direct inhibitor of STAT3

Sara Verdura^{a,b,1}, Elisabet Cuyàs^{a,b,1}, Laura Llorach-Parés^c, Almudena Pérez-Sánchez^d, Vicente Micol^{d,e}, Alfons Nonell-Canals^c, Jorge Joven^f, Manuel Valiente^g, Melchor Sánchez-Martínez^c, Joaquim Bosch-Barrera^{h,i,**}, Javier A. Menendez^{a,b,*}

^a Program Against Cancer Therapeutic Resistance (ProCURE), Metabolism and Cancer Group, Catalan Institute of Oncology, Girona, Spain

^b Molecular Oncology Group, Girona Biomedical Research Institute (IDIBGI), Girona, Spain

^c Mind the Byte, Barcelona, Spain

^d Instituto de Biología Molecular y Celular (IBMC), Miguel Hernández University (UMH), Elche, Alicante, Spain

^e CIBER, Fisiopatología de la Obesidad y la Nutrición, CIBERobn, Instituto de Salud Carlos III (CB12/03/30038), Spain

^f Unitat de Recerca Biomèdica, Hospital Universitari de Sant Joan, IISPV, Rovira i Virgili University, Reus, Spain

^g Brain Metastasis Group, Molecular Oncology Program, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

^h Department of Medical Oncology, Catalan Institute of Oncology, Girona, Spain

ⁱ Department of Medical Sciences, Medical School, University of Girona, Girona, Spain

ARTICLE INFO

Keywords:

Silibinin
STAT3
Cancer
Metastasis

ABSTRACT

We herein combined experimental and computational efforts to delineate the mechanism of action through which the flavonolignan silibinin targets STAT3. Silibinin reduced IL-6 inducible, constitutive, and acquired feedback activation of STAT3 at tyrosine 705 (Y705). Silibinin attenuated the inducible phospho-activation of Y705 in GFP-STAT3 genetic fusions without drastically altering the kinase activity of the STAT3 upstream kinases JAK1 and JAK2. A comparative computational study based on docking and molecular dynamics simulation over 14 different STAT3 inhibitors (STAT3i) predicted that silibinin could directly bind with high affinity to both the Src homology-2 (SH2) domain and the DNA-binding domain (DBD) of STAT3. Silibinin partially overlapped with the cavity occupied by other STAT3i in the SH2 domain to indirectly prevent Y705 phosphorylation, yet showing a unique binding mode. Moreover, silibinin was the only STAT3i predicted to establish direct interactions with DNA in its targeting to the STAT3 DBD. The prevention of STAT3 nuclear translocation, the blockade of the binding of activated STAT3 to its consensus DNA sequence, and the suppression of STAT3-directed transcriptional activity confirmed silibinin as a direct STAT3i. The unique characteristics of silibinin as a bimodal SH2- and DBD-targeting STAT3i make silibinin a promising lead for designing new, more effective STAT3i.

1. Introduction

The aberrant activation of signal transducer and activator of transcription 3 (STAT3) contributes to cancer initiation and progression in a multi-faceted manner via promotion of cell proliferation/survival, invasion/migration, angiogenesis, and immune-evasion (Chang et al., 2013; Sansone and Bromberg, 2012; Yu et al., 2009, 2014). Feedback activation of STAT3 additionally mediates tumor resistance to a broad spectrum of cancer therapies, including radiotherapy, conventional chemotherapy, and modern targeted therapies (Lee et al., 2014; Poli and Camporeale, 2015; Tan et al., 2014; Zhao et al., 2016). STAT3 activation associates also with the generation and maintenance of

cancer stem cells (CSC), a particularly aggressive type of malignant cell defined in terms of functional traits including tumor/metastasis-initiating capacity and therapy resistance (Kroon et al., 2013; Misra et al., 2018; Schroeder et al., 2014; Wang et al., 2018). Not surprisingly, the activation status of STAT3 is a strong predictor of poor prognosis and is an independent risk factor for tumor recurrence and post-therapy progression (Chen et al., 2013; Liu et al., 2012; Tong et al., 2017; Wu et al., 2016). These observations have motivated great efforts over the last decade to clinically exploit the beneficial effects of inhibiting STAT3 in human malignancies. Accordingly, a large number of STAT3 inhibitors (STAT3i) have been developed as potential cancer therapeutics (Fagard et al., 2013; Furtek et al., 2016a; b; Jin et al., 2016; Miklossy et al.,

