



**IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS
THAT CAN PREVENT CARDIOVASCULAR DISEASES**
Esther Sala Argüello

ISBN:
Dipòsit Legal: T. 1030-2011

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tesisenxarxa.net) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tesisenred.net) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tesisenxarxa.net) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Esther Sala Argüello

***In silico* methodologies for the design of functional foods that
can prevent cardiovascular diseases**

PhD DOCTORAL THESIS

Directed by Dr.Gerard Pujadas Anguiano

Biochemistry & Biotechnology Department



UNIVERSITAT ROVIRA I VIRGILI

Tarragona

2011

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011



Departament de Bioquímica I Biotecnologia

C/Marcel·lí Domingo s/n
Campus Sescelades
43007 Tarragona
Tel. : +34 977 559 521
Fax : +34 977 558 232

FAIG CONSTAR que aquest treball titulat “In silico methodologies for the design of functional foods that can prevent cardiovascular diseases”, que presenta Esther Sala Argüello per a l’obtenció del títol de Doctor, ha estat realitzat sota la meva direcció al Departament de Bioquímica i Biotecnologia d’aquesta universitat i que aconsegueix els requeriments per poder optar a Menció Europea.

Tarragona, 3 de Març de 2011

Director de la tesi doctoral,

Dr. Gerard Pujadas Anguiano

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

CONTENTS

| | | |
|------------|---|------------|
| I | INTRODUCTION | 1 |
| II | CONTEXT AND GOALS | 17 |
| III | RESULTS | 21 |
| | Inhibitor of nuclear factor kappa-B kinase subunit beta | 23 |
| 1. | Identification of Human IKK-2 Inhibitors of Natural Origin (Part I): Modeling of the IKK-2 Kinase Domain, Virtual screening and Activity Assays. | 25 |
| 1.1. | Supporting Information (manuscript 1) | 41 |
| 2. | Identification of IKK-2 Inhibitors of Natural Origin (Part II): In silico prediction of IKK-2 inhibitors in natural extracts with known anti-inflammatory activity. | 61 |
| 2.1. | Supporting Information (manuscript 2) | 81 |
| 3. | 3D-QSAR Study of Pyrine Derivates as IKK-2 Inhibitors. | 87 |
| | 11 beta hydroxysteroid dehydrogenase type 1 | 107 |
| 4. | A Structure-Based <i>In Silico</i> Methodology for Predicting 11beta-HSD2 Inhibitors. | 109 |
| 4.1. | Supporting Information (manuscript 4) | 135 |
| 5. | <i>In silico</i> prediction of selective 11 β -HSD1 inhibitors of Natural Origin. | 159 |
| IV | SUMMARIZING DISCUSSION | 189 |
| V | CONCLUSIONS | 195 |

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

I. INTRODUCTION

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

The primary role of any diet is to provide enough nutrients to meet metabolic requirements. Nevertheless, recent findings support the hypothesis that some diet ingredients, in addition to meeting nutrition needs, may modulate various physiological functions and may play detrimental or beneficial roles in some diseases¹⁻⁴. Thus, according to the European Commission concerted action on **Functional Food Science (FuFoSE)**, a food can be regarded as *functional* if it contains a bioactive component (whether or not a nutrient but from natural origin) and has been demonstrated to beneficially affect one or more target functions in the body. This efficacy must be beyond adequate nutritional effects in a way that is relevant to either the maintenance or promotion of a state of well-being or health or to the reduction of the risk of a pathologic process or of disease consequences⁵. Therefore, the working hypothesis is that the bioactive molecules that confer *functionality* to a food do so, fundamentally, through the activation or inhibition of specific targets (*e.g.*, nuclear receptors or enzymes) in a way that produces the above-mentioned beneficial effects⁶⁻⁹.

In the year 2000, the worldwide market of functional foods generated US\$ 33 billion. In 2005, this total was increased to US\$ 73.5 billion¹⁰, and in some occidental countries, it is expected that the portion of the total food market corresponding to functional foods will double during the period of 2006-2020¹¹. Nevertheless, to accomplish these predictions, it is necessary to (1) minimize the expenses needed to demonstrate the positive effect of new food ingredients on health (regulation EC No 1924/2006 of the European Union states that foods can only be marketed as *functional foods* if they have been scientifically proven to have a positive effect on health); and (2) introduce in the bioactive extracts market new food ingredients with a better activity/cost ratio than the ones currently on use. Unfortunately, bioactivity screening in extracts exclusively by *in vitro* or *in vivo* approaches is a complex and expensive process that is difficult for the functional food industry to afford in a context characterized by the competence of private brands and the general economic situation. However, virtual screening workflows may play an essential role in significantly lowering the R&D expenses associated with these goals by (a) allowing studying without a significant budget increment any number of extracts; (b) predicting which extracts have the desired bioactivity; (c) predicting the level of the extracts' bioactivity; (d) allowing a predictive-based selection of the extracts whose bioactivity will be experimentally analyzed (which contributes to minimizing the number of assays needed to demonstrate a positive effect on health); and (e) shortening the time needed to patent and transfer the most interesting results to the functional food industry.

The goal of virtual screening techniques is to identify the most promising candidates on which one can focus experimental efforts by eliminating molecules that do not possess the required features¹².

There are successful examples, especially in the pharmaceutical industry, of the use of virtual screening techniques for the discovery of bioactive products¹²⁻¹⁴. An important difference between the pharmaceutical industry and the food industry lies in the type of molecules that are involved. While any molecule is likely to have interest as a drug (including natural molecules), in the field of functional foods, the repertoire of molecules to be screened is restricted to those found in nature. Although there are more than 200,000 known natural compounds¹⁵, it is estimated that only 5-15% of the 250,000 described plant species have been used in phytochemical and pharmacological research¹⁶. Although natural products have been used as sources of new drugs¹⁷⁻¹⁹, their use in the functional food sector is restricted to a very limited repertoire of molecules. Therefore, virtual screening techniques can be of use for identifying undescribed bioactivities for known natural molecules and, subsequently, for increasing the repertoire of ingredients that can be used in functional food development. Nevertheless, the use of such computational techniques in functional food design is only just emerging²⁰. At this point, it is key to remark that the identification of the natural source where you can find these compounds is as important as the discovery of new bioactive compounds itself, as it makes little sense to chemically synthesize a natural product for use in a functional food. In fact, the use of a natural extract that is rich in the desired compound or compounds seems to be the ideal solution. Unfortunately, this crucial information is only available in very expensive natural molecule databases like Reaxys (<https://www.reaxys.com/>) and the Dictionary of Natural Products (<http://dnp.chemnetbase.com/>).

The current PhD thesis aims to use virtual screening workflows to identify food ingredients with previously undescribed activity in preventing, retarding and/or reversing the metabolic risk factors (*i.e.*, abdominal obesity, insulin resistance, hypertension, prothrombotic state, and atherogenic dyslipidemia) associated with metabolic syndrome. Metabolic syndrome increases the risk of developing cardiovascular diseases and diabetes and affects one in five people, making it one of the greatest healthcare challenges of the 21st century and a leading cause of death in developed countries. Inflammation is a complex immune response that is triggered by damage to living tissues that protect the organism from infections and injury. The two “types” of inflammatory response include the following: **(a)** one called acute inflammation (which lasts only a few days) and **(b)** a response of longer duration that is referred to as chronic inflammation. The result of each inflammatory reaction may be beneficial (defense against agents interfering with homeostasis) or harmful (causing damage to cells and tissues)^{21, 22}. Chronic inflammation is directly involved in the combination of metabolic risk factors that characterize the metabolic syndrome²³⁻²⁶. During last few years, different studies have highlighted the relationship between type 2 diabetes and chronic inflammation.^{24, 27-30} There are also growing evidence that suggests that the cooperation between metabolic and inflammatory pathways is disrupted in the pathogenesis of the metabolic

complications of obesity^{31, 32}. Therefore, targeting the central molecules involved in the integration of metabolic and immune/inflammatory responses appears to have good prospects for a successful therapeutic approach in the metabolic syndrome³³. Thus, the present PhD thesis will focus on predicting which natural products can be of therapeutic use for modulating the activity of two relevant targets of chronic inflammation treatment: (a) inhibitor nuclear-factor κ B kinase 2 (also called IKK-2 or IKK- β); and (b) 11 β -hydroxysteroid dehydrogenase type 1 (so-called 11 β -HSD1).

The human inhibitor nuclear factor κ B subunit beta

IKK-2 is a cytoplasmic serine-threonine protein kinase belonging to the IKK complex and is the primary component responsible for activating NF- κ B in response to various inflammatory stimuli. This kinase complex consists of **(a)** two catalytic subunits (*i.e.*, IKK- α or IKK-1 and IKK- β or IKK-2); **(b)** one regulatory subunit, called IKK- γ or NEMO; **(c)** NIK; and **(d)** other various known and unknown proteins³⁴. NF- κ B is normally retained in the cytoplasm as an inactive form bound to the inhibitor kappa B α (I κ B α). When is stimulated by TNF α , IL-1 or LPS, the IKK complex catalyzes the phosphorylation of the N-terminal regulatory serines on I κ B α and I κ B β and, subsequently, leads to the release of different dimers of NF- κ B and the fast activation of the transcription of a variety of genes encoding cytokines (including TNF- α , IL-1, IL-6 and ICAM), growth factors, adhesion molecules and procoagulatory factors³⁵. Although NF- κ B plays an essential and beneficial role in normal physiology for immune and inflammation responses, the constitutive activation of NF- κ B has also been found to be associated with most cancers and other chronic inflammation-related pathologies like diabetes and cardiovascular disease³⁶. Thus, for instance, the dysregulation of the NF- κ B system is involved in early and later stages of the inflammatory-proliferative process of atherogenesis and, therefore, plays an important role in atherosclerosis^{21, 27, 37}. There are two pathways that lead to the activation of NF- κ B: (a) the canonical (or classical); and (b) the non-canonical (or alternative). In the canonical pathway (see Figure 1), a cellular stimulation by inflammatory cytokines activates IKK-2, which subsequently phosphorylates I κ B α in an NEMO-dependent manner. The phosphorylated I κ B α can no longer bind to (and inactivate) NF- κ B, and it is released from the complex and metabolized. Then, the *free* NF- κ B can translocate into the nucleus and activate the transcription of genes involved in inflammation. This canonical NF- κ B pathway has been implicated as an important pathway in coordinating the expression of multiple inflammatory and innate immune genes. For that reason, the canonical pathway is directly related with inflammatory stimuli^{38, 39}. In contrast to the canonical pathway, the non-canonical pathway (see Figure 1) is mainly stimulated by the ligation of LT β R, BAFFR and CD40R, which is mediated by NIK and is strictly dependent on the IKK-1 homodimer but independent of IKK-2 and NEMO. The target for IKK-1 homodimers is the NF- κ B

p52/p100 protein. C-terminal phosphorylation is essential for p100 processing to p52, which is also dependent on polyubiquitination and proteasomal degradation. However, the phosphorylation-dependent ubiquitination of p100 results in the degradation of the C-terminal only and leaves the N-terminal portion intact, thus producing the p52 polypeptide. As p100 is most commonly associated with RelB, the activation of this *alternative* pathway results in the nuclear translocation of the p52-RelB dimers and has been suggested to play a central role in the expression of genes involved in developing and maintaining secondary lymphoid organs⁴⁰.

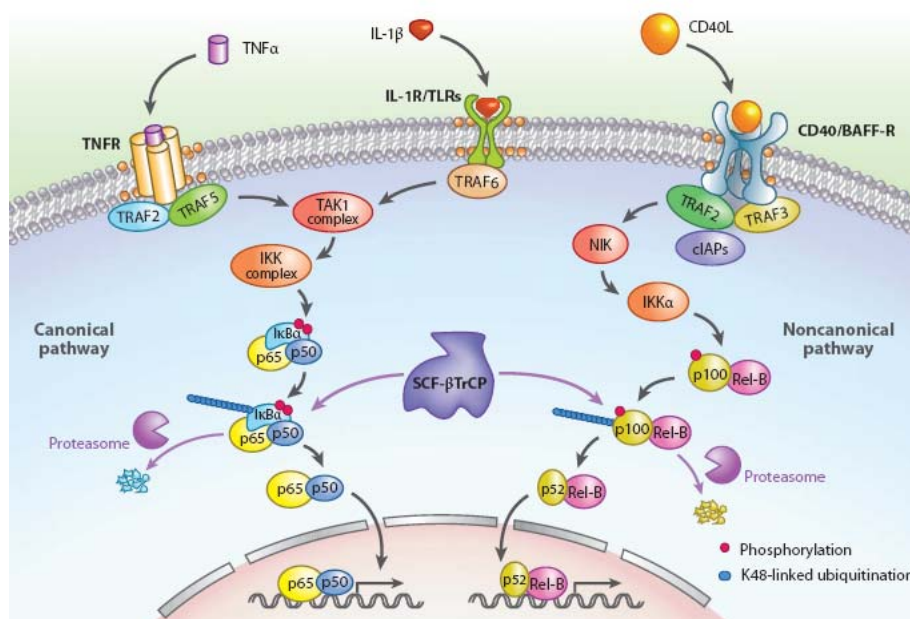


Figure 1. Two pathways lead to NF-κB activation. On the left, the canonical NF-κB activation pathway is shown. In that pathway, the stimulation of the TNF receptor (TNFR), the IL-1 receptor (IL-1R) and the Toll-like receptors (TLRs) activate the TAK1 complex through TRAF proteins. TAK1 then activates IKK, which in turn phosphorylates IκB proteins and targets them for polyubiquitination by the SCF-βTrCP E3 ligase complex. Once IκB has been ubiquitinated, it is degraded by the proteasome, allowing the p50/p65 NF-κB dimer to enter the nucleus and activate gene transcription. On the right, the non-canonical NF-κB pathway is shown. In that pathway, the stimulation of a subset of receptors, including the BAFF receptor, leads to the stabilization of the kinase NIK, followed by the activation of IKKα that phosphorylates p100 and leads to its ubiquitination by the SCF-βTrCP complex. Once p100 has been ubiquitinated, it is targeted for proteasomal processing to p52. Then, the p52/REL-B dimer translocates into the nucleus to activate gene transcription⁴¹.

The elucidation of the key role of activating IKK-2 in cell proliferation, inflammation and the inhibition of insulin signaling (see Figure 2) has led pharmaceutical companies to undertake a great effort to develop specific inhibitors for this enzyme from several chemical families (*e.g.*, pyrimidines, quinazolines, pyridines, thiophenecarboxamides, benzimidazoles, indoles and

carboline derivatives)^{36, 42-45}. In this thesis, we present our own efforts in the discovery of novel natural and selective inhibitors of IKK-2.

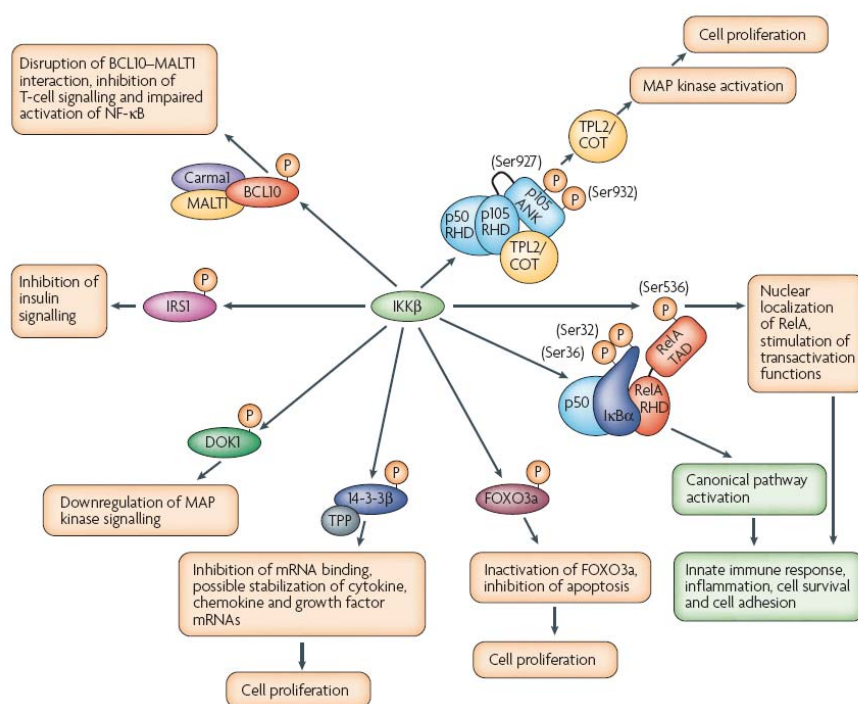


Figure 2. Consequences of IKK-2 activation. In addition to activating the canonical NF- κ B-signaling pathway through phosphorylation (P) of the I κ B proteins, IKK-2 phosphorylates affect several other substrates, including NF- κ B subunits. For example, the phosphorylation of RelA at Ser536 can result in the nuclear localization and stimulation of its transactivation functions by promoting interactions with co-activator proteins. In some unactivated cells, the p105 NF- κ B subunit is found in a complex with the mitogenactivated protein (MAP)/extracellular signal-regulated kinase (ERK) (MEK) kinase TPL2 (also known as COT). IKK-2 phosphorylates p105, resulting in its degradation. This degradation causes the release of TPL2, resulting in the activation of the pro-proliferative MAP kinase-signaling pathway. Another way in which IKK-2 might exert its anti-apoptotic effects is through the phosphorylation and inhibition of the tumor suppressor FOXO3a. By contrast, the phosphorylation of the adaptor protein DOK1 is thought to contribute to its ability to inhibit MAP kinase signaling and cell proliferation. When complexed with tristetraprolin (TPP), an AU-rich element (ARE) binding protein that regulates mRNA stability, IKK-2 phosphorylates and inhibits the 14-3-3 β protein. IKK-2 phosphorylation inhibits TPP-14-3-3 β ARE binding and might therefore stabilize cytokine, chemokine and growth factor transcripts. IKK-2 can also inhibit insulin signaling by targeting insulin-receptor substrate-1 (IRS1). The phosphorylation of BCL10 as part of the T-cell receptor (TCR) Carma1 BCL10-MALT1 complex disrupts the BCL10-MALT1 interaction, therefore attenuating T-cell signaling. As TCR signaling activates the NF- κ B pathway, BCL10 phosphorylation by IKK-2 serves as a negative feedback loop⁴⁶.