* Corresponding author. Catalan Institute of Oncology (ICO), Girona Biomedical Research Institute (IDIBGI), Edifici M2, Parc Hospitalari Martí i Julià, E-17190 Salt, Girona, Spain.

** Corresponding author. Catalan Institute of Oncology (ICO), Hospital Dr. Josep Trueta de Girona, Avda. de França s/n, 17007, Girona, Spain

E-mail addresses: jbosch@iconcologia.net (J. Bosch-Barrera), jmenendez@iconcologia.net, jmenendez@idibgi.org (J.A. Menendez).

¹ These authors contributed equally.



Metformin Is a Direct SIRT1-Activating Compound: Computational Modeling and Experimental Validation

Elisabet Cuyàs^{1,2†}, Sara Verdura^{1,2†}, Laura Llorach-Parés^{3†}, Salvador Fernández-Arroyo⁴, Jorge Joven⁴, Begoña Martín-Castillo^{2,5}, Joaquim Bosch-Barrera^{2,6,7}, Joan Brunet^{7,8,9}, Alfons Nonell-Canals³, Melchor Sanchez-Martinez³ and Javier A. Menendez^{1,2*}

OPEN ACCESS

Edited by:

Frederic Bost,
Centre National de la Recherche
Scientifique (CNRS), France

Reviewed by:

Pierre De Meyts,
de Duve Institute, Belgium
Yves Combarnous,
Centre National de la
Recherche Scientifique (CNRS),
France

*Correspondence:

Javier A. Menendez
jmenendez@iconcologia.net;
jmenendez@idibgi.org

† These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Cellular Endocrinology,
a section of the journal
Frontiers in Endocrinology

Received: 05 July 2018

Accepted: 19 October 2018

Published: 06 November 2018

Citation:

Cuyàs E, Verdura S, Llorach-Parés L,
Fernández-Arroyo S, Joven J,
Martín-Castillo B, Bosch-Barrera J,
Brunet J, Nonell-Canals A,
Sanchez-Martinez M and
Menendez JA (2018) Metformin Is a
Direct SIRT1-Activating Compound:
Computational Modeling and
Experimental Validation.
Front. Endocrinol. 9:657.
doi: 10.3389/fendo.2018.00657

¹ ProCURE (Program Against Cancer Therapeutic Resistance), Metabolism and Cancer Group, Catalan Institute of Oncology, Girona, Spain, ² Girona Biomedical Research Institute (IDIBGI), Girona, Spain, ³ Mind the Byte, Barcelona, Spain, ⁴ Unitat de Recerca Biomèdica, Hospital Universitari de Sant Joan, Institut d'Investigació Sanitària Pere Virgili (IISPV), Rovira i Virgili University, Reus, Spain, ⁵ Unit of Clinical Research, Catalan Institute of Oncology (ICO), Girona, Spain, ⁶ Department of Medical Sciences, Medical School, University of Girona, Girona, Spain, ⁷ Medical Oncology, Catalan Institute of Oncology (ICO), Dr. Josep Trueta University Hospital, Girona, Spain, ⁸ Hereditary Cancer Programme, Catalan Institute of Oncology (ICO), Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet del Llobregat, Barcelona, Spain, ⁹ Hereditary Cancer Programme, Catalan Institute of Oncology (ICO), Girona Biomedical Research Institute (IDIBGI), Girona, Spain