The human 11 β -hydroxysteroid dehydrogenase type 1

Glucocorticoids (GCs) are well-known ubiquitous hormones that play a key role in modulating immune and inflammatory responses, regulating energy metabolism, in cardiovascular homeostasis and in the body's responses to stress by binding to the glucocorticoid receptor (GR; a member of the nuclear receptor superfamily)³³. Thus, it has been stated that GCs modulate the expression of up to 20% of the genes in the mammalian genome⁴⁷. Normalization of GC levels can reverse metabolic syndrome by improving obesity, insulin resistance, hypertension and the lipid and lipoprotein profiles⁴⁸. In contrast to the action of insulin, GCs stimulate the production of glucose, switching the homeostatic balance toward catabolism. Thus, GCs promote gluconeogenesis but inhibit β -cell insulin secretion and peripheral glucose uptake. They also increase protein breakdown and lipolysis with consequent fatty acid mobilization. The noticeable resemblance between the symptoms of the metabolic syndrome and the symptoms of hypercortisolism or Cushing syndrome (*i.e.*, obesity, insulin resistance, hypertension and an unfavorable lipid and lipoprotein profile)³³ have initiated intensive investigations on the potential role of GCs in the current obesity epidemic⁴⁹.

GCs have become some of the most important and frequently prescribed drugs for the control of acute and chronic inflammation, allergies and autoimmune diseases, the treatment of leukemia and lymphomas and in the prevention of graft rejection after organ transplantation. The development of novel steroids and improved therapeutic strategies based on GCs is therefore highly desirable, but it is clear that such a development will depend on a greater understanding of how GCs act at the molecular level.

The principal GC is cortisol, whose concentration in target tissues together with the activation of GR is tightly regulated by tissue-specific enzymes called 11 β -hydroxysteroid dehydrogenases. Thus, corticosteroid 11- β -dehydrogenase isozyme type 1 (11 β -HSD1) is a NADPH-dependent enzyme that is predominantly expressed in liver, adipose tissue and in skeletal muscles, where it increases intracellular GC action by catalyzing the reduction of inactive 11-ketoglucocorticoids (*i.e.*, cortisone in human) to active 11 β -hydroxyglucocorticoids (*i.e.*, cortisol in human). On the other hand, corticosteroid 11- β -dehydrogenase isozyme type 2 (11 β -HSD2) is a NAD⁺-dependent enzyme that is highly expressed in classical aldosterone selective target tissues (like distal nephron, colon, sweat glands and the placenta), where it catalyzes the opposite reaction to that of 11 β -HSD1 and thus prevents the binding of cortisol to the mineralocorticoid receptor and subsequently protects it from GCs⁵⁰. These two isoenzymes belong to one of the largest enzyme superfamilies, the short-chain dehydrogenases/reductases (SDR)^{51,52}.

Serum GCs readily pass through cell membranes and exert their intracellular functions by binding to the GR. However, GR only binds with high affinity to the GCs reduced form, resulting in the dissociation of heat shock protein (HSP) complexes and subsequent phosphorylation⁴⁹. GR then translocates to the nucleus, where it dimerizes and either interacts with other transcription factors or binds to glucocorticoid response elements (GREs) upstream of GR-regulated genes (*e.g.*, IκB). These interactions forms a cascade that involves important kinases and interleukines, including IKK-2, IL-1, protein kinase (MAPK), kinase (MKK)4/7 and c-Jun amino-terminal kinase (JNK)⁵³ (see Figure 3).

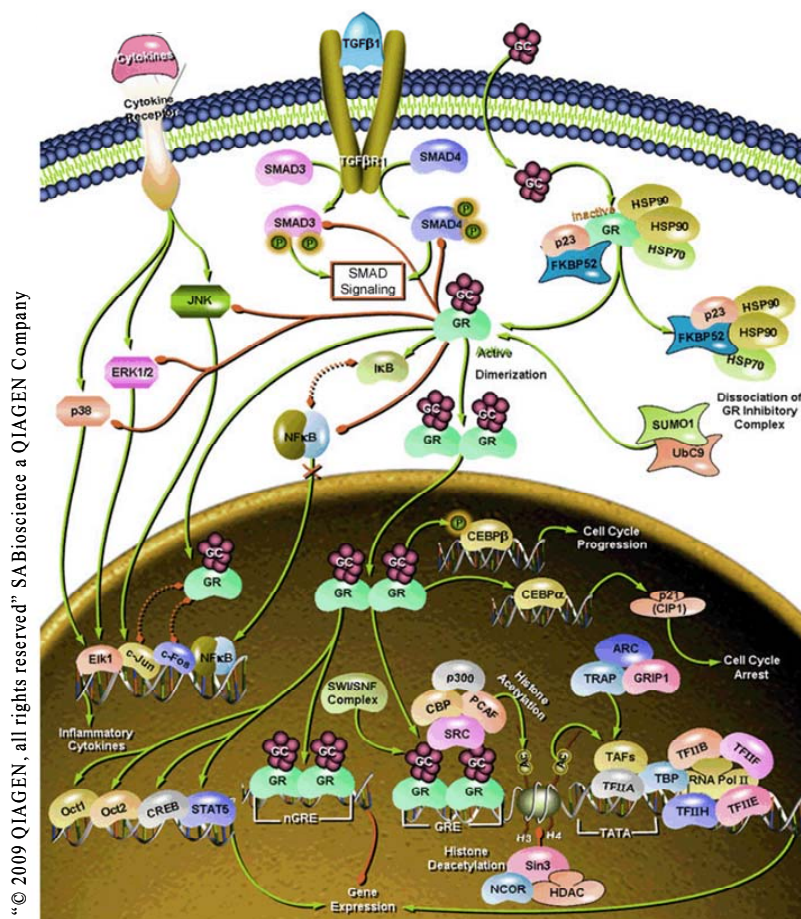


Figure 3. GR activation and its cascade effects.

A growing body of evidence suggests that increased 11β-HSD1 activity within target tissues may promote insulin resistance, obesity, hypertension and dyslipidemia (see Figure 4). Therefore, looking for 11β-HSD1 inhibitors has been, in recent years, an active area of study^{33, 49, 54-64}.

Unfortunately, there are many examples of 11β -HSD1 inhibitors that cause side effects like sodium retention and hypertension because they also inhibit 11β -HSD2^{33, 65-68}. Thus, an important task when finding new 11β -HSD1 inhibitors consists in evaluating their activity on 11β -HSD2 by means of *in vitro* or *in vivo* assays^{58, 69-73}.

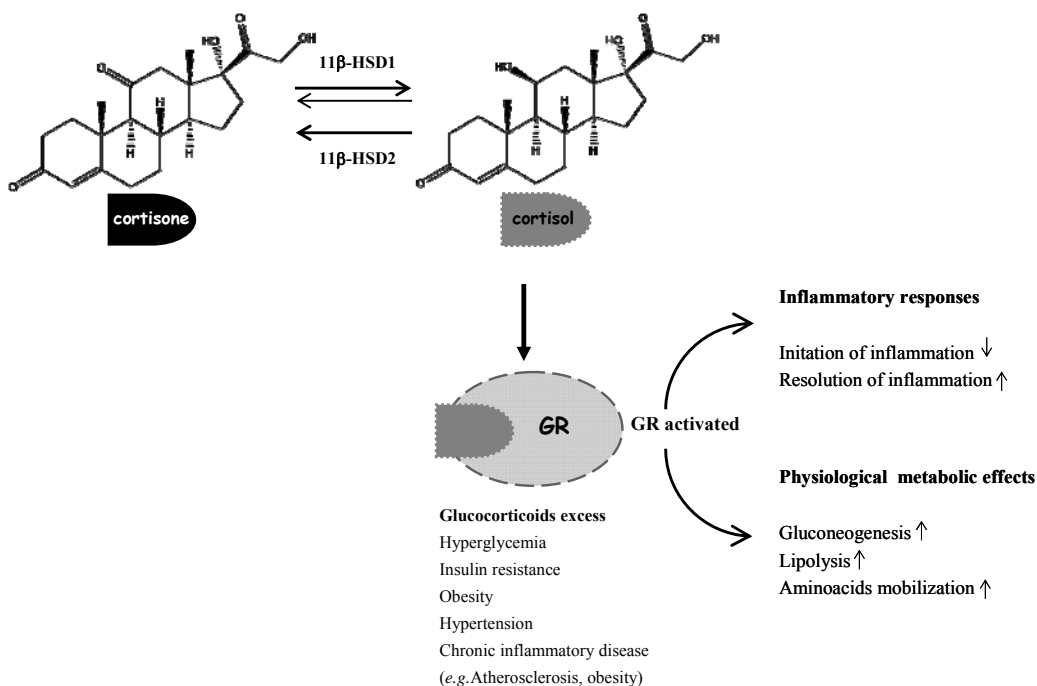


Figure 4. 11β -HSD1 reduces receptor-inactive cortisone to cortisol, which can bind to GR. The activated GR modulates the expression of numerous genes involved in nutrient metabolism and inflammation.

References

1. Terra, X.; Pallarés, V.; Ardèvol, A.; Bladé, C.; Fernández-Larrea, J.; Pujadas, G.; Salvadó, J.; Arola, L.; Blay, M. Modulatory effect of grape-seed procyanidins on local and systemic inflammation in diet-induced obesity rats. *J Nutr Biochem* **2010**.
2. Baiges, I.; Palmfeldt, J.; Bladé, C.; Gregersen, N.; Arola, L. Lipogenesis is decreased by grape seed proanthocyanidins according to liver proteomics of rats fed a high fat diet. *Mol Cell Proteomics* **2010**, 9, 1499-513.
3. Montagut, G.; Bladé, C.; Blay, M.; Fernández-Larrea, J.; Pujadas, G.; Salvadó, M. J.; Arola, L.; Pinent, M.; Ardèvol, A. Effects of a grape seed procyanidin extract (GSPE) on insulin resistance. *J Nutr Biochem* **2010**, 21, 961-7.
4. Quesada, H.; del Bas, J. M.; Pajuelo, D.; Díaz, S.; Fernandez-Larrea, J.; Pinent, M.; Arola, L.; Salvadó, M. J.; Bladé, C. Grape seed proanthocyanidins correct dyslipidemia associated with a high-fat diet in rats and repress genes controlling lipogenesis and VLDL assembling in liver. *Int J Obes (Lond)* **2009**, 33, 1007-12.
5. Roberfroid, M. B. Concepts and strategy of functional food science: the European perspective. *Am J Clin Nutr* **2000**, 71, 1660S-4S; discussion 1674S-5S.
6. Del Bas, J. M.; Ricketts, M. L.; Vaque, M.; Sala, E.; Quesada, H.; Ardevol, A.; Salvado, M. J.; Blay, M.; Arola, L.; Moore, D. D.; Pujadas, G.; Fernandez-Larrea, J.; Blade, C. Dietary procyanidins enhance transcriptional activity of bile acid-activated FXR in vitro and reduce triglyceridemia in vivo in a FXR-dependent manner. *Mol Nutr Food Res* **2009**, 53, 805-14.
7. Terra, X.; Valls, J.; Vitrac, X.; Merrillon, J.-M.; Arola, L.; Ardevol, A.; Blade, C.; Fernandez-Larrea, J.; Pujadas, G.; Salvado, J.; Blay, M. Grape-seed procyanidins act as antiinflammatory agents in endotoxin-stimulated RAW 264.7 macrophages by inhibiting NFκB signaling pathway. *J Agric Food Chem* **2007**, 55, 4357-65.
8. Del Bas, J. M.; Fernández-Larrea, J.; Blay, M.; Ardèvol, A.; Salvadó, M. J.; Arola, L.; Bladé, C. Grape seed procyanidins improve atherosclerotic risk index and induce liver CYP7A1 and SHP expression in healthy rats. *FASEB J* **2005**, 19, 479-81.
9. Puiggros, F.; Llopiz, N.; Ardevol, A.; Blade, C.; Arola, L.; Salvado, M. J. Grape seed procyanidins prevent oxidative injury by modulating the expression of antioxidant enzyme systems. *J Agric Food Chem* **2005**, 53, 6080-6.
10. Justfood. Global market review of functional foods-forecasts to 2012. 84 p. Available from: http://www.alacrastore.com/acm/2023_sample.pdf **2006**.
11. Siró, I.; Kápolna, E.; Kápolna, B.; Lugasi, A. Functional food. Product development, marketing and consumer acceptance--a review. *Appetite* **2008**, 51, 456-67.

12. Koppen, H. Virtual screening - What does it give us? *Current Opinion in Drug Discovery & Development* **2009**, 12, 397-407.
13. Lewis, S. N.; Bassaganya-Riera, J.; Bevan, D. R. Virtual Screening as a Technique for PPAR Modulator Discovery. *Ppar Research* **2010**.
14. Cerqueira, N. M.; Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Virtual screening of compound libraries. *Methods Mol Biol* **2009**, 572, 57-70.
15. Tulp, M.; Bohlin, L. Rediscovery of known natural compounds: nuisance or goldmine? *Bioorg Med Chem* **2005**, 13, 5274-82.
16. Balandrin, M. F.; Kinghorn, A. D.; Farnsworth, N. R. PLANT-DERIVED NATURAL-PRODUCTS IN DRUG DISCOVERY AND DEVELOPMENT - AN OVERVIEW. *Acs Symposium Series* **1993**, 534, 2-12.
17. Rollinger, J. M.; Stuppner, H.; Langer, T. Virtual screening for the discovery of bioactive natural products. In *Natural Compounds as Drugs*, Petersen, F.; Amstutz, R., Eds. Birkhäuser Basel: Basel, 2008; Vol. I, pp 211-249.
18. Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the last 25 years. *Journal of Natural Products* **2007**, 70, 461-477.
19. Schuster, D.; Wolber, G. Identification of bioactive natural products by pharmacophore-based virtual screening. *Curr Pharm Des* **2010**, 16, 1666-81.
20. Martinez-Mayorga, K.; Medina-Franco, J. L. Chemoinformatics-applications in food chemistry. *Adv Food Nutr Res* **2009**, 58, 33-56.
21. David A Evans, J. B. H. a. S. D. Phenolics, inflammation and nutrigenomics. In *Journal of the Science of Food and Agriculture*: 2006; Vol. 86, pp 2503-2509.
22. Egner, M. S. a. U. Structural Aspects of Drugability and Selectivity of Protein Kinases in Inflammation. In *Anti-Inflammatory & Anti-Allergy Agents in Medicinal Chemistry*: 2007; Vol. 6, pp 5-17.
23. Mathieu, P.; Lemieux, I.; Després, J. P. Obesity, inflammation, and cardiovascular risk. *Clin Pharmacol Ther* **2010**, 87, 407-16.
24. Reilly, M.; Rader, D. The metabolic syndrome: more than the sum of its parts? *Circulation* **2003**, 108, 1546-51.
25. Hotamisligil, G. S.; Erbay, E. Nutrient sensing and inflammation in metabolic diseases. *Nat Rev Immunol* **2008**, 8, 923-34.
26. Roche, H.; Phillips, C.; Gibney, M. The metabolic syndrome: the crossroads of diet and genetics. *Proc Nutr Soc* **2005**, 64, 371-7.
27. Hotamisligil, G. Inflammation and metabolic disorders. *Nature* **2006**, 444, 860-7.
28. Mach, F. Inflammation is a crucial feature of atherosclerosis and a potential target to reduce cardiovascular events. In *Atherosclerosis: Diet and Drugs*, 2005; Vol. 170, pp 697-722.