Metformin has been proposed to operate as an agonist of SIRT1, a nicotinamide adenine dinucleotide (NAD⁺)-dependent deacetylase that mimics most of the metabolic responses to calorie restriction. Herein, we present an *in silico* analysis focusing on the molecular docking and dynamic simulation of the putative interactions between metformin and SIRT1. Using eight different crystal structures of human SIRT1 protein, our computational approach was able to delineate the putative binding modes of metformin to several pockets inside and outside the central deacetylase catalytic domain. First, metformin was predicted to interact with the very same allosteric site occupied by resveratrol and other sirtuin-activating compounds (STATCs) at the amino-terminal activation domain of SIRT1. Second, metformin was predicted to interact with the NAD⁺ binding site in a manner slightly different to that of SIRT1 inhibitors containing an indole ring. Third, metformin was predicted to interact with the C-terminal regulatory segment of SIRT1 bound to the NAD⁺ hydrolysis product ADP-ribose, a “C-pocket”-related mechanism that appears to be essential for mechanism-based activation of SIRT1. Enzymatic assays confirmed that the net biochemical effect of metformin and other biguanides such as a phenformin was to improve the catalytic efficiency of SIRT1 operating in conditions of low NAD⁺ *in vitro*. Forthcoming studies should confirm the mechanistic relevance of our computational insights into how the putative binding modes of metformin to SIRT1 could explain its ability to operate as a direct SIRT1-activating compound. These findings might have important implications for understanding how metformin might confer health benefits *via* maintenance of SIRT1 activity during the aging process when NAD⁺ levels decline.

Keywords: metformin, SIRT1, aging, NAD⁺, NAD loss

ORIGINAL ARTICLE

An olive oil phenolic is a new chemotype of mutant isocitrate dehydrogenase 1 (IDH1) inhibitors

Sara Verdura^{1,2,†}, Elisabet Cuyàs^{1,2,†}, Jesús Lozano-Sánchez^{3,4,†,•}, Cristian Bastidas-Velez^{2,†}, Laura Llorach-Parés⁵, Salvador Fernández-Arroyo⁶, Anna Hernández-Aguilera⁶, Jorge Joven⁶, Alfons Nonell-Canals⁵, Joaquim Bosch-Barrera^{2,7,8}, Begoña Martín-Castillo^{2,9}, Luciano Vellon¹⁰, Melchor Sanchez-Martinez^{5,•}, Antonio Segura-Carretero^{3,4,•} and Javier A. Menendez^{1,2,*}

¹Program Against Cancer Therapeutic Resistance (ProCURE), Metabolism and Cancer Group, Catalan Institute of Oncology, Girona, Spain, ²Girona Biomedical Research Institute (IDIBGI), Edifici M2, Parc Hospitalari Martí i Julià, E-17190 Salt, Girona, Spain, ³Department of Analytical Chemistry, Faculty of Sciences, University of Granada, Granada, Spain, ⁴Research and Development Functional Food Centre (CIDAF), PTS Granada, Granada, Spain, ⁵Mind the Byte, Barcelona, Spain, ⁶Unitat de Recerca Biomèdica, Hospital Universitari de Sant Joan, IISPV, Rovira i Virgili University, Reus, Spain, ⁷Department of Medical Sciences, Medical School, University of Girona, Girona, Spain, ⁸Medical Oncology and ⁹Unit of Clinical Research, Catalan Institute of Oncology, Girona, Spain and ¹⁰Stem Cells Laboratory, Institute of Biology and Experimental Medicine (IBYME-CONICET), Buenos Aires, Argentina

*To whom correspondence should be addressed. Tel: +34 872 987 087 ext. 50; Fax: +34 972 217 344; Email: jmenendez@iconcologia.net or jmenendez@idibgi.org

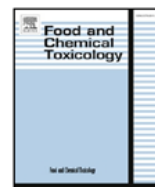
[†]These authors contributed equally.

Abstract

Mutations in the isocitrate dehydrogenase 1 (IDH1) gene confer an oncogenic gain-of-function activity that allows the conversion of α -ketoglutarate (α -KG) to the oncometabolite R-2-hydroxyglutarate (2HG). The accumulation of 2HG inhibits α -KG-dependent histone and DNA demethylases, thereby generating genome-wide hypermethylation phenotypes with cancer-initiating properties. Several chemotypes of mutant IDH1/2-targeted inhibitors have been reported, and some of them are under evaluation in clinical trials. However, the recognition of acquired resistance to such inhibitors within a few years of clinical use raises an urgent need to discover new mutant IDH1 antagonists. Here, we report that a naturally occurring phenolic compound in extra-virgin olive oil (EVOO) selectively inhibits the production of 2HG by neomorphic IDH1 mutations. *In silico* docking, molecular dynamics, including steered simulations, predicted the ability of the oleoside decarboxymethyl oleuropein aglycone (DOA) to preferentially occupy the allosteric pocket of mutant IDH1. DOA inhibited the enzymatic activity of recombinant mutant IDH1 (R132H) protein in the low micromolar range, whereas >10-fold higher concentrations were required to inhibit the activity of wild-type (WT) IDH1. DOA suppressed 2HG overproduction in engineered human cells expressing a heterozygous IDH1-R132H mutation. DOA restored the 2HG-suppressed activity of histone demethylases as it fully reversed the hypermethylation of H3K9me3 in IDH1-mutant cells. DOA epigenetically restored the expression of *PD-L1*, an immunosuppressive gene silenced in IDH1 mutant cells via 2HG-driven DNA hypermethylation. DOA selectively blocked colony formation of IDH1 mutant cells while sparing WT IDH1 isogenic counterparts. In sum, the EVOO-derived oleoside DOA is a new, naturally occurring chemotype of mutant IDH1 inhibitors.

Received: August 24, 2018; Revised: October 9, 2018; Accepted: November 13, 2018

© The Author(s) 2018. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com.



The extra virgin olive oil phenolic oleacein is a dual substrate-inhibitor of catechol-O-methyltransferase

Elisabet Cuyàs^{a,b}, Sara Verdura^{a,b}, Jesús Lozano-Sánchez^{c,d}, Ignacio Viciano^e,
 Laura Llorach-Parés^e, Alfons Nonell-Canals^e, Joaquim Bosch-Barrera^{b,f,g}, Joan Brunet^{f,g,h,i},
 Antonio Segura-Carretero^{c,d}, Melchor Sanchez-Martinez^e, José Antonio Encinar^{j,*},
 Javier A. Menendez^{a,b,**}

^a ProCURE (Program Against Cancer Therapeutic Resistance), Metabolism & Cancer Group, Catalan Institute of Oncology, Girona, Spain

^b Girona Biomedical Research Institute (IDIBGI), Girona, Spain

^c Department of Analytical Chemistry, Faculty of Sciences, University of Granada, Granada, Spain

^d Research and Development Functional Food Centre (CIDAF), PTS Granada, Granada, Spain

^e Mind the Byte, Barcelona, Spain

^f Medical Oncology, Catalan Institute of Oncology (ICO) Dr. Josep Trueta University Hospital, Girona, Spain

^g Department of Medical Sciences, Medical School University of Girona, Girona, Spain

^h Hereditary Cancer Programme, Catalan Institute of Oncology (ICO), Bellvitge Institute for Biomedical Research (IDIBELL) L'Hospitalet del Llobregat, Barcelona, Spain

ⁱ Hereditary Cancer Programme, Catalan Institute of Oncology (ICO) Girona Biomedical Research Institute (IDIBGI), Girona, Spain

^j Institute of Research, Development and Innovation in Biotechnology of Elche (IDiBE) and Molecular and Cell Biology Institute (IBMC), Miguel Hernández University (UMH), Elche, Spain

ARTICLE INFO

Keywords:

Extra virgin olive oil
 Polyphenols
 Secoiridoids
 Oleacein
 COMT
 Cancer

ABSTRACT

Catechol-containing polyphenols present in coffee and tea, while serving as excellent substrates for catechol-O-methyltransferase (COMT)-catalyzed O-methylation, can also operate as COMT inhibitors. However, little is known about the relationship between COMT and the characteristic phenolics present in extra virgin olive oil (EVOO). We here selected the EVOO dihydroxy-phenol oleacein for a computational study of COMT-driven methylation using classic molecular docking/molecular dynamics simulations and hybrid quantum mechanical/molecular mechanics, which were supported by *in vitro* activity studies using human COMT. Oleacein could be superimposed onto the catechol-binding site of COMT, maintaining the interactions with the atomic positions involved in methyl transfer from the S-adenosyl-L-methionine cofactor. The transition state structure for the *meta*-methylation in the O5 position of the oleacein benzenediol moiety was predicted to occur preferentially. Enzyme analysis of the conversion ratio of catechol to O-alkylated guaiacol confirmed the inhibitory effect of oleacein on human COMT, which remained unaltered when tested against the protein version encoded by the functional Val¹⁵⁸Met polymorphism of the COMT gene. Our study provides a theoretical determination of how EVOO dihydroxy-phenols can be metabolized via COMT. The ability of oleacein to inhibit COMT adds a new dimension to the physiological and therapeutic utility of EVOO secoiridoids.

1. Introduction

Human catechol-O-methyltransferase (COMT) is a phase II detoxifying enzyme (UniProt code P21964) that catalyzes the transfer of a methyl moiety from the S-adenosyl-L-methionine (SAM) cofactor to one of the hydroxyl groups present in endogenous neurotransmitters (e.g.,

catecholamines) and hormones (e.g., estradiol), and also xenobiotic substances that incorporate catecholic structures (Bai et al., 2007; Mannisto and Kaakkola, 1999; Zhu and Conney, 1998).

Various catechol-containing coffee and tea polyphenols have been described as excellent substrates for COMT-mediated O-methylation (Zhu and Liehr, 1996; Zhu et al., 2000, 2001, 2009). Catechol-

* Corresponding author. Instituto de Investigación, Desarrollo e Innovación en Biotecnología Sanitaria de Elche (IDiBE), Av. de la Universidad, Edif. Torregaitan, Despacho 2.08, E-03202, Elche, Alicante, Spain.

** Corresponding author. Catalan Institute of Oncology (ICO), Girona Biomedical Research Institute (IDIBGI), Edifici M2, Parc Hospitalari Martí i Julià, E-17190, Salt, Girona, Spain.

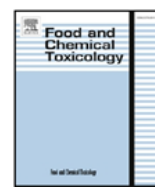
E-mail addresses: jant.encinar@umh.es (J.A. Encinar), jmenendez@iconcologia.net, jmenendez@idibgi.org (J.A. Menendez).

<https://doi.org/10.1016/j.fct.2019.03.049>

Received 26 February 2019; Received in revised form 25 March 2019; Accepted 26 March 2019

Available online 29 March 2019

0278-6915/ © 2019 Published by Elsevier Ltd.



Computational de-orphanization of the olive oil biophenol oleacein: Discovery of new metabolic and epigenetic targets

Elisabet Cuyàs^{a,b}, David Castillo^c, Laura Llorach-Parés^d, Jesús Lozano-Sánchez^{e,f}, Sara Verdura^{a,b}, Alfons Nonell-Canals^d, Joan Brunet^{g,h,i,j}, Joaquim Bosch-Barrera^{b,g,h}, Jorge Joven^k, Rafael Valdés^c, Melchor Sanchez-Martinez^d, Antonio Segura-Carretero^{e,f}, Javier A. Menendez^{a,b,*}

^a ProCURE (Program Against Cancer Therapeutic Resistance), Metabolism & Cancer Group, Catalan Institute of Oncology, Girona, Spain

^b Girona Biomedical Research Institute (IDIBGI), Girona, Spain

^c Dreamgenics, Oviedo, Spain

^d Mind the Byte, Barcelona, Spain

^e Department of Analytical Chemistry, Faculty of Sciences, University of Granada, Granada, Spain