29. Libby, P. Inflammation and cardiovascular disease mechanisms. *Am J Clin Nutr* **2006**, 83, 456S-460S.
30. Das, S. K.; Chakrabarti, R. Antiobesity therapy: emerging drugs and targets. *Curr Med Chem* **2006**, 13, 1429-60.
31. Hotamisligil, G. S. Role of endoplasmic reticulum stress and c-Jun NH₂-terminal kinase pathways in inflammation and origin of obesity and diabetes. *Diabetes* **2005**, 54 Suppl 2, S73-8.
32. Hotamisligil, G. S. Endoplasmic reticulum stress and inflammation in obesity and type 2 diabetes. *Novartis Found Symp* **2007**, 286, 86-94; discussion 94-8, 162-3, 196-203.
33. Wamil, M.; Seckl, J. R. Inhibition of 11beta-hydroxysteroid dehydrogenase type 1 as a promising therapeutic target. *Drug Discov Today* **2007**, 12, 504-20.
34. Tak, P. P.; Firestein, G. S. NF-kappaB: a key role in inflammatory diseases. *J Clin Invest* **2001**, 107, 7-11.
35. Karin, M.; Delhase, M. The I kappa B kinase (IKK) and NF-kappa B: key elements of proinflammatory signalling. *Semin Immunol* **2000**, 12, 85-98.
36. Ziegelbauer, K.; Gantner, F.; Lukacs, N.; Berlin, A.; Fuchikami, K.; Niki, T.; Sakai, K.; Inbe, H.; Takeshita, K.; Ishimori, M.; Komura, H.; Murata, T.; Lowinger, T.; Bacon, K. A selective novel low-molecular-weight inhibitor of IkappaB kinase-beta (IKK-beta) prevents pulmonary inflammation and shows broad anti-inflammatory activity. *Br J Pharmacol* **2005**, 145, 178-92.
37. Brand, K.; Page, S.; Rogler, G.; Bartsch, A.; Brandl, R.; Knuechel, R.; Page, M.; Kaltschmidt, C.; Baeuerle, P. A.; Neumeier, D. Activated transcription factor nuclear factor-kappa B is present in the atherosclerotic lesion. *J Clin Invest* **1996**, 97, 1715-22.
38. Mireille Delhase, M. H. a. Y. C., Michael Karin. Positive and Negative Regulation of Ikb Kinase Activity Through IKKb Subunit Phosphorylation. In SCIENCE: 1999.
39. Strnad, J.; Burke, J. R. IkappaB kinase inhibitors for treating autoimmune and inflammatory disorders: potential and challenges. *Trends Pharmacol Sci* **2007**, 28, 142-8.
40. Brasier, A. R. The NF-kappaB regulatory network. *Cardiovasc Toxicol* **2006**, 6, 111-30.
41. Skaug, B.; Jiang, X.; Chen, Z. J. The role of ubiquitin in NF-kappaB regulatory pathways. *Annu Rev Biochem* **2009**, 78, 769-96.
42. Murata, T.; Shimada, M.; Sakakibara, S.; Yoshino, T.; Masuda, T.; Shintani, T.; Sato, H.; Koriyama, Y.; Fukushima, K.; Nunami, N.; Yamauchi, M.; Fuchikami, K.; Komura, H.; Watanabe, A.; Ziegelbauer, K. B.; Bacon, K. B.; Lowinger, T. B. Synthesis and structure-activity relationships of novel IKK-beta inhibitors. Part 3: Orally active anti-inflammatory agents. *Bioorganic & Medicinal Chemistry Letters* **2004**, 14, 4019-4022.
43. Karin, M.; Yamamoto, Y.; Wang, Q. M. The IKK NF-kappa B system: a treasure trove for drug development. *Nat Rev Drug Discov* **2004**, 3, 17-26.

44. Sugiyama, H.; Yoshida, M.; Mori, K.; Kawamoto, T.; Sogabe, S.; Takagi, T.; Oki, H.; Tanaka, T.; Kimura, H.; Ikeura, Y. Synthesis and structure activity relationship studies of benzothieno[3,2-b]furan derivatives as a novel class of IKKbeta inhibitors. *Chem Pharm Bull (Tokyo)* **2007**, *55*, 613-24.
45. Coish, P. D. G.; Wickens, P. L.; Lowinger, T. B. Small molecule inhibitors of IKK kinase activity. *Expert Opinion on Therapeutic Patents* **2006**, *16*, 1-12.
46. Perkins, N. D. Integrating cell-signalling pathways with NF-kappaB and IKK function. *Nat Rev Mol Cell Biol* **2007**, *8*, 49-62.
47. Galon, J.; Franchimont, D.; Hiroi, N.; Frey, G.; Boettner, A.; Ehrhart-Bornstein, M.; O'Shea, J. J.; Chrousos, G. P.; Bornstein, S. R. Gene profiling reveals unknown enhancing and suppressive actions of glucocorticoids on immune cells. *FASEB J* **2002**, *16*, 61-71.
48. Inagaki, K.; Otsuka, F.; Miyoshi, T.; Watanabe, N.; Suzuki, J.; Ogura, T.; Makino, H. Reversible pituitary dysfunction in a patient with Cushing's syndrome discovered as adrenal incidentaloma. *Endocr J* **2004**, *51*, 201-6.
49. Staab, C. A.; Maser, E. 11beta-Hydroxysteroid dehydrogenase type 1 is an important regulator at the interface of obesity and inflammation. *J Steroid Biochem Mol Biol* **2010**, *119*, 56-72.
50. Seckl JR, W. B. Minireview: 11beta-hydroxysteroid dehydrogenase type 1- a tissue-specific amplifier of glucocorticoid action. *Endocrinology* **2001**, *142*, 1371-6.
51. Bray, J. E.; Marsden, B. D.; Oppermann, U. The human short-chain dehydrogenase/reductase (SDR) superfamily: A bioinformatics summary. *Chemico-Biological Interactions* **2009**, *178*, 99-109.
52. Kavanagh, K. L.; Jornvall, H.; Persson, B.; Oppermann, U. Medium- and short-chain dehydrogenase/reductase gene and protein families : the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell Mol Life Sci* **2008**, *65*, 3895-906.
53. Pace, T. W.; Miller, A. H. Cytokines and glucocorticoid receptor signaling. Relevance to major depression. *Ann N Y Acad Sci* **2009**, *1179*, 86-105.
54. Morton, N. M. Obesity and corticosteroids: 11beta-hydroxysteroid type 1 as a cause and therapeutic target in metabolic disease. *Mol Cell Endocrinol* **2010**, *316*, 154-64.
55. Zhu, Y.; Olson, S. H.; Hermanowski-Vosatka, A.; Mundt, S.; Shah, K.; Springer, M.; Thieringer, R.; Wright, S.; Xiao, J.; Zokian, H.; Balkovec, J. M. 4-Methyl-5-phenyl triazoles as selective inhibitors of 11beta-hydroxysteroid dehydrogenase type I. *Bioorg Med Chem Lett* **2008**, *18*, 3405-11.
56. Rosenstock, J.; Banarer, S.; Fonseca, V. A.; Inzucchi, S. E.; Sun, W.; Yao, W.; Hollis, G.; Flores, R.; Levy, R.; Williams, W. V.; Seckl, J. R.; Huber, R.; Investigators, I.-P. The 11-beta-

hydroxysteroid dehydrogenase type 1 inhibitor INCB13739 improves hyperglycemia in patients with type 2 diabetes inadequately controlled by metformin monotherapy. *Diabetes Care* **2010**, *33*, 1516-22.

57. Tiwari, A. INCB-13739, an 11beta-hydroxysteroid dehydrogenase type 1 inhibitor for the treatment of type 2 diabetes. *IDrugs* **2010**, *13*, 266-75.

58. Jean, D. J. S.; Yuan, C.; Bercot, E. A.; Cupples, R.; Chen, M.; Fretland, J.; Hale, C.; Hungate, R. W.; Komorowski, R.; Veniant, M.; Wang, M. H.; Zhang, X. P.; Fotsch, C. 2-(S)-phenethylaminothiazolones as potent, orally efficacious inhibitors of 11 beta-hydroxysteroid dehydrogenase type 1. *Journal of Medicinal Chemistry* **2007**, *50*, 429-432.

59. Schnackenberg, C. G. 11beta-hydroxysteroid dehydrogenase type 1 inhibitors for metabolic syndrome. *Curr Opin Investig Drugs* **2008**, *9*, 295-300.

60. St Jean, D. J.; Wang, M.; Fotsch, C. Inhibitors of 11beta-HSD1: a potential treatment for the metabolic syndrome. *Curr Top Med Chem* **2008**, *8*, 1508-23.

61. Ge, R.; Huang, Y.; Liang, G.; Li, X. 11beta-hydroxysteroid dehydrogenase type 1 inhibitors as promising therapeutic drugs for diabetes: status and development. *Curr Med Chem* **2010**, *17*, 412-22.

62. Boyle, C. D.; Kowalski, T. J. 11beta-hydroxysteroid dehydrogenase type 1 inhibitors: a review of recent patents. *Expert Opin Ther Pat* **2009**, *19*, 801-25.

63. Rollinger, J. M.; Kratschmar, D. V.; Schuster, D.; Pfisterer, P. H.; Gummy, C.; Aubry, E. M.; Brandstötter, S.; Stuppner, H.; Wolber, G.; Odermatt, A. 11beta-Hydroxysteroid dehydrogenase 1 inhibiting constituents from *Eriobotrya japonica* revealed by bioactivity-guided isolation and computational approaches. *Bioorg Med Chem* **2010**, *18*, 1507-15.

64. Gummy, C.; Thurnbichler, C.; Aubry, E. M.; Balazs, Z.; Pfisterer, P.; Baumgartner, L.; Stuppner, H.; Odermatt, A.; Rollinger, J. M. Inhibition of 11beta-hydroxysteroid dehydrogenase type 1 by plant extracts used as traditional antidiabetic medicines. *Fitoterapia* **2009**, *80*, 200-5.

65. Wang, H.; Ruan, Z.; Li, J. J.; Simpkins, L. M.; Smirk, R. A.; Wu, S. C.; Hutchins, R. D.; Nirschl, D. S.; Van Kirk, K.; Cooper, C. B.; Sutton, J. C.; Ma, Z.; Golla, R.; Seethala, R.; Salyan, M. E.; Nayeem, A.; Krystek, S. R.; Sheriff, S.; Camac, D. M.; Morin, P. E.; Carpenter, B.; Robl, J. A.; Zahler, R.; Gordon, D. A.; Hamann, L. G. Pyridine amides as potent and selective inhibitors of 11beta-hydroxysteroid dehydrogenase type 1. *Bioorg Med Chem Lett* **2008**, *18*, 3168-72.

66. Rohde, J. J.; Pliushchev, M. A.; Sorensen, B. K.; Wodka, D.; Shuai, Q.; Wang, J.; Fung, S.; Monzon, K. M.; Chiou, W. J.; Pan, L.; Deng, X.; Chovan, L. E.; Ramaiya, A.; Mullally, M.; Henry, R. F.; Stolarik, D. F.; Imade, H. M.; Marsh, K. C.; Beno, D. W.; Fey, T. A.; Droz, B. A.; Brune, M. E.; Camp, H. S.; Sham, H. L.; Frevert, E. U.; Jacobson, P. B.; Link, J. T. Discovery and metabolic stabilization of potent and selective 2-amino-N-(adamant-2-yl) acetamide 11beta-hydroxysteroid dehydrogenase type 1 inhibitors. *J Med Chem* **2007**, *50*, 149-64.

67. Walker, E. A.; Stewart, P. M. 11beta-hydroxysteroid dehydrogenase: unexpected connections. *Trends Endocrinol Metab* **2003**, 14, 334-9.
68. Ferrari, P. The role of 11 β -hydroxysteroid dehydrogenase type 2 in human hypertension. *Biochim Biophys Acta* **2010**, 1802, 1178-87.
69. Wan, Z. K.; Chenail, E.; Xiang, J.; Li, H. Q.; Ipek, M.; Bard, J.; Svenson, K.; Mansour, T. S.; Xu, X.; Tian, X.; Suri, V.; Hahm, S.; Xing, Y.; Johnson, C. E.; Li, X.; Qadri, A.; Panza, D.; Perreault, M.; Tobin, J. F.; Saiah, E. Efficacious 11beta-hydroxysteroid dehydrogenase type I inhibitors in the diet-induced obesity mouse model. *J Med Chem* **2009**, 52, 5449-61.
70. Julian, L. D.; Wang, Z.; Bostick, T.; Caille, S.; Choi, R.; DeGraffenreid, M.; Di, Y.; He, X.; Hungate, R. W.; Jaen, J. C.; Liu, J.; Monshouwer, M.; McMinn, D.; Rew, Y.; Sudom, A.; Sun, D.; Tu, H.; Ursu, S.; Walker, N.; Yan, X.; Ye, Q.; Powers, J. P. Discovery of novel, potent benzamide inhibitors of 11beta-hydroxysteroid dehydrogenase type 1 (11beta-HSD1) exhibiting oral activity in an enzyme inhibition ex vivo model. *J Med Chem* **2008**, 51, 3953-60.
71. Beseda, I.; Czollner, L.; Shah, P. S.; Khunt, R.; Gaware, R.; Kosma, P.; Stanetty, C.; Del Ruiz-Ruiz, M. C.; Amer, H.; Mereiter, K.; Da Cunha, T.; Odermatt, A.; Classen-Houben, D.; Jordis, U. Synthesis of glycyrrhetic acid derivatives for the treatment of metabolic diseases. *Bioorg Med Chem* **2010**, 18, 433-54.
72. Yang, H.; Shen, Y.; Chen, J.; Jiang, Q.; Leng, Y.; Shen, J. Structure-based virtual screening for identification of novel 11beta-HSD1 inhibitors. *Eur J Med Chem* **2009**, 44, 1167-71.
73. Classen-Houben, D.; Schuster, D.; Da Cunha, T.; Odermatt, A.; Wolber, G.; Jordis, U.; Kuenburg, B. Selective inhibition of 11beta-hydroxysteroid dehydrogenase 1 by 18alpha-glycyrrhetic acid but not 18beta-glycyrrhetic acid. *J Steroid Biochem Mol Biol* **2009**, 113, 248-52.

II. CONTEXT AND GOALS

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

One of the main challenges in nutrigenomics is improving the efficiency of the selection (which is currently time consuming and expensive) of new bioactive compounds in order to expedite the development of new functional foods. Computational techniques, such as virtual screening, may play an essential role in accelerating the early stages of the discovery of new bioactive substances by efficiently searching for compounds that could activate or inhibit a known target.

The **initial hypothesis** of this PhD thesis is that, by modulating specific target functions in the body, molecules that act as IKK-2 or 11 β -HSD1 inhibitors have beneficial physiological effects that can be of interest for preventing, retarding and/or reversing the metabolic syndrome. Because of their anti-inflammatory properties, natural extracts that contain these molecules have a promising role as ingredients in new functional foods.

Therefore, this PhD thesis will focus on the development of virtual screening workflows that predict with a high level of success natural products that can inhibit IKK-2 or 11 β -HSD1. To achieve this goal, the following steps must be accomplished:

- Use virtual screening to identify potential hIKK-2 inhibitors of natural origin that can compete with ATP.
- Evaluate the reliability of the virtual screening protocol by experimentally testing the *in vitro* activity of selected NP hits on hIKK-2.
- Find natural extracts with known anti-inflammatory activity that contain at least one molecule that has been predicted to be an hIKK-2 inhibitor and find hits from the virtual screening protocol that have yet to be described as hIKK-2 inhibitors and that are present in such anti-inflammatory extracts.
- Contribute to knowledge of the structural-activity relationships of hIKK-2 inhibitors by using pyridine derivatives for developing a 3D-QSAR model that (a) correlates the experimental activity with the structures of the inhibitors used to develop the model and (b) is able to predict the IC₅₀ of a large set of other pyridine derivatives not used during the development of the model.
- Develop a structure-based *in silico* methodology that can be used for discerning in a molecular database between molecules that inhibit 11 β -HSD2 and those that do not.
- Use virtual screening to identify potential 11 β -HSD1 inhibitors of natural origin that do not inhibit 11 β -HSD2.

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

III. RESULTS

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

Inhibitor of nuclear factor kappa- B kinase subunit beta (hIKK-2)

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

Manuscript 1

Identification of Human IKK-2 Inhibitors of Natural Origin (part I): Modeling of the IKK-2 Kinase Domain, Virtual Screening and Activity Assays

Esther Sala, Laura Guasch, Justyna Iwaszkiewicz, Miquel Mulero, Maria-Josepa Salvadó, Montserrat Pinent, Vincent Zoete, Aurélien Grosdidier, Santiago Garcia-Vallvé, Olivier Michielin and Gerard Pujadas

PLoS ONE 6(2): e16903. doi:10.1371/journal.pone.0016903

(February 2011)

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

Identification of Human IKK-2 Inhibitors of Natural Origin (Part I): Modeling of the IKK-2 Kinase Domain, Virtual Screening and Activity Assays

Esther Sala¹, Laura Guasch¹, Justyna Iwazskiewicz³, Miquel Mulero¹, Maria-Josepa Salvadó¹, Montserrat Pinet¹, Vincent Zoete³, Aurélien Grosdidier³, Santiago Garcia-Vallvé^{1,2}, Olivier Michielin³, Gerard Pujadas^{1,2*}

1 Grup de Recerca en Nutrigenòmica, Departament de Bioquímica i Biotecnologia, Universitat Rovira i Virgili, Campus de Sescelades, Tarragona, Catalonia, Spain, **2** Centre Tecnològic de Nutrició i Salut, Reus, Catalonia, Spain, **3** Molecular Modeling Group, Swiss Institute of Bioinformatics, Quartier UNIL-Sorge, Lausanne, Switzerland

Abstract

Background: Their large scaffold diversity and properties, such as structural complexity and drug similarity, form the basis of claims that natural products are ideal starting points for drug design and development. Consequently, there has been great interest in determining whether such molecules show biological activity toward protein targets of pharmacological relevance. One target of particular interest is hIKK-2, a serine-threonine protein kinase belonging to the IKK complex that is the primary component responsible for activating NF- κ B in response to various inflammatory stimuli. Indeed, this has led to the development of synthetic ATP-competitive inhibitors for hIKK-2. Therefore, the main goals of this study were (a) to use virtual screening to identify potential hIKK-2 inhibitors of natural origin that compete with ATP and (b) to evaluate the reliability of our virtual-screening protocol by experimentally testing the *in vitro* activity of selected natural-product hits.