^f Research and Development Functional Food Centre (CIDAF), PTS Granada, Granada, Spain

^g Medical Oncology, Catalan Institute of Oncology (ICO), Dr. Josep Trueta University Hospital, Girona, Spain

^h Department of Medical Sciences, Medical School University of Girona, Girona, Spain

ⁱ Hereditary Cancer Programme, Catalan Institute of Oncology (ICO), Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet del Llobregat, Barcelona, Spain

^j Hereditary Cancer Programme, Catalan Institute of Oncology (ICO) Girona Biomedical Research Institute (IDIBGI), Girona, Spain

^k Campus of International Excellence Southern Catalonia, Unitat de Recerca Biomèdica, Hospital Universitari de Sant Joan, Institut d'Investigació Sanitària, Pere Virgili, Universitat Rovira i Virgili, Reus, Spain

ARTICLE INFO

Keywords:

Olive oil
Oleacein
Chemoinformatics
Metabolism
Epigenetics

ABSTRACT

The health promoting effects of extra virgin olive oil (EVOO) relate to its unique repertoire of phenolic compounds. Here, we used a chemoinformatics approach to computationally identify endogenous ligands and assign putative biomolecular targets to oleacein, one of the most abundant secoiridoids in EVOO. Using a structure-based virtual profiling software tool and reference databases containing more than 9000 binding sites protein cavities, we identified 996 putative oleacein targets involving more than 700 proteins. We subsequently identified the high-level functions of oleacein in terms of biomolecular interactions, signaling pathways, and protein-protein interaction (PPI) networks. Delineation of the oleacein target landscape revealed that the most significant modules affected by oleacein were associated with metabolic processes (e.g., glucose and lipid metabolism) and chromatin-modifying enzymatic activities (i.e., histone post-translational modifications). We experimentally confirmed that, in a low-micromolar physiological range (< 20 μmol/l), oleacein was capable of inhibiting the catalytic activities of predicted metabolic and epigenetic targets including nicotinamide N-methyltransferase, ATP-citrate lyase, lysine-specific demethylase 6A, and N-methyltransferase 4. Our computational de-orphanization of oleacein provides new mechanisms through which EVOO biophenols might operate as chemical prototypes capable of modulating the biologic machinery of healthy aging.

1. Introduction

The ability of the “Mediterranean diet”, which reflects the dietary patterns found in olive-growing areas of the Mediterranean basin, to significantly reduce aging-related morbidity and promote increased life expectancy can be largely attributed to the unique nutraceutical properties of extra virgin olive oil (EVOO) (Colomer and Menendez, 2006; Menendez and Lupu, 2006; Escrich et al., 2007; López-Miranda et al.,

2010; Fernández del Río et al., 2016; Piroddi et al., 2017). The positive influence of EVOO on human health has been historically ascribed to its high content of monounsaturated fatty acids (e.g., oleic acid; 18:1n-9). However, it has been shown that other oleic acid-rich oils but lacking the characteristic functional components of EVOO (e.g., biophenols) such as high-oleic canola or high-oleic sunflower oils do not share the same ability to improve, for example, cardiovascular prognosis, and to concurrently lowering the incidence of cancer and neurodegeneration

* Corresponding author. Catalan Institute of Oncology (ICO), Girona Biomedical Research Institute (IDIBGI), Edifici M2, Parc Hospitalari Martí i Julià, E-17190, Salt, Girona, Spain.

E-mail addresses: jmenendez@iconcologia.net, jmenendez@idibgi.org (J.A. Menendez).

<https://doi.org/10.1016/j.fct.2019.05.037>









Received 14 March 2019; Received in revised form 23 May 2019; Accepted 24 May 2019

Available online 29 May 2019

0278-6915/ © 2019 Elsevier Ltd. All rights reserved.

Communication

Extra Virgin Olive Oil Contains a Phenolic Inhibitor of the Histone Demethylase LSD1/KDM1A

Elisabet Cuyàs ^{1,2}, Juan Gumuzio ³, Jesús Lozano-Sánchez ^{4,5} , David Carreras ^{2,6}, Sara Verdura ^{1,2} , Laura Llorach-Parés ⁷ , Melchor Sanchez-Martinez ⁷ , Elisabet Selga ^{6,8,9} , Guillermo J. Pérez ^{6,8} , Fabiana S. Scornik ^{6,8}, Ramon Brugada ^{2,6,8,10}, Joaquim Bosch-Barrera ^{2,11,12} , Antonio Segura-Carretero ^{4,5}, Ángel G. Martin ³, José Antonio Encinar ¹³  and Javier A. Menendez ^{1,2,*}

¹ ProCURE (Program Against Cancer Therapeutic Resistance), Metabolism & Cancer Group, Catalan Institute of Oncology, 17007 Girona, Spain

² Girona Biomedical Research Institute (IDIBGI), 17190 Girona, Spain

³ StemTek Therapeutics, 48160 Bilbao, Spain

⁴ Department of Analytical Chemistry, Faculty of Sciences, University of Granada, 18071 Granada, Spain

⁵ Research and Development Functional Food Centre (CIDAF), PTS Granada, 18100 Granada, Spain

⁶ Cardiovascular Genetics Centre, Department of Medical Sciences, University of Girona, 17071 Girona, Spain

⁷ Mind the Byte, 08007 Barcelona, Spain

⁸ Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), 28029 Madrid, Spain

⁹ Faculty of Medicine, University of Vic-Central University of Catalonia (UVic-UCC), 08500 Vic, Spain

¹⁰ Dr. Josep Trueta Hospital of Girona, 17007 Girona, Spain

¹¹ Medical Oncology, Catalan Institute of Oncology (ICO), 17007 Girona, Spain

¹² Department of Medical Sciences, Medical School University of Girona, 17071 Girona, Spain

¹³ Institute of Research, Development and Innovation in Biotechnology of Elche (IDiBE) and Molecular and Cell Biology Institute (IBMC), Miguel Hernández University (UMH), 03202 Elche, Spain

* Correspondence: jmenendez@iconcologia.net or jmenendez@idibgi.org

Received: 5 July 2019; Accepted: 17 July 2019; Published: 19 July 2019



Abstract: The lysine-specific histone demethylase 1A (LSD1) also known as lysine (K)-specific demethylase 1A (KDM1A) is a central epigenetic regulator of metabolic reprogramming in obesity-associated diseases, neurological disorders, and cancer. Here, we evaluated the ability of oleacein, a biophenol secoiridoid naturally present in extra virgin olive oil (EVOO), to target LSD1. Molecular docking and dynamic simulation approaches revealed that oleacein could target the binding site of the LSD1 cofactor flavin adenosine dinucleotide with high affinity and at low concentrations. At higher concentrations, oleacein was predicted to target the interaction of LSD1 with histone H3 and the LSD1 co-repressor (RCOR1/CoREST), likely disturbing the anchorage of LSD1 to chromatin. AlphaScreen-based in vitro assays confirmed the ability of oleacein to act as a direct inhibitor of recombinant LSD1, with an IC₅₀ as low as 2.5 μmol/L. Further, oleacein fully suppressed the expression of the transcription factor SOX2 (SEX determining Region Y-box 2) in cancer stem-like and induced pluripotent stem (iPS) cells, which specifically occurs under the control of an LSD1-targeted distal enhancer. Conversely, oleacein failed to modify ectopic SOX2 overexpression driven by a constitutive promoter. Overall, our findings provide the first evidence that EVOO contains a naturally occurring phenolic inhibitor of LSD1, and support the use of oleacein as a template to design new secoiridoid-based LSD1 inhibitors.

Keywords: phenolics; secoiridoids; cancer; cancer stem cells; SOX2; metabolism; neurological disorders