Methodology/Principal Findings: We thus predicted that 1,061 out of the 89,425 natural products present in the studied database would inhibit hIKK-2 with good ADMET properties. Notably, when these 1,061 molecules were merged with the 98 synthetic hIKK-2 inhibitors used in this study and the resulting set was classified into ten clusters according to chemical similarity, there were three clusters that contained only natural products. Five molecules from these three clusters (for which no anti-inflammatory activity has been previously described) were then selected for *in vitro* activity testing, in which three out of the five molecules were shown to inhibit hIKK-2.

Conclusions/Significance: We demonstrated that our virtual-screening protocol was successful in identifying lead compounds for developing new inhibitors for hIKK-2, a target of great interest in medicinal chemistry. Additionally, all the tools developed during the current study (i.e., the homology model for the hIKK-2 kinase domain and the pharmacophore) will be made available to interested readers upon request.

Citation: Sala E, Guasch L, Iwazskiewicz J, Mulero M, Salvadó M-J, et al. (2011) Identification of Human IKK-2 Inhibitors of Natural Origin (Part I): Modeling of the IKK-2 Kinase Domain, Virtual Screening and Activity Assays. PLoS ONE 6(2): e16903. doi:10.1371/journal.pone.0016903

Editor: Vladimir Uversky, University of South Florida College of Medicine, United States of America

Received: November 19, 2010; **Accepted:** January 14, 2011; **Published:** February 24, 2011

Copyright: © 2011 Sala et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by Grant Number AGL2008-00387/ALI from the Ministerio de Educación y Ciencia of the Spanish Government, Grant Number ASTF 61-2009 from the EMBO Short-Term Fellowship program and the ACC10 (TECCT10-1-0008) program (Generalitat de Catalunya). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: gerard.pujadas@urv.cat

Introduction

Natural products (NPs) are a valuable source of inspiration as lead compounds for the design and development of new drug candidates [1]. In fact, over 60% of the current anticancer drugs are natural-product-related molecules (i.e., either NPs or derivatives/analogues that have been inspired by a natural compound) [2]. This trend is a direct consequence of the large scaffold diversity of these molecules [2–4], which, together with properties such as structural complexity and drug similarity, support the claim that they are ideal starting points for medicinal chemistry [2,5]. Despite these clear advantages, it is estimated that only 5 to 15% of the approximately 250,000 described higher plant species have been tested for some type of biological activity [4] (and the

use of other organisms such as micro-organisms, insects, fungi or marine species for this aim is just beginning) [6,7]. It has recently been suggested that bioinformatics/chemoinformatics tools could be used to screen large NP databases and so identify new bioactive molecules for specific targets [1,4,8]. One target of interest for NPs is the human inhibitor nuclear-factor κ B kinase 2 (hIKK-2). This target is a serine-threonine protein kinase belonging to the IKK complex and is the primary component responsible for activating the nuclear-factor κ B transcription factor (NF- κ B) in response to inflammatory stimuli. Indeed, the importance of the NF- κ B pathway in regulating the expression of the genes controlling cellular immune and inflammatory responses has motivated research groups in both academia and the pharmaceutical industry to devote increasing efforts toward developing synthetic

ATP-competitive inhibitors for hIKK-2 [9–36], which could be of therapeutic use, e.g., in patients affected by chronic inflammatory diseases [37].

The hIKK-2 subunit is a polypeptide of 756 residues with a kinase domain in its 15–300 sequence segment. At present, there are only two entries in the Protein Data Bank (<http://www.pdb.org>; PDB) [38] for hIKK-2 (*i.e.*, PDB codes 3BRT and 3BRV), and in both cases, they correspond to the NEMO-binding region located at the C-terminal end of the protein [39] (therefore, no experimental structure for the kinase domain of hIKK-2 is known). Moreover, obtaining a homology model for this kinase domain is not trivial because the more similar protein in the PDB (*i.e.*, the catalytic domain of human ZIP kinase; PDB code 1YRP) has a pairwise sequence alignment with hIKK-2 of only 217 residues long, with a low percentage of sequence identity (*i.e.*, 32%). Despite this limitation, the great pharmacological interest of this target has led scientists to develop either ligand-based approaches [14,40–42] or homology models for this domain [13–16,19,27,35,42–44] and to use them in virtual-screening (VS) or drug-design projects. Unfortunately, these pharmacophore hypotheses or homology models are not in the public domain.

Therefore, the main goals of this study were to **(a)** use VS to identify potential hIKK-2 inhibitors of natural origin that can compete with ATP and **(b)** evaluate the reliability of our VS protocol by experimentally testing the *in vitro* activity of selected NP hits. To achieve these goals, we **(1)** developed a homology model for the hIKK-2 kinase domain which could stand the test of our validation criteria, **(2)** docked ATP-competitive molecules known to be potent and specific inhibitors of hIKK-2 with this model [10,11,13,15,16,18,20–31], **(3)** identified which of the resulting poses were *knowledge-based coherent* by analyzing whether they satisfied the experimentally known generic binding features of ATP-competitive inhibitors of kinases [45], **(4)** used the knowledge-based coherent poses to derive a structure-based common pharmacophore containing the key intermolecular interactions between hIKK-2 and its inhibitors, **(5)** obtained exclusion volumes from our homology model and added them to the pharmacophore, **(6)** validated the selectivity of the resulting pharmacophore and of the VS process using a large database of kinase decoys [46] and ATP-competitive inhibitors for hIKK-2 that were not used during the pharmacophore building [47], **(7)** used the previously validated structure-based pharmacophore and VS protocol to find ATP-competitive inhibitors for hIKK-2 in a database of NPs [48], and, finally, **(8)** proved the reliability of the prediction by testing the inhibitory effect of some selected hits on hIKK-2 *in vitro*. In addition, all the tools developed during the current study (*i.e.*, the homology model for the hIKK-2 kinase domain and the pharmacophore) are available to interested readers upon request.

Results and Discussion

Homology-model description

The protein kinase domains have a common fold consisting of two lobes linked by a flexible hinge region (see Figure 1A). Thus, the N-terminal lobe (from Trp15 to Ala95 in hIKK-2) is formed by a β -sheet (consisting of five antiparallel β -strands) and an α -helix that contains Glu61, which is one of the most important residues when hIKK-2 is in the active form because it makes polar contacts with the catalytic residue Lys44 (located in β 3) [49,50]. The N lobe also contains the highly conserved GXGXXG motif called the glycine-rich loop (G-loop) or P-loop, corresponding to the Gly22-Thr23-Gly24-Gly25-Phe26-Gly27 sequence segment in hIKK-2.

The C-terminal lobe (from Gln100 to Phe300 in hIKK-2) is highly helical and contains the activation segment formed by the HRD motif and the activation and catalytic loops. The HRD motif (from His143 to Asp145 in hIKK-2) is highly conserved in the kinases, and its arginine residue acts as the catalytic base that drives the phosphorylation event [50,51]. The activation loop (A-loop; from Lys171 to Leu189 in hIKK-2) is the most flexible and diverse part of the activation segment [50] and includes residue(s) that when phosphorylated enable activity in many kinases (Ser177 and Ser181 for hIKK-2) [49,52]. Finally, the catalytic loop (from Ile165 to Ala170 in hIKK-2) contains the DFG motif (but Asp166, Leu167 and Gly168 for hIKK-2) in most kinases, and its Asp residue forms polar contacts with all three ATP phosphates (either directly or through coordinating magnesium atoms).

The ATP binding site is a narrow, hydrophobic pocket located between the two lobes; its floor is formed by a C-terminal β -sheet, and its roof is formed by the G-loop, which fastens the γ -phosphate of ATP and serves as a regulatory flap above the ATP-binding site [45,51,53–55].

The most important residue in the hinge region of the protein kinase domain (from Met96 to Cys99 in hIKK-2) is the gatekeeper (*i.e.*, Met96 in hIKK-2) [18], which is located behind the hydrophobic pocket of the ATP-binding site. In that location, the size or volume of the side chain dictates its access to this pocket and therefore defines the potential inhibitor selectivity of the ATP-binding site. Moreover, this hinge region usually has a hydrogen donor flanked by hydrogen acceptors derived from the protein backbone, where they form a characteristic hydrogen-bond motif (in fact, the hinge region provides the most important hydrogen bonds with the ATP-site inhibitors). The residues involved in this hydrogen-bond motif are Glu97 and Cys99 in hIKK-2 (called gk+1 and gk+3 considering their relative location to the gatekeeper residue, *i.e.*, gk, in the hIKK-2 sequence) [18].

Figure 1B shows the multiple sequence alignment used to guide the modeling of the best 3D structure obtained for the hIKK-2 kinase domain, which is shown in Figure 1A. Because we were interested in finding new ATP-competitive inhibitors for hIKK-2, we selected two kinase structures as templates (2CN5 [56] and 2J90 [57]), each cocrystallized with a different ATP-competitive inhibitor (ADP and pyridone 6, respectively), both sharing a 32% sequence identity with the kinase domain of hIKK-2 and a 50% identity with the residues of the ATP-binding site of hIKK-2 (where these residues correspond to those in the G-loop, hinge region, HRD motif and catalytic loop).

The Ramachandran plot of this hIKK-2 homology model (see Figure 1C) shows that only 2.5% of the non-glycine residues (*i.e.*, Asn89, Gln110, Cys114, Arg144, Leu160, Thr185, Leu189 and Ala190) are in disallowed regions. Remarkably, none of these residues are located in the ATP-binding-site area and, therefore, its incorrect conformation did not affect our protein-ligand docking results. Moreover, when the ANOLEA results for the hIKK-2 model are compared with those for the two templates (see Table 1), the results show that the packing quality of our homology model is comparable to the packing qualities of the templates. Thus, although ANOLEA reports that 8 out of the 22 hIKK-2 high-energy residues (*i.e.*, from Phe182 to Leu189) are located in the A-loop (a very important area for kinase activity), it also shows that the equivalent residues in the two template structures also have high energies. The rest of high-energy residues in the hIKK-2 homology model are located in areas that are not crucial for its activity or structural stability.

Finally, the quality of this homology model was tested by using it as a rigid receptor in protein-ligand docking experiments with known hIKK-2 inhibitors and evaluating whether the docking

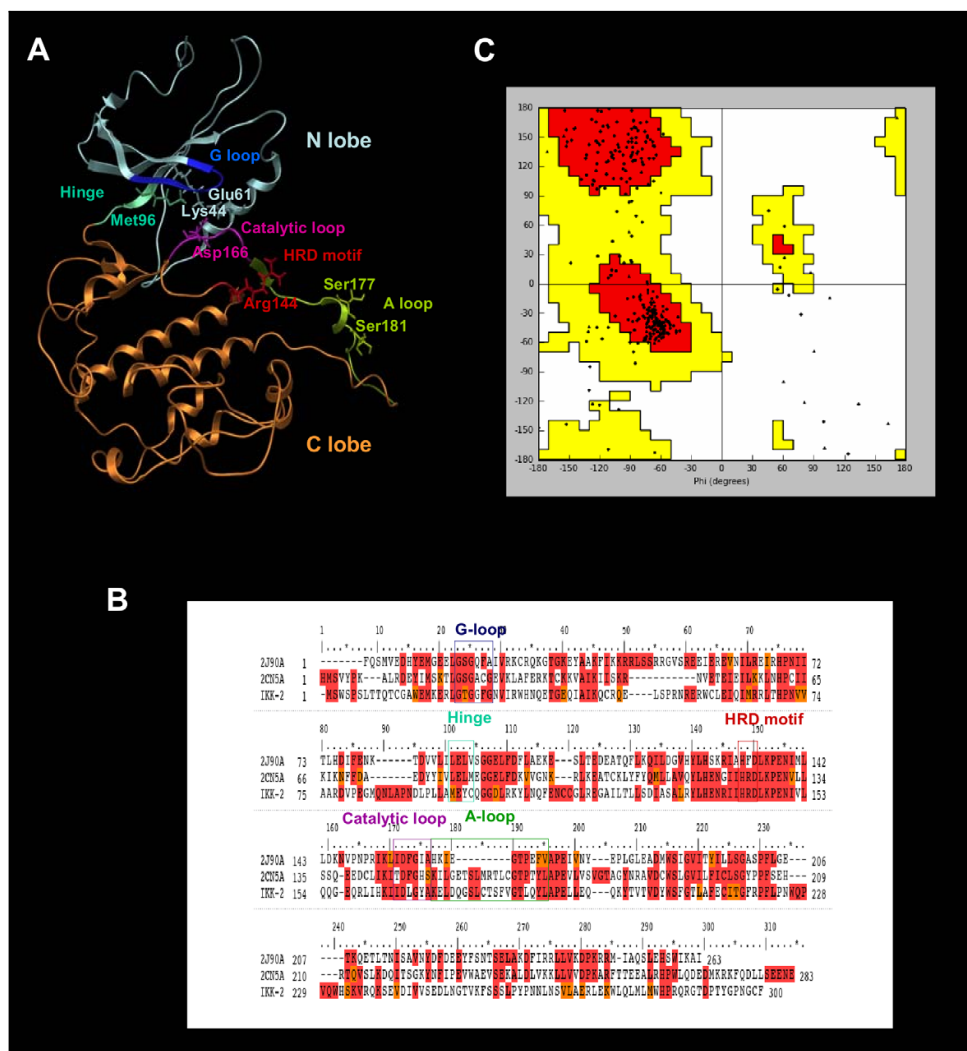


Figure 1. The homology model for the the kinase domain of hIKK-2. The best homology model for the kinase domain of hIKK-2 obtained in our study (Figure 1A), the multiple sequence alignment used to guide its modeling (Figure 1B) and its Ramachandran plot (Figure 1C). The relative locations of the G-loop, hinge region, HRD motif, catalytic loop and A-loop are shown in the model and in the sequence multialignment (using boxes) with the same color schema (i.e., G-loop in blue; hinge region in cyan; HRD motif in red; catalytic loop in magenta and A-loop in green). The residues in the alignment are colored according to residue similarity. Circles in the Ramachandran plot show the location of non-glycine residues.
 doi:10.1371/journal.pone.0016903.g001

poses reproduced the interactions described in the literature as important for kinase inhibitors, specifically, those with key residues in the hIKK-2 ATP binding pocket [13,14,44,45,53,58,59]. Thus, 36 known inhibitors of hIKK-2 (listed in Table S1) were docked with our hIKK-2 homology model without imposing constraints that forced poses to make specific intermolecular interactions with the target. Next, the resultant hIKK-2 complexes were analyzed with the help of LigandScout to determine which complexes exhibit target-inhibitor intermolecular interactions equivalent to those described in prior studies [14,15,31,44]. This knowledge-based analysis enabled us to find at least one knowledge-based coherent pose for 21 out of the 36 hIKK-2 inhibitors assayed

(regardless of their eHits scores). Therefore, we concluded that (a) our best homology model correctly describes the ATP binding pocket in hIKK-2 (see Figures 2A and 2B) and (b) our protein-ligand docking strategy (using the hIKK-2 homology model as a rigid body and only allowing rotation around ligand single bonds of the ligand) is valid for describing the interactions between a large number of known selective ATP-competitive hIKK-2 inhibitors from different chemical families and this target. Notably, this fast protein-ligand docking strategy is appropriate for use in a VS workflow, where other more sophisticated and accurate protein-ligand docking strategies, such as induced-fit docking, are inapplicable [44].

Table 1. ANOLEA results for the two templates (2CN5 and 2J90) and the hIKK-2 homology model.

| | 2CN5A | 2J90A | hIKK-2 |
|---|-----------|------------|-----------|
| Total amino acids with high energy (percentage) | 22 (7.77) | 31 (11.79) | 22 (7.33) |
| Total number of non-local atomic interactions | 32310 | 30226 | 33992 |
| Total non-local energy of the protein (E/kT units) | -1878 | -1946 | -1956 |
| Non-local normalized energy Z-score | 0.04 | -0.49 | 0.09 |

doi:10.1371/journal.pone.0016903.t001

Figures 2C, 2D, 2E and 2F display the knowledge-based coherent docking poses for four selected hIKK-2 inhibitors [*i.e.*, **13** [27] (Figure 2C), **12** [18] (Figure 2D), **4a** [23] (Figure 2E) and **14** [18] (Figure 2F)] that (**a**) belong to three different chemical

families [inhibitors **14** [18] (Figure 2F) and **12** [18] (Figure 2D) share the same structural scaffold], (**b**) have different types of intermolecular interactions with the target and (**c**) share some of the hIKK-2 residues with which they interact (for instance, all of

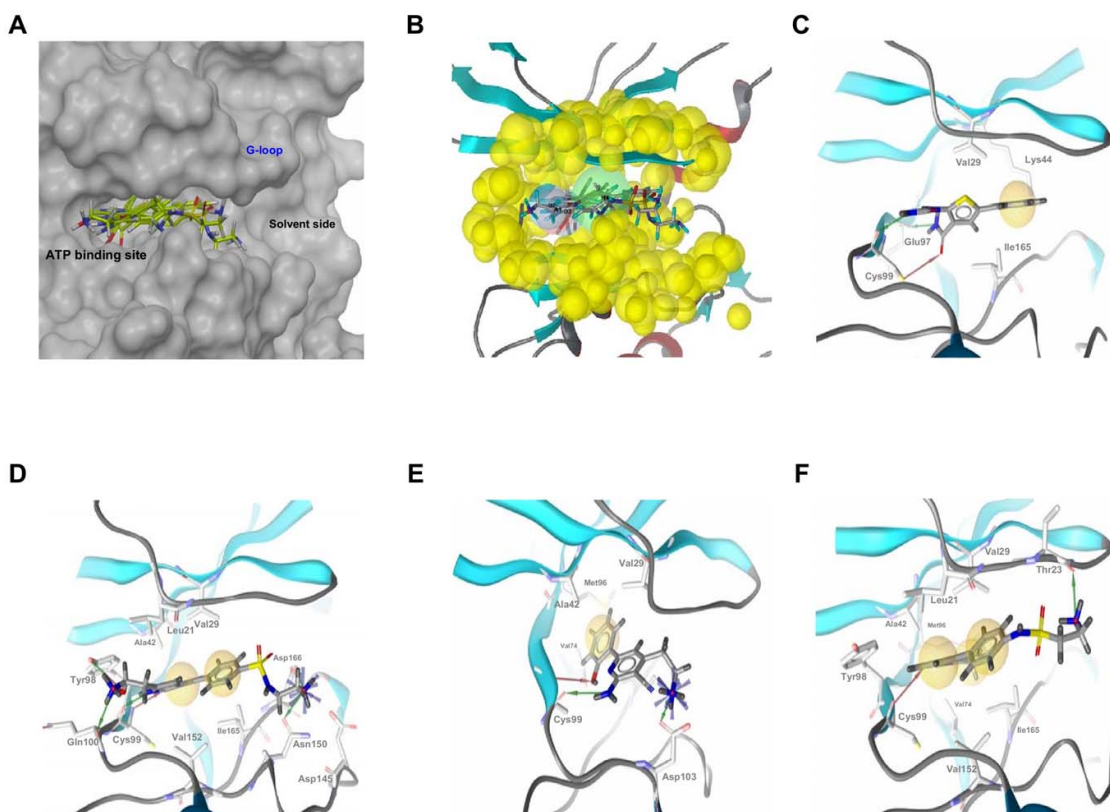


Figure 2. The ATP-binding site, the structure-based common pharmacophore and the best poses for different hIKK-2 inhibitors. The ATP-binding site (Figure 2A), the structure-based common pharmacophore (Figure 2B) and the best poses of the various chemical hIKK-2 inhibitors that were used during the homology-model validation and the structure-based common pharmacophore construction (Figures 2C, 2D, 2E and 2F). In Figures 2C, 2D, 2E and 2F, hydrogen-bond donors, hydrogen-bond acceptors and hydrophobic interactions are shown as red arrows, green arrows and yellow spheres, respectively. Figure 2A shows the ATP-binding pocket of our hIKK-2 homology model and the best poses obtained when docking it with the hIKK-2 inhibitors **13** [27]; **12** [18]; **4a** [23] and **14** [18]. Figure 2B shows the locations on the ATP-binding site of the sites that form the structure-based common pharmacophore developed in the present work and the shell of excluded volumes (in yellow) that schematically represents the locations of the surrounding residues. This pharmacophore is formed by two hydrogen-bond donors (in blue), one hydrogen-bond acceptor (in red) and one hydrophobic region (in green), with tolerances (*i.e.*, radii) of 1.5, 1.5 and 3.0 Å, respectively. Figures 2C, 2D, 2E and 2F show the intermolecular interactions between hIKK-2 and the best poses from inhibitors **13**, **12**, **4a** and **14**, respectively. Figures 2A and 2B were drawn with Maestro v9.0.211 (Schrodinger LLC., Portland, USA; <http://www.schrodinger.com>), whereas Figures 2C, 2D, 2E and 2F were drawn with LigandScout v2.03.

doi:10.1371/journal.pone.0016903.g002

them have a hydrogen-bond interaction with the Cys99 backbone and a hydrophobic interaction with Val29). Thus, the analysis of this figure revealed the following:

- Inhibitor **13** [27] (see Figure 2C) is a thiophene carboxamide with substituted ureas that exhibits the most fundamental intermolecular interactions with the target: **(a)** hydrogen bonding between the urea groups and the main-chain oxygens from Glu97 and Cys99 [27,28,31] in the hinge region (and also between the carbonyl oxygen atom from **13** and the thiol group from Cys99) and **(b)** hydrophobic interaction with residues Val29, Lys44 and Ile165.
- Inhibitor **12** [18] (see Figure 2D) is an azaindole sulfonamide that exhibits more hydrogen bonds with hIKK-2 than the other inhibitors. Specifically, it has **(a)** two hydrogen bonds between the two nitrogen atoms from the azaindole ring and the main-chain N and O from Cys99 (in agreement with the results of *Liddle et al.* [18]), **(b)** one hydrogen bond between its amide group and the side-chain hydroxyl group of Tyr98, **(c)** one hydrogen bond between the nitrogen from the amide group and the backbone oxygen atom of Gln100, and **(d)** one hydrogen bond on the other side of the binding pocket with the Asn150 and Asp166 side chains. Here, we note that the relevance of Cys99 and Gln100 in this intermolecular interaction has been reported [18]. Furthermore, inhibitor **12** has hydrophobic interactions with the Leu21, Val29, Ala42, Asp145, Val152 and Ile165 side chains.
- Inhibitor **4a** [23] (Figure 2E) is a pyridine derivate that has a very different chemical scaffold than the other inhibitors studied and, moreover, it belongs to a family of very active hIKK-2 inhibitors. The Cys99 backbone atoms are involved in two hydrogen-bonding interactions, one of which is between the nitrogen and the hydroxyl group in the 2' position of the benzyl moiety of **4a** (which is an important group for this family of hIKK-2 inhibitors [22–24,37]) and the other is between the carbonyl oxygen and one of two amine group of **4a**. Asp103 also contributes to stabilizing inhibitor **4a** on the other side of the ATP pocket (close to the solvent-accessibility region of the binding site) by forming another hydrogen bond between its side-chain oxygen, which accepts the hydrogen from the second **4a** amine group. The theoretical complex also shows that Val29, Ala42, Val74 and Met96 have hydrophobic interactions with the **4a** benzyl ring.
- Inhibitor **14** [18] (Figure 2F) and inhibitor **12** (Figure 2D) share the same structural scaffold but have different residue interactions with the hIKK-2 homology model. Thus, only one residue in the hinge region (Cys99) forms a hydrogen bond with **14** (instead of the three hydrogen bonds made by **12**; see above). Here, the hydrogen bond is between the Cys99 carbonyl oxygen and the nitrogen from the azaindole six-membered ring of **14**. Additionally, the amine group of **14** forms a hydrogen bond with the Thr23 hydroxyl group (located in the G-loop). The hydrophobic interactions of **14** are with Leu21, Val29, Ala42, Val74, Met96, Tyr98, Val152 and Ile165.

Structure-based pharmacophore description

The validation of the homology model for hIKK-2 showed that our protein-ligand docking strategy found 43 knowledge-based coherent poses for 21 out of the 36 hIKK-2 inhibitors assayed. Then, a visual inspection of the chemical features involved in the intermolecular interactions between hIKK-2 and these 43 poses was used to identify those common to most of them and which,

therefore, could be used to build a common pharmacophore describing the mechanism of the ligand-target interaction. The resulting common structure-based pharmacophore, shown in Figure 2B, contains one hydrogen-bond acceptor, two hydrogen-bond donors and one hydrophobic site that were common to most of the 43 poses. The spatial location of the hydrogen-bond donors and acceptor sites was nearly coincident for all poses, but this was not the case for the hydrophobic site that occupies a large volume in the ATP-binding site. Thus, in the context of the common structure-based pharmacophore, this was translated into a larger radius for the hydrophobic feature (*i.e.*, 3.0 Å) than for the acceptor and donor radii (*i.e.*, 1.5 Å), as a consequence of the huge structural variability found for the inhibitor moieties that bound to the narrow hydrophobic pocket located between the two hIKK-2 lobes (see Figure 2A). Moreover, to increase the power of the pharmacophore to discriminate between hIKK-2 inhibitors and inactive ligands that could adopt poses that accomplish the pharmacophore features, a shell of excluded volumes was added that schematically represents the locations of the residues forming the ATP-binding site in our hIKK-2 homology model (see Figure 2B). These excluded volumes were especially useful during the second step of the VS workflow (see Figure 3), *i.e.*, the identification of molecules having at least one conformer that, irrespective of its initial spatial orientation, can fit into the pharmacophore (because this shell can be used to exclude from the VS those inactive molecules that have at least one conformer that matches the sites in the pharmacophore but is sterically hindered by other residues in the binding site).

Virtual-screening workflow: validation and application to the NP subset of the ZINC database

The capacity of our VS workflow to distinguish between hIKK-2 inhibitors and molecules that do not inhibit hIKK-2 was evaluated by applying it to a set consisting of 62 known hIKK-2 inhibitors (different from the 36 used during the *structure-based* pharmacophore generation; see Table S2) and 10,036 kinase decoys obtained from the Directory of Useful Decoys (DUD; <http://dud.docking.org>) [46]. Figure 3 shows how many actives and decoys *survived* each VS step. Thus, the ADME/Tox, the pharmacophore-docking-pharmacophore and the shape and electrostatic-potential comparison filters produced enrichment factors of 1.4, 6.3 and 4.5, respectively, with a global enrichment factor of 39.3 for the full VS process. Therefore, these results show that our VS protocol is sufficiently selective to discern between those molecules that can inhibit hIKK-2 from those that do not affect its activity. Consequently, it was used to predict that 1,061 out of the 89,425 molecules in the ZINC Natural Products Database were potential hIKK-2 inhibitors (see Figure 3).

Finding new scaffolds for hIKK-2 inhibitors

One of the most important challenges of any VS workflow is its ability to find molecules with the required activity but without trivial similarity (in terms of chemical structure) to known active compounds. Thus, to determine which of the 1,061 potential hIKK-2 inhibitors predicted by our VS workflow could be considered as new lead molecules that may be of use in drug design and development, we merged these 1,061 NPs with the 98 known hIKK-2 inhibitors used either for validation or for pharmacophore-generation purposes, and the resulting set was classified into clusters (data not shown). Notably, three out of the ten clusters obtained consisted exclusively of NPs (*i.e.*, Clusters 1, 2 and 3 contain 120, 88 and 38 NPs, respectively) and, therefore, these 246 molecules are scaffold-hopping candidates for hIKK-2 inhibition (see Table S3). To prove the reliability of our

| Workflow steps | Validation | | | Virtual Screening | | |
|---|------------------------------|---|-----|-------------------|-----|-------|
| | Actives | Decoys | EF | ZINC Database | | |
| Starting Database | 62 | 10036 | - | 89425 | | |
| ADME/Toxicity analysis FaF-Drug2 | 58 | 6816 | 1.4 | 53656 | | |
| Structure-based Pharmacophore Phase v3.1 | Rigid Docking eHiTs v2009 | Structure-based Pharmacophore Phase v3.1 | 26 | 462 | 6.3 | 10892 |
| Similarity & Electrostatic analysis EON v2.0.1 | 14 | 44 | 4.5 | 1061 | | |

Figure 3. The VS workflow used in the present work. The data beside each VS step show the number of molecules that *survived* it. The *ACTIVES* and *DECOYS* columns refer to known hIKK-2 and kinase decoys used during VS validation, respectively. The *ZINC NPs* column refers to data obtained with the ZINC NP subset (http://wiki.compbio.ucsf.edu/wiki/index.php/Natural_products_database). Enrichment factors were calculated during the validation of each step of the VS protocol as the quotient between the fraction of actives in the sample that survived the VS step and the fraction of actives in the sample before the VS step.
doi:10.1371/journal.pone.0016903.g003

predictions, we selected 5 of these 246 molecules (ZINC00058225, ZINC01669260 and ZINC16946275 from Cluster 1; ZINC03683886 from Cluster 2; and ZINC03871389 from Cluster 3; see Figure 4) and tested their effects on the hIKK-2 activity using an *in vitro* assay (see Figure 5). The results of this experiment showed that three out of the five molecules (*i.e.*, ZINC01669260 from Cluster 1, ZINC03683886 from Cluster 2 and ZINC03871389 from Cluster 3) inhibited hIKK-2, with IC_{50} values ranging from 183.8 to 3,325 μ M. Thus, the highest inhibition rate achieved for ZINC03871389 (ethacridine, ethodin or rivanol) was 93% at a 15-mM inhibitor concentration. Similarly, the highest inhibition rates achieved for ZINC01669260 (lupinine) and ZINC03683886 were 91 and 92% at 10- and 15-mM inhibitor concentrations, respectively. Remarkably, a SciFinder search (Chemical Abstracts Service, Columbus, Ohio, USA; <http://www.cas.org/products/sfacad>) of the literature in which the three active molecules are cited revealed that none of them have been described as anti-inflammatory drugs. Therefore, this result supports the ability of the electrostatic-potential and shape comparison performed by EON to smooth differences in chemical structures and translate them into criteria important for their intermolecular interactions with the ligand-binding site (see Figure 6).

Structural analysis of the inhibition of hIKK-2 by ZINC01669260, ZINC03683886 and ZINC03871389

The docking of ZINC01669260, ZINC03683886 and ZINC03871389 in the ATP-binding site of our hIKK-2 model showed that the three molecules share the same form of intermolecular interaction with the enzyme because they match the structure-based common pharmacophore in the same orientation (see Figure 7A) and because they form a hydrogen bond with the carbonyl oxygen of the gk+3 residue (*i.e.*, Cys99; see Figure 7B, 7C and 7D; Cys99 also uses its backbone nitrogen to make a second hydrogen bond with ZINC01669260; Figure 7C). Moreover, ZINC03871389 and ZINC03683886 also interact with the side-chain oxygen from Asp103 (see Figure 7B and 7D), and only

ZINC03871389 is able to form a third hydrogen bond with the main-chain oxygen from the gk+1 residue (*i.e.*, Glu97; see Figure 7B). With respect to the hydrophobic interactions, they were only detected for ZINC03871389 and ZINC03683886 (see Figure 7B and 7D) and involve the acridine moiety of ZINC03871389 (which interacts with Val29, Ala42, Val152 and Ile165) and the methyl group from ZINC03683886 (which interacts with Val29, Ala42, Val74, Val152 and Ile165). Notably, our results show that there is a direct correlation between the number and strength of intermolecular interactions and the inhibitory activity of the three molecules. Thus, ZINC03871389, ZINC03683886 and ZINC01669260 make three/four, two/five and one/zero hydrogen bonds/hydrophobic interactions with key residues on the hIKK-2 ATP binding pocket, and their corresponding IC_{50} values are 183.8, 2,074 and 3,325 μ M, respectively.

Conclusions

The challenge of any VS protocol consists of using *in silico* tools to predict which of the molecules in a database have the required activity against a specific target. Thus, the results of the present study demonstrate that our VS protocol successfully identified hIKK-2 inhibitors with no chemical-structure similarities to known actives and, therefore, it is suitable for scaffold hopping on this target. Moreover, this is the first time that anti-inflammatory activity has been described for ZINC03683886, lupinine (*i.e.*, ZINC01669260) and ethodin (*i.e.*, ZINC03871389; a molecule that is known to cause uterine contractions [59,60] and that has been also described as inhibitor of the human carboxylesterase 1 [61]). In this sense, the results of the present study outperform those of a recent *in silico* survey on the ChemDiv database (<http://us.chemdiv.com>) that found hIKK-2 inhibitory activity for only two out of 29 hits selected for *in vitro* testing the VS workflow success rate [42]. Moreover, one of these two molecules shares the same core scaffold with BMS-345541 (a well-known hIKK-2 inhibitor [62]) and, therefore, only the other molecule can

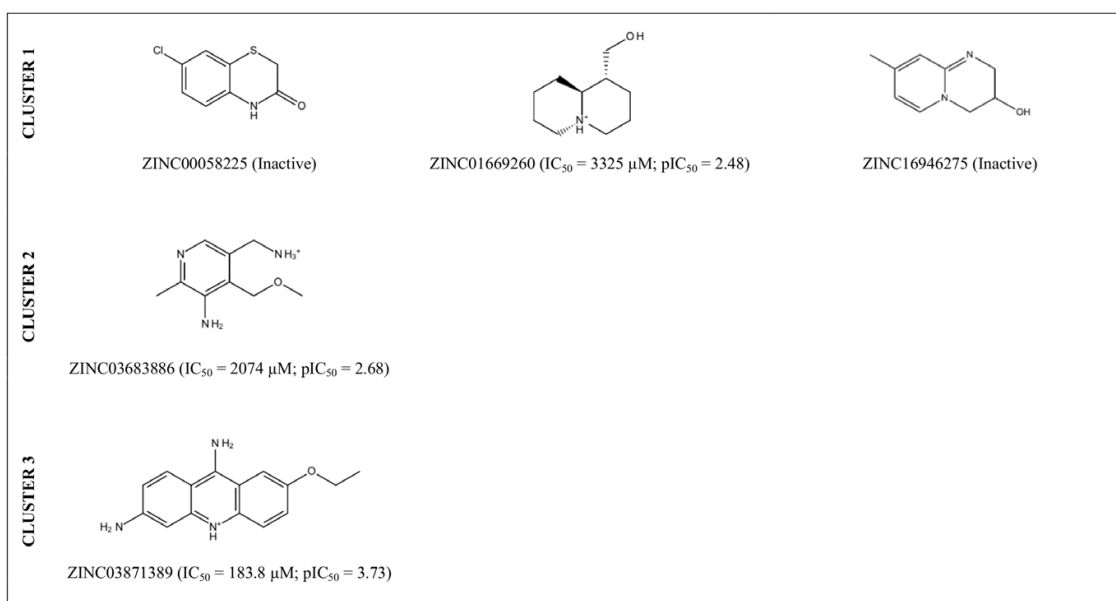


Figure 4. 2D structures of the VS hits for which hIKK-2-inhibition activity was tested in an *in vitro* assay. The IC_{50} and pIC_{50} values are also shown for those molecules that inhibit hIKK-2 and were calculated with GraphPad Prism v4.0. doi:10.1371/journal.pone.0016903.g004

be considered a lead compound for the development of new hIKK-2 inhibitors.

Furthermore, although the pIC_{50} of the three hit molecules that showed *in vitro* activity is significantly lower than that for most known hIKK-2 inhibitors used in the present study (see Tables S1 and S2), it is important that these three molecules (**a**) can be used as lead compounds for developing more potent inhibitors using structural-activity relationship studies and (**b**) were selected based on their commercial availability, cost and purity and with the primary goal of testing the performance of our VS protocol. Therefore, it is possible that there are other molecules in the remaining 241 included in Clusters 1, 2 and 3

(see Table S3) that show higher activity towards hIKK-2 than ZINC03871389, ZINC03683886 and ZINC01669260 and could be better lead compounds for drug development than these three molecules. In this sense, our results indicate that the activity correlates well with the *ET_combo* score provided by EON (see Figure 4 and Table S3) and, therefore, this finding suggests that molecules with *ET_combo* values higher than 1.2 (e.g., ZINC12410246, ZINC00485744 and ZINC00071662 from Clusters 1, 2 and 3, respectively) are optimal starting points for the rational drug design of potent and selective hIKK-2 inhibitors with new chemical scaffolds.

Supporting information

Details on hIKK-2 synthetic inhibitors used in the present work are described in Table S1 (for pharmacophore-generation purposes) and S2 (for validation purposes). Table S3 reports the ZINC codes for the 246 hit molecules that are predicted to inhibit hIKK-2 and that are scaffold-hopping candidates for hIKK-2 inhibition, because the Tanimoto similarities between their MOLPRINT 2D fingerprints and those from the hIKK-2 inhibitors in Table S1 and S2 are significantly low.

Methods

Homology-model building of the 3D-structure for the kinase domain of hIKK-2

The sequence of hIKK-2 was obtained from the UniProt database (<http://www.uniprot.org>, accession number O14920). The protein templates used to build its homology model were PDB structures that simultaneously satisfied the following criteria: (**a**) highest sequence similarity with hIKK-2, (**b**) X-ray-derived structures with acceptable quality (*i.e.*, resolution $< 2.50 \text{ \AA}$ and R-value $< 0.210 \text{ \AA}$), (**c**) complexation with inhibitors that bind to the ATP-binding site (because we were interested in the activated

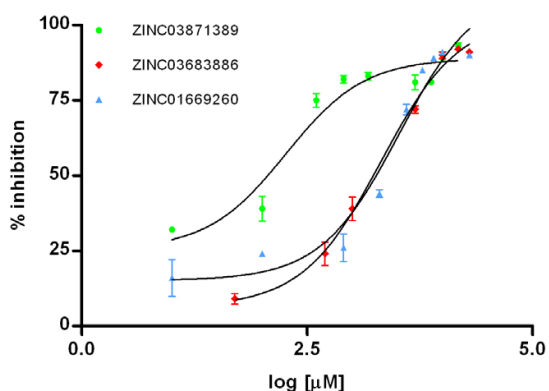


Figure 5. Dose-response results for the *in vitro* inhibition of hIKK-2 by ZINC01669260, ZINC03683886 and ZINC03871389 (with $n = 3$). The figure was drawn with GraphPad Prism v4.0. doi:10.1371/journal.pone.0016903.g005

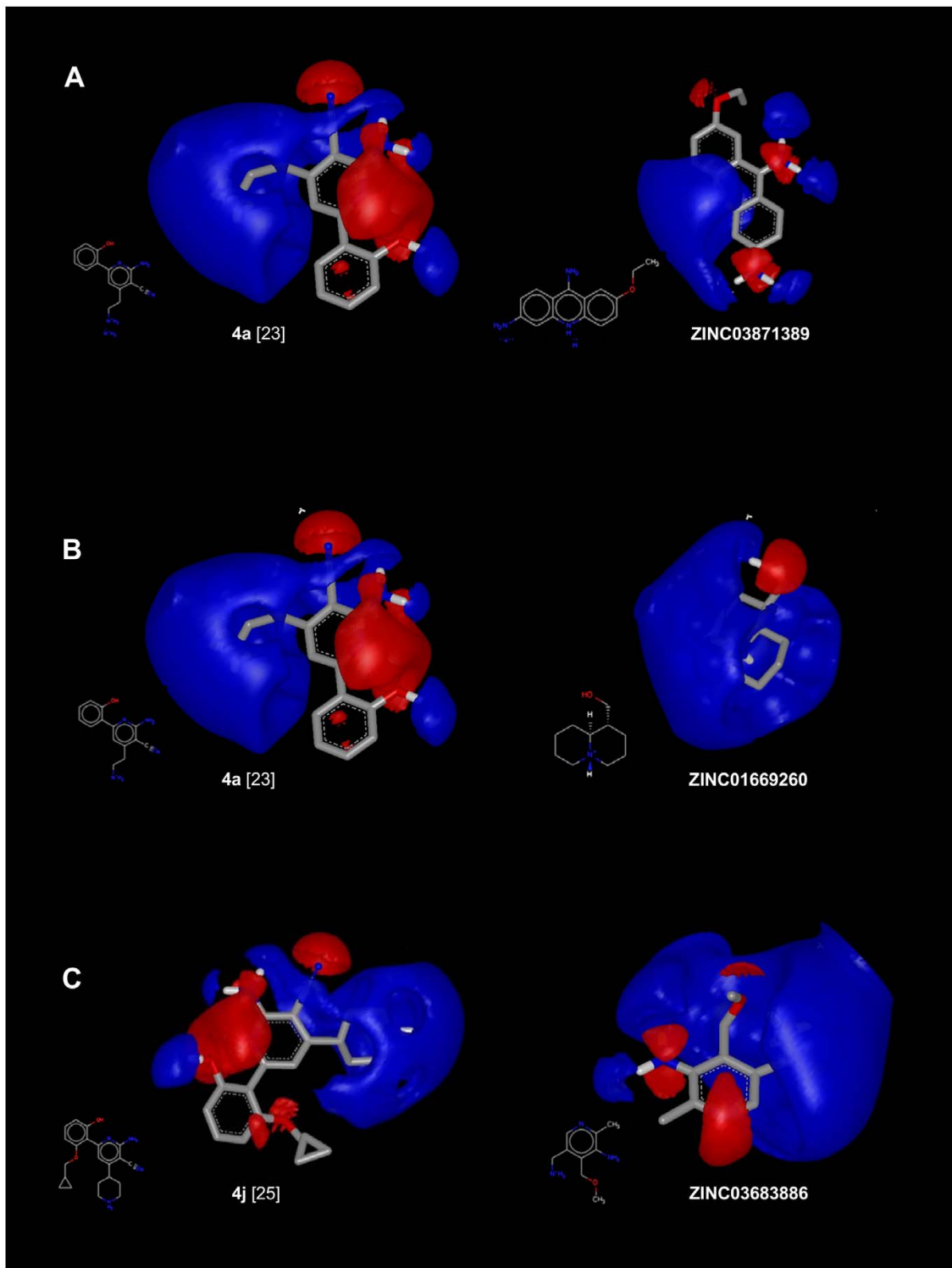


Figure 6. EON results for the best comparison between known hIKK-2 inhibitors and ZINC01669260, ZINC03683886 and ZINC03871389. The best electrostatic-potential and shape comparisons between the docked poses from ZINC01669260, ZINC03683886 and ZINC03871389 and those from the known hIKK-2 inhibitors used by EON as the queries of the last VS filter. For clarity, only the electrostatic potential fields are shown. This Figure was drawn with VIDA v4.03 (OpenEye Scientific Software, Inc., Santa Fe, New Mexico, USA; <http://www.eyesopen.com>). doi:10.1371/journal.pone.0016903.g006

form of our templates), **(d)** closest proximity to hIKK-2 [63] in the human kinome tree (although IKK-2 is a serine/threonine kinase without a defined family in this tree), and **(e)** structures with no mutations in the active site of the kinase domain. The two kinases that fulfilled these criteria and were thus chosen as templates belong to the CAMK kinase family [63], have a 32% sequence identity with the kinase domain of hIKK-2, have a 50% identity with the residues of the ATP-binding site of hIKK-2 (where these residues correspond to those in the G-loop, hinge region, HRD motif and catalytic loop); they are the serine/threonine-protein

kinase CHK2 (PDB code 2CN5 [56]) and the death-associated protein kinase 3 or ZIP kinase (PDB code 2J90 [57]). The superposition of both template structures revealed some differences in their loop structures: **(a)** the G-loop is more open in 2J90 than in 2CN5 and **(b)** the A-loop coordinates were not present for 2J90. Armed with this information and following a knowledge-based approach, we used MODELLER v9.5 [64] to align the sequences and construct several models by changing their alignments with the kinase domain of hIKK-2. The resulting models were subsequently refined with MODELLER in a two-step

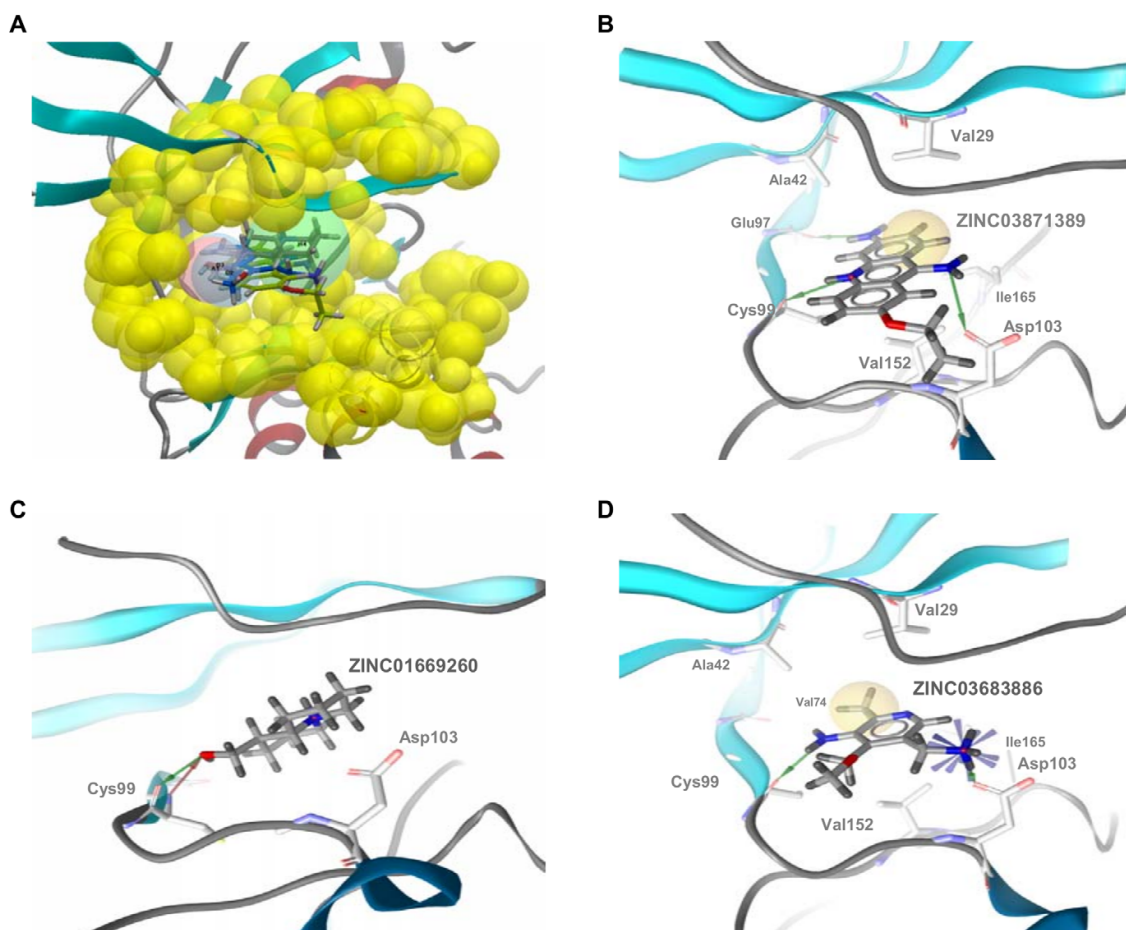


Figure 7. Glide XP results for ZINC01669260, ZINC03683886 and ZINC03871389. Glide XP results for ZINC01669260, ZINC03683886 and ZINC03871389 in the context of the structure-based common pharmacophore (panel A) and the hIKK-2 ATP-binding site (panels B, C and D). Thus, Figure 7A shows that Glide XP was able to find at least one pose for ZINC01669260, ZINC03683886 and ZINC03871389 that matched the pharmacophore. The intermolecular interactions of these poses with the hIKK-2 homology model are shown in Figure 7B (ZINC03871389), 7C (ZINC01669260) and 7D (ZINC03683886). Hydrophobic contacts and hydrogen-bond donors/acceptors are shown as yellow spheres and red/green arrows, respectively. doi:10.1371/journal.pone.0016903.g007

process (*i.e.*, loops were refined first without the use of the templates following an *ab initio* methodology [65], and α -helices were the refined when loop processing was finished using the templates) and using default running conditions. Next, the quality of the resulting models was evaluated by **(a)** assessing the packing quality with the **Atomic Non-Local Environment Assessment** (ANOLEA) webserver (<http://protein.bio.puc.cl/cardex/servers/anolea>) [66], **(b)** analyzing the outliers in the Ramachandran plot; **(c)** examining the ATP-binding site structure (because it must be optimized to built a receptor-based pharmacophore), and **(d)** performing a protein-ligand rigid-docking study with the eHiTS v2009 [67] software (SimBioSys Inc., Toronto, Canada; <http://www.simbiosys.ca/ehits>) using various known hIKK-2 inhibitors [47,48] and visually analyzing the knowledge-based coherence of the results. In this sense, the capability of eHiTS to predict a knowledge-based coherent pose was assessed by redocking the two complexes used as templates to obtain the homology models. Thus, eHiTS predicted poses for the ADP and pyridone 6 with root-mean-square deviations relative to their experimental poses in 2CN5 and 2J90 of 1.23 and 0.38 Å, respectively.

Structure-based pharmacophore generation

The best homology model obtained for the hIKK-2 kinase domain was subsequently used as the receptor in a protein-ligand docking with 36 known selective ATP-competitive inhibitors for hIKK-2 from different chemical families (see Table S1) [10,11,13,15,16,18,20–31]. The docking was done with eHiTS v2009 [67] by considering the receptor to be a rigid body and the ligands as flexible (*i.e.*, free rotation was allowed around the single bonds of the ligand). Default docking conditions were selected except for the size of the sides of the cubic box encompassing the ATP-binding pocket, which was increased from 10 Å to 15 Å. Thereafter, all subsequent protein-ligand dockings in this work, either for validation or VS purposes, were performed with eHiTS v2009 using the same running conditions, hIKK-2 area and homology-model orientation. Glide XP v5.6 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>), running with default conditions, was used to predict the poses of three NPs that our VS workflow predicted as hIKK-2 inhibitors and for which activity was confirmed by the *in vitro* assay.

The protein-ligand interactions of the poses obtained by the docking process (a maximum of 32 poses per ligand) were analyzed with LigandScout v2.03 (Inteligand GmbH, Vienna, Austria, <http://www.inteligand.com/ligandscout>) [68]. Next, the poses that satisfied the generic binding features of ATP-competitive kinase inhibitors [45] [*i.e.*, the formation of hydrogen bonds with the main chain of residues in the hinge region (segment 96–99 in hIKK-2 sequence) and hydrophobic interactions with the hydrophobic pocket of hIKK-2 (*e.g.* Val29, Lys44, Val152, and Ile65)] were selected as knowledge-based coherent (regardless of their eHiTS scores). As a result of this analysis, 43 poses of the 21 hIKK-2 inhibitors were identified as knowledge-based coherent, and their corresponding sites (*i.e.*, functional groups used by the poses in their intermolecular interactions with the hIKK-2 kinase domain) were used to derive a common structure-based pharmacophore. This pharmacophore contains one hydrogen-bond acceptor, two hydrogen-bond donors and one hydrophobic site common to most of the 43 poses. Moreover, the pharmacophore was completed with a shell of excluded volumes that schematically represents the location of the residues forming the ATP-binding site in our hIKK-2 homology model. This shell was built by applying the *Receptor-Based Excluded Volumes* graphic front-end in Phase v3.1 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) [69] to our hIKK-2 kinase domain model and by setting the *Sphere filters* parameter

values to **(a)** ignore receptor atoms with surfaces within **0.25 Å** of the ligand surface and **(b)** limit excluded-volume shell thickness to **10 Å**. The remaining parameters used in the front-end were the defaults.

Ligand setup

All the ligands used in this work (either for validation methodology or for VS purposes) were processed in the same way. The 3D ligand structures were either downloaded from public databases [*i.e.*, DUD [46], the BindingDB (<http://www.bindingdb.org>) [47] and the ZINC Natural Products Database (http://wiki.compbio.ucsf.edu/wiki/index.php/Natural_products_database) [48]] or constructed with ChemDraw Ultra v11.0 (CambridgeSoft Corporation, Cambridge, MA, USA; <http://www.cambridgesoft.com>) [70]. These 3D structures were then incorporated into LigPrep v2.3 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) and improved by cleaning. The cleaning process was carried out with the following parameters: **(a)** the force field used was OPLS 2005; **(b)** all possible ionization states at pH 7.0±2.0 were generated with Ionizer; **(c)** the *desalt* option was activated; **(d)** tautomers were generated for all ionization states at pH 7.0±2.0; **(e)** chiralities, when present, were determined from the 3D structure; and **(f)** one low-energy ring conformation per ligand was generated. Conformations and sites for the resulting ligand structures were determined during the generation of the corresponding Phase [69] databases with the *Generate Phase Database* graphic front-end. The parameter values used during this conformer generation were by default with the exception of the maximum number of conformers per structure, which was increased from 100 (the default value) to 200. The conformer sites were generated with definitions made by adding the ability to consider aromatic rings as hydrophobic groups to the default built-in Phase definitions.

Virtual-screening workflow

The NP subset of the ZINC database (<http://zinc.docking.org/vendor0/npd>) [48] contains the 3D structure of 89,425 commercially available NPs and NP derivatives and was used as the source of molecules to which our VS schema was applied in the search for new hIKK-2 inhibitors. Initially, these 89,425 molecules were submitted to an ADME/Tox filter with the FAF-Drugs2 tool [71], which aimed to discard molecules that either had poor ADME properties or were potentially toxic. Thus, the drug-like properties of a compound were evaluated using the Lipinski rule [72], and only one violation of this rule was allowed. This rule is based on a set of property values (*i.e.*, the number of hydrogen-bond donors and acceptors, the molecular weight and the logP) that were derived from drugs with good ADME characteristics [72]. Therefore, molecules that pass the Lipinski rule are expected to be active in humans after oral admission. Moreover, molecules containing toxic groups were filtered using the 204 substructures for “warhead” chelators, frequent hitters, promiscuous inhibitors and other undesirable functional groups available in FAF-Drugs2 tool [71].

Molecules with appropriate ADME/Tox properties were then set up with LigPrep and used to generate a Phase database. Next, this database was filtered with Phase through the structure-based common pharmacophore using the following running conditions: **(a)** search in the conformers database, **(b)** do not score *in place* the conformers into the structure-based common pharmacophore (*i.e.*, allow reorientation of the conformers to determine if they match the pharmacophore or not), **(c)** match at least three out of the four sites of the structure-based common pharmacophore, **(d)** do not consider any site as mandatory, **(e)** do not prefer partial matches involving more sites and **(f)** use the excluded volumes from the

structure-based common pharmacophore. The rest of options and parameter values used during that search were the default values.

Those ligands with at least one hit in this Phase search were then docked to our hIKK-2 homology model. Then, the resulting ligand poses were again filtered with Phase through the structure-based common pharmacophore using the same filtering conditions as in the first Phase run, with the exception that no reorientation of the poses was allowed during the search (*i.e.*, the *score in place* option was used) to find docking poses that were compatible with the pharmacophore.

Finally, the poses that were hits in this second pharmacophore screening were submitted to a shape and electrostatic-potential comparison with 43 poses from 21 known hIKK-2 inhibitors (see Table S1) that (**a**) were also obtained by docking with our hIKK-2 homology model and (**b**) also matched the structure-based common pharmacophore without reorientation. This comparison was done with EON v2.0.1 (OpenEye Scientific Software, Inc., Santa Fe, New Mexico, USA; <http://www.eyesopen.com>) using the *ET_combo* score as the similarity criterion. The *ET_combo* is the sum of two calculations: (**a**) the *Shape Tanimoto*, which is a quantitative measure of three-dimensional overlap [where 1 corresponds to a perfect overlap (*i.e.*, the same shape)] [73] and (**b**) the *Poisson-Boltzman Electrostatic Tanimoto*, which compares the electrostatic potential of two small molecules (where 1 corresponds to identical potentials and negative values correspond to the overlap of positive and negative charges) [74]. Only those NPs with *ET_combo* values equal to or greater than 0.850 were considered as candidates for experimental tests of their activity on hIKK-2.

Virtual-screening workflow validation

The ability of the VS workflow to identify hIKK-2 inhibitors in a database of molecules was tested by applying the same filtering method described for the NPs (*i.e.*, the same methodology and parameter values) to: (**a**) a group of 62 known inhibitors different from the 36 used for the *structure-based* pharmacophore generation (see Table S2) and (**b**) an external group of 10,139 kinase decoys for Cdk2 and p38 (CMGC family at the human kinome tree [63]) that were obtained from DUD [46].

Hit selection for further experimental assays on hIKK-2 activity

A structural-similarity analysis of the NPs with *ET_combo* values equal to or greater than 0.850 and all known hIKK-2 inhibitors used in this work (either for validation or for pharmacophore-generation purposes) was performed by executing a Schrödinger script that clusters molecules based on Tanimoto similarities between MOLPRINT 2D fingerprints. We obtained ten clusters and, notably, three of them consisted only of NPs. Thus, five compounds from these three clusters were selected [based on their commercial availability, cost and purity ($\geq 95\%$)] for experimental assay of their effects on hIKK-2 activity. These compounds were ZINC03683886, ZINC00242004 and ZINC03871389, bought from InterBioScreen, Ltd (<http://www.ibscreen.com>), and ZINC00058225 and ZINC01669260, bought from TimTec, Inc. (<http://www.timtec.net>).

In vitro assay of the effect of selected compounds on the hIKK-2 activity

The effect of the selected compounds on hIKK-2 kinase activity was determined using the Calbiochem® K-LISA™ IKK β Inhibitor Screening Kit, designed for rapid *in vitro* screening of IKK-2 inhibitors. Briefly, the kit consists of an ELISA-based activity assay

that uses a 50-amino-acid GST-IkK α fusion-polypeptide substrate that includes the Ser³² and Ser³⁶ IKK-2 phosphorylation sites. The GST-IkK α substrate and IKK-2 are incubated in the presence of IKK-2 inhibitors in the wells of a glutathione-coated 96-well plate, which allows for substrate phosphorylation and capture in a single step. The phosphorylated GST-IkK α substrate is detected using an anti-phospho-IkK α (Ser³²/Ser³⁶) antibody followed by an HRP-conjugate and color development with the TMB substrate. The ELISA stop solution is then used to stop the color development, and the absorbance is read at 450 nm (with a reference wavelength of 540–600 nm). The absorbance is directly related to the IKK-2 activity level. Moreover, following the kit instructions, one positive control and three negative controls were included by varying the absence or presence of the commercial hIKK-2 inhibitor IV (Calbiochem catalogue number: 401481), the ATP/MgCl₂ mix and the hIKK-2 enzyme. The results of these controls confirmed the absence of nonspecific or artificial inhibition. Three repetitions were made for all the activity assays.

IC₅₀ calculation

The potential of the molecules to inhibit hIKK-2 activity was quantified by calculating their IC₅₀ values. This parameter was determined using GraphPad Prism v4.0 for Windows (GraphPad Software, San Diego CA, USA; <http://www.graphpad.com>) by fitting the experimental data from the *in vitro* assay to a nonlinear regression function using a four-parameter logistic equation.

Supporting Information

Table S1 hIKK-2 inhibitors from different chemical families docked with the hIKK-2 homology model. These 36 hIKK-2 inhibitors were docked with our hIKK-2 homology model without imposing constraints that forced poses to make specific intermolecular interactions with the target. Next, the resulting hIKK-2 complexes were analyzed with the help of LigandScout to determine which complexes exhibited target-inhibitor intermolecular interactions equivalent to those described in prior studies. This knowledge-based analysis enabled us to identify at least one knowledge-based coherent pose (43 in total) for 21 out of the 36 hIKK-2 inhibitors assayed (regardless of their scoring by eHiTS; see their values in the *Knowledge-based coherent docking poses* column). By analyzing the chemical features used by each pose in its intermolecular interaction with hIKK-2, a common pharmacophore was derived that describes the mechanism of the ligand-target interaction. The *Cluster* column shows the cluster in which each molecule was classified after running a Schrödinger script that clusters molecules based on Tanimoto similarities between MOLPRINT 2D fingerprints (using the Knime v.2.0.3 module in the Schrödinger software package). The molecules distributed in these clusters are the natural products obtained as hits in our virtual-screening protocol and all known hIKK-2 inhibitors used in the present work [either for validation (see Table S2) or for pharmacophore-generation purposes]. The pIC₅₀ values were obtained from the literature. (PDF)

Table S2 hIKK-2 inhibitors used during the validation of the virtual-screening workflow. These 62 hIKK-2 inhibitors (different from the 36 used during the structure-based pharmacophore generation; see Table S1) were used to test the ability of the virtual-screening workflow to identify hIKK-2 inhibitors in a database of molecules. The *Cluster* column shows the cluster into which each molecule was classified after running a Schrödinger script that clusters molecules based on Tanimoto similarities

between MOLPRINT 2D fingerprints (using the Knime v.2.0.3 module in the Schrödinger software package). The molecules distributed in these clusters are the natural products obtained as hits in our virtual-screening protocol and all known hIKK-2 inhibitors used in the present work (either for validation or for pharmacophore-generation purposes). The pIC₅₀ values were obtained from the literature.
(PDF)

Table S3 Scaffold-hopping candidates for hIKK-2 inhibition predicted by our study. The ZINC codes for the 246 hit molecules predicted to inhibit hIKK-2 and belonging to clusters consisting exclusively of natural products. For each hit molecule, the best results of the shape and electrostatic-potential comparisons with 43 poses from 21 known hIKK-2 inhibitors (see Table S1) is shown. Thus, the Tanimoto values for the comparison between the electrostatic potentials of the molecules (using an outer dielectric of 80) are shown in the **PB** columns, whereas the values for the comparison between shapes are shown in the **Shape** columns. The sum of the PB and Shape values is reported in the **Combo** columns. Hits from each cluster are sorted according their decreasing combo value. All of these hit molecules are scaffold-hopping candidates for hIKK-2 inhibition because the Tanimoto similarities between their MOLPRINT 2D fingerprints and those from the hIKK-2 inhibitors in Tables S1 and S2 are quite low. ZINC00058225, ZINC01669260 and ZINC16946275 from Cluster 1, ZINC03683886 from Cluster 2,

and ZINC03871389 from Cluster 3 were selected to experimentally test the success rate of our predictions using an *in vitro* assay (in bold in Table S3). The results of this experiment showed that three out of the five molecules (*i.e.*, ZINC01669260 from Cluster 1, ZINC03683886 from Cluster 2 and ZINC03871389 from Cluster 3) inhibited hIKK-2, with IC₅₀ values ranging from 183.8 to 3,325 μM.
(PDF)

Acknowledgments

This manuscript has been edited for English-language fluency by American Journal Experts. We thank Dr. Daniel D. Robinson (Schrödinger, LLC) for his useful advice during the homology modeling of hIKK-2 and technician Vanessa Grifoll of our group for her support and help in the experimental part of the *in vitro* assay. The authors wish to thank the Servei de Disseny de Fàrmacs (Drug Design Service) of the Catalonia Supercomputer Center (CESCA) for providing access to Schrödinger software and the Protein Modeling Facility of the Lausanne University for help during the homology modeling of hIKK-2.

Author Contributions

Conceived and designed the experiments: GP SG-V MM OM VZ. Performed the experiments: ES LG JI. Analyzed the data: ES LG JI GP SG-V MM OM VZ M-JS MP. Contributed reagents/materials/analysis tools: AG. Wrote the manuscript: ES GP SG-V VZ OM.

References

1. Rollinger JM, Stuppner H, Langer T (2008) Virtual screening for the discovery of bioactive natural products. In: Petersen F, Amstutz R, eds. Natural Compounds as Drugs. Basel: Birkhäuser Basel. pp 211–249.
2. Cragg GM, Newman DJ (2009) Nature: a vital source of leads for anticancer drug development. *Phytochem Rev* 8: 313–331.
3. Harvey AL (2008) Natural products in drug discovery. *Drug Discov Today* 13: 894–901.
4. Rollinger JM, Langer T, Stuppner H (2006) Strategies for efficient lead structure discovery from natural products. *Curr Med Chem* 13: 1491–1507.
5. Kumar K, Waldmann H (2009) Synthesis of Natural Product Inspired Compound Collections. *Angew Chem Int Edit* 48: 3224–3242.
6. Williams PG (2009) Panning for chemical gold: marine bacteria as a source of new therapeutics. *Trends Biotechnol* 27: 45–52.
7. Zhou XW, Gong ZH, Su Y, Lin J, Tang KX (2009) Cordyceps fungi: natural products, pharmacological functions and developmental products. *J Pharm Pharmacol* 61: 279–291.
8. Rollinger JM, Langer T, Stuppner H (2006) Integrated *in silico* tools for exploiting the natural products' bioactivity. *Planta Med* 72: 671–678.
9. Castro AC, Dang LC, Soucy F, Grenier L, Mazdiyasn H, et al. (2003) Novel IKK inhibitors: β-carbolines. *Bioorg Med Chem Lett* 13: 2419–2422.
10. Clare M, Fletcher T, Hamper BC, Hanson G, Heier RF, et al. (2005) Substituted pyrazole urea compounds for the treatment of inflammation. *WO/2005/037797*.
11. Karin M, Yamamoto Y, Wang QM (2004) The IKK NF-κB system: a treasure trove for drug development. *Nat Rev Drug Discov* 3: 17–26.
12. Bhagwat SS (2009) Kinase inhibitors for the treatment of inflammatory and autoimmune disorders. *Purinergic Signal* 5: 107–115.
13. Nagarajan S, Doddareddy M, Choo H, Cho YS, Oh KS, et al. (2009) IKKβ inhibitors identification part I: homology model assisted structure based virtual screening. *Bioorg Med Chem Lett* 17: 2759–2766.
14. Nagarajan S, Choo H, Cho YS, Oh KS, Lee BH, et al. (2010) IKKβ inhibitors identification part II: ligand and structure-based virtual screening. *Bioorg Med Chem Lett* 18: 3951–3960.
15. Christopher JA, Avitabile BG, Bamborough P, Champigny AC, Cutler GJ, et al. (2007) The discovery of 2-amino-3,5-diarylbenzamide inhibitors of IKK-α and IKK-β kinases. *Bioorg Med Chem Lett* 17: 3972–3977.
16. Bingham AH, Davenport RJ, Gowers L, Knight RL, Lowe C, et al. (2004) A novel series of potent and selective IKK2 inhibitors. *Bioorg Med Chem Lett* 14: 409–412.
17. Bingham AH, Davenport RJ, Fosbeary R, Gowers L, Knight RL, et al. (2008) Synthesis and structure-activity relationship of aminopyrimidine IKK2 inhibitors. *Bioorg Med Chem Lett* 18: 3622–3627.
18. Liddle J, Bamborough P, Barker MD, Campos S, Cousins RP, et al. (2009) 4-Phenyl-7-azaindoles as potent and selective IKK2 inhibitors. *Bioorg Med Chem Lett* 19: 2504–2508.
19. Christopher JA, Bamborough P, Alder C, Campbell A, Cutler GJ, et al. (2009) Discovery of 6-aryl-7-alkoxyisoquinoline inhibitors of IκB kinase-β (IKK-β). *J Med Chem* 52: 3098–3102.
20. Beaulieu F, Ouellet C, Ruediger EH, Belema M, Qiu YP, et al. (2007) Synthesis and biological evaluation of 4-amino derivatives of benzimidazoquinoline, benzimidazoquinoline, and benzopyrazoloquinazoline as potent IKK inhibitors. *Bioorg Med Chem Lett* 17: 1233–1237.
21. Belema M, Bunker A, Nguyen V, Beaulieu F, Ouellet C, et al. (2007) Synthesis and structure-activity relationship of imidazo[1,2-a]thieno[3,2-c]pyrazines as IKK-β inhibitors. *Bioorg Med Chem Lett* 17: 4284–4289.
22. Murata T, Shimada M, Sakakibara S, Yoshino T, Kadono H, et al. (2003) Discovery of novel and selective IKK-β serine-threonine protein kinase inhibitors. Part 1. *Bioorg Med Chem Lett* 13: 913–918.
23. Murata T, Shimada M, Kadono H, Sakakibara S, Yoshino T, et al. (2004) Synthesis and structure-activity relationships of novel IKK-β inhibitors. Part 2: Improvement of *in vitro* activity. *Bioorg Med Chem Lett* 14: 4013–4017.
24. Murata T, Shimada M, Sakakibara S, Yoshino T, Masuda T, et al. (2004) Synthesis and structure-activity relationships of novel IKK-β inhibitors. Part 3: Orally active anti-inflammatory agents. *Bioorg Med Chem Lett* 14: 4019–4022.
25. Ziegelbauer K, Gantner F, Lukacs N, Berlin A, Fuchikami K, et al. (2005) A selective novel low-molecular-weight inhibitor of IκB kinase-β (IKK-β) prevents pulmonary inflammation and shows broad anti-inflammatory activity. *Br J Pharmacol* 145: 178–192.
26. Waelchli R, Bollbuck B, Bruns C, Buhl T, Eder J, et al. (2006) Design and preparation of 2-benzamido-pyrimidines as inhibitors of IKK. *Bioorg Med Chem Lett* 16: 108–112.
27. Baxter A, Brough S, Cooper A, Floetmann E, Foster S, et al. (2004) Hit-to-lead studies: the discovery of potent, orally active, thiophenecarboxamide IKK-2 inhibitors. *Bioorg Med Chem Lett* 14: 2817–2822.
28. Morwick T, Berry A, Brickwood J, Cardozo M, Catron K, et al. (2006) Evolution of the thienopyridine class of inhibitors of IκB kinase-β. Part I: hit-to-lead strategies. *J Med Chem* 49: 2898–2908.
29. Bonafoux D, Bonar S, Christine L, Clare M, Donnelly A, et al. (2005) Inhibition of IKK-2 by 2-[(aminocarbonyl)amino]-5-acetylenyl-3-thiophenecarboxamides. *Bioorg Med Chem Lett* 15: 2870–2875.
30. Sommers CD, Thompson JM, Guzova JA, Bonar SL, Rader RK, et al. (2009) Novel tight-binding inhibitory factor-κB kinase (IKK-2) inhibitors demonstrate target-specific anti-inflammatory activities in cellular assays and following oral and local delivery in an *in vivo* model of airway inflammation. *J Pharmacol Exp Ther* 330: 377–388.
31. Sugiyama H, Yoshida M, Mori K, Kawamoto T, Sogabe S, et al. (2007) Synthesis and structure activity relationship studies of benzothieno[3,2-b]furan derivatives as a novel class of IKKβ inhibitors. *Chem Pharm Bull (Tokyo)* 55: 613–624.

New hIKK-2 Inhibitors of Natural Origin

32. Bhattarai BR, Ko JH, Shrestha S, Kafle B, Cho H, et al. (2010) Inhibition of IKK- β : A new development in the mechanism of the anti-obesity effects of PTP1B inhibitors SA18 and SA32. *Bioorg Med Chem Lett* 20: 1075–1077.
33. Crombie AL, Sum FW, Powell DW, Hopper DW, Torres N, et al. (2010) Synthesis and biological evaluation of tricyclic anilinopyrimidines as IKK β inhibitors. *Bioorg Med Chem Lett* 20: 3821–3825.
34. De Silva D, Mitchell MD, Keelan JA (2010) Inhibition of choriodecidual cytokine production and inflammatory gene expression by selective I κ B kinase (IKK) inhibitors. *Brit J Pharmacol* 160: 1808–1822.
35. Shimizu H, Tanaka S, Toki T, Yasumatsu I, Akimoto T, et al. (2010) Discovery of imidazo 1,2-b pyridazine derivatives as IKK β inhibitors. Part 1: Hit-to-lead study and structure-activity relationship. *Bioorg Med Chem Lett* 20: 5113–5118.
36. Mbalaviele G, Sommers CD, Bonar SL, Mathialagan S, Schindler JF, et al. (2009) A Novel, Highly Selective, Tight Binding I κ B Kinase-2 (IKK-2) Inhibitor: A Tool to Correlate IKK-2 Activity to the Fate and Functions of the Components of the Nuclear Factor- κ B Pathway in Arthritis-Relevant Cells and Animal Models. *J Pharmacol Exp Ther* 329: 14–25.
37. Sala E, Guasch L, Vaqué M, Mateo-Sanz JM, Blay M, et al. (2009) 3D-QSAR Study of Pyridine Derivates as IKK-2 Inhibitors. *QSAR Comb Sci* 28: 678–695.
38. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, et al. (2002) The Protein Data Bank. *Acta Crystallogr D* 58: 899–907.
39. Rushe M, Silvian L, Bixler S, Chen LL, Cheung A, et al. (2008) Structure of a NEMO/IKK-associating domain reveals architecture of the interaction site. *Structure* 16: 798–808.
40. Nagarajan S, Ahmed A, Choo H, Cho YS, Oh KS, et al. (2010) 3D QSAR pharmacophore model based on diverse IKK β inhibitors. *J Mol Model*. doi:10.1007/s00894-010-0714-8.
41. Noha SM, Atanasov AG, Schuster D, Markt P, Fakhrudin N, et al. (2010) Discovery of a novel IKK- β inhibitor by ligand-based virtual screening techniques. *Bioorg Med Chem Lett*. doi:10.1016/j.bmcl.2010.10.051.
42. Nagarajan S, Choo H, Cho YS, Shin KJ, Oh KS, et al. (2010) IKK β inhibitor identification: a multi-filter driven novel scaffold. *BMC Bioinformatics*. doi:10.1186/1471-2105-11-S7-S15.
43. Avila CM, Romeiro NC, Sant'Anna CMR, Barreiro EJ, Fraga CAM (2009) Structural insights into IKK β inhibition by natural products staurosporine and quercetin. *Bioorg Med Chem Lett* 19: 6907–6910.
44. Lauria AIM, Fazzari M, Tutone M, Di Blasi F, Mingoaia F, et al. (2010) IKK- β inhibitors: an analysis of drug-receptor interaction by using Molecular Docking and Pharmacophore 3D-QSAR approaches. *J Mol Graph Model* 29: 72–81.
45. Ghose AK, Herberich T, Pippin DA, Salvino JM, Mallamo JP (2008) Knowledge based prediction of ligand binding modes and rational inhibitor design for kinase drug discovery. *J Med Chem* 51: 5149–5171.
46. Huang N, Shoichet BK, Irwin JJ (2006) Benchmarking sets for molecular docking. *J Med Chem* 49: 6789–6801.
47. Liu TQ, Lin YM, Wen X, Jorissen RN, Gilson MK (2007) BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res* 35: D198–D201.
48. Irwin JJ, Shoichet BK (2005) ZINC - A free database of commercially available compounds for virtual screening. *J Chem Inf Model* 45: 177–182.
49. Schmid JA, Birbach A (2008) I κ B kinase β (IKK β /IKK2/IKKBK β)—a key molecule in signaling to the transcription factor NF- κ B. *Cytokine Growth Factor Rev* 19: 157–165.
50. Kornev A, Haste N, Taylor S, Eyck L (2006) Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A* 103: 17783–17788.
51. Rabiller M, Getik M, Klüter S, Richters A, Tückmantel S, et al. (2010) Proteus in the world of proteins: conformational changes in protein kinases. *Arch Pharm (Weinheim)* 343: 193–206.
52. Delhase M, Hayakawa M, Chen Y, Karin M (1999) Positive and negative regulation of I κ B kinase activity through IKK β subunit phosphorylation. *Science* 284: 309–313.
53. Sowadski JM, Epstein LF, Lankiewicz L, Karlsson R (1999) Conformational diversity of catalytic cores of protein kinases. *Pharmacol Ther* 82: 157–164.
54. Engh RA, Bossemeyer D (2002) Structural aspects of protein kinase control-role of conformational flexibility. *Pharmacol Ther* 93: 99–111.
55. Eglén R, Reisine T (2009) The current status of drug discovery against the human kinome. *Assay Drug Dev Technol* 7: 22–43.
56. Oliver AW, Paul A, Boxall KJ, Barrie SE, Aherne GW, et al. (2006) Trans-activation of the DNA-damage signalling protein kinase Chk2 by T-loop exchange. *EMBO J* 25: 3179–3190.
57. Pike AC, Rellos P, Niesen FH, Turnbull A, Oliver AW, et al. (2008) Activation segment dimerization: a mechanism for kinase autophosphorylation of non-consensus sites. *EMBO J* 27: 704–714.
58. Noble ME, Endicott JA, Johnson LN (2004) Protein kinase inhibitors: insights into drug design from structure. *Science* 303: 1800–1805.
59. Rudolph MI, Cabanillas A, Gomez P, Garcia MA, Villan L (1997) On the mechanism of action of ethodin in inducing myometrium contractions. *Gen Pharmacol* 28: 381–385.
60. Rudolph MI, de los Angeles García M, Sepulveda M, Brandon E, Reinicke K, et al. (1997) Ethodin: pharmacological evidence of the interaction between smooth muscle and mast cells in the myometrium. *J Pharmacol Exp Ther* 282: 256–261.
61. Bencharit S, Morton CL, Hyatt JL, Kuhn P, Danks MK, et al. (2003) Crystal structure of human carboxylesterase 1 complexed with the Alzheimer's drug tacrine: from binding promiscuity to selective inhibition. *Chem Biol* 10: 341–349.
62. Burke JR, Pattoli MA, Gregor KR, Brassil PJ, MacMaster JF, et al. (2003) BMS-345541 is a highly selective inhibitor of I κ B kinase that binds at an allosteric site of the enzyme and blocks NF- κ B-dependent transcription in mice. *J Biol Chem* 278: 1450–1456.
63. Manning G, Whyte DB, Martínez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. *Science* 298: 1912–1934.
64. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, et al. (2006) Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* Chapter 5: Unit 5.6.
65. Bonneau R, Baker D (2001) *Ab initio* protein structure prediction: progress and prospects. *Annu Rev Biophys Biomol Struct* 30: 173–189.
66. Melo F, Feytmans E (1998) Assessing protein structures with a non-local atomic interaction energy. *J Mol Biol* 277: 1141–1152.
67. Zsoldos Z, Reid D, Simon A, Sadjad BS, Johnson AP (2006) eHiTS: an innovative approach to the docking and scoring function problems. *Curr Protein Pept Sci* 7: 421–435.
68. Wolber G, Langer T (2005) LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J Chem Inf Model* 45: 160–169.
69. Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, et al. (2006) PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J Comput Aided Mol Des* 20: 647–671.
70. Mills N (2006) ChemDraw ultra 10.0. *J Am Chem Soc* 128: 13649–13650.
71. Lagorce D, Sperandio O, Galons H, Miteva MA, Villoutreix BO (2008) FAF-Drugs2: free ADME/tox filtering tool to assist drug discovery and chemical biology projects. *BMC Bioinformatics*. doi:10.1186/1471-2105-9-396.
72. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 46: 3–26.
73. Rush TS, 3rd, Grant JA, Mosyak L, Nicholls A (2005) A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J Med Chem* 48: 1489–1495.
74. Naylor E, Arredouani A, Vasudevan SR, Lewis AM, Parkesh R, et al. (2009) Identification of a chemical probe for NAADP by virtual screening. *Nat Chem Biol* 5: 220–226.

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

SUPPORTING INFORMATION

Manuscript 1

Identification of Human IKK-2 Inhibitors of Natural Origin (part I): Modeling of the IKK-2 Kinase Domain, Virtual Screening and Activity Assays

Esther Sala, Laura Guasch, Justyna Iwaszkiewicz, Miquel Mulero, Maria-Josepa Salvadó, Montserrat Pinent, Vincent Zoete, Aurélien Grosdidier, Santiago Garcia-Vallvé, Olivier Michielin and Gerard Pujadas

PLoS ONE 6(2): e16903. doi:10.1371/journal.pone.0016903
(February 2011)

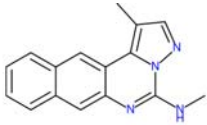
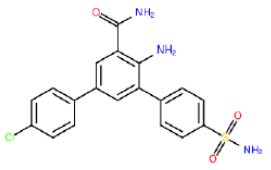
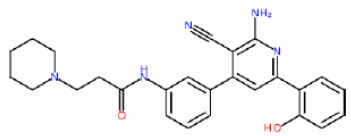
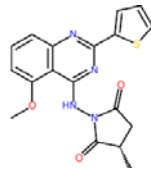
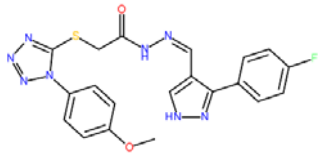
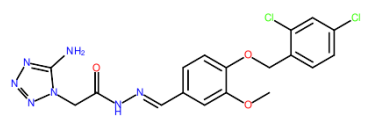
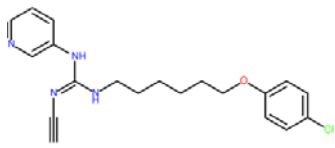
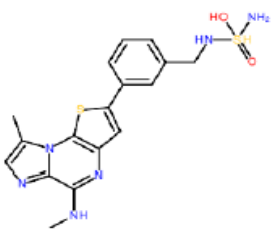
UNIVERSITAT ROVIRA I VIRGILI

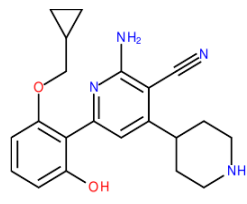
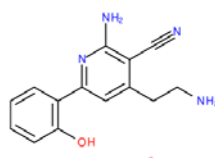
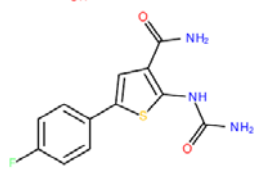
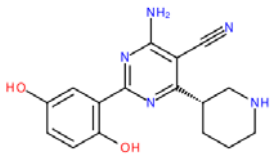
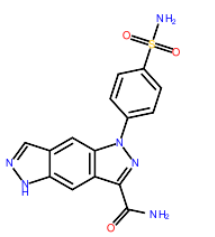
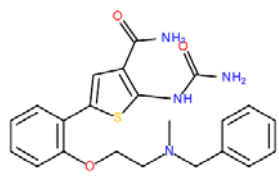
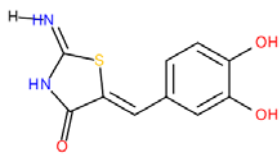
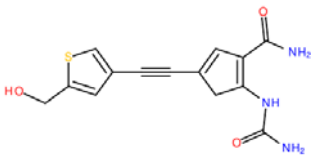
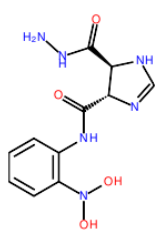
IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

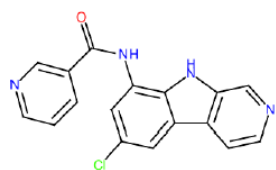
Esther Sala Argüello

ISBN:/DL:T. 1030-2011

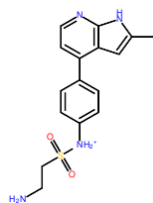
Table S1. hIKK-2 inhibitors from different chemical families docked with the hIKK-2 homology model.

| 2D structure | Molecule name | pIC ₅₀ | cluster | Knowledge-based coherent docking poses |
|---|-----------------------|-------------------|---------|--|
|  | 13a [1] | 7,96 | 6 | 3 |
|  | 8h [2] | 7.0 | 6 | 0 |
|  | 26 [3] | 6,22 | 6 | 0 |
|  | SCP-839 [4] | 8,10 | 6 | 0 |
|  | M8012-3312 [4] | ND | 6 | 0 |
|  | M8008-5430 [4] | ND | 6 | 0 |
|  | CHS-828 [4] | 8,10 | 6 | 0 |
|  | 22d [5] | 8,52 | 6 | 1 |

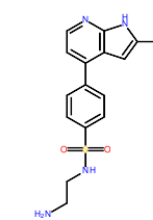
| | | | | |
|---|-------------------------|------|---|---|
|  | 4j [6] | 8,07 | 6 | 1 |
|  | 4a [7] | 7,15 | 6 | 4 |
|  | SKB-TPCA1 [4] | 7,74 | 6 | 3 |
|  | Pharmacia [4] | 6,50 | 6 | 2 |
|  | Pharmacia_02 [4] | 6,17 | 6 | 4 |
|  | Astra Zeneca [4] | 7,40 | 6 | 1 |
|  | M4891-3155 [4] | 5,75 | 5 | 2 |
|  | 1b [8] | 6,71 | 9 | 2 |
|  | M2295-0236 [4] | ND | 6 | 0 |



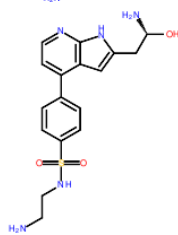
PS-1145 [9] 6,82 6 0



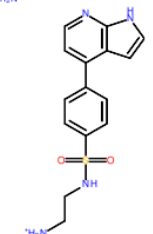
14 [10] 7,8 6 3



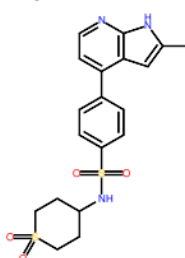
13 [10] 7,7 6 2



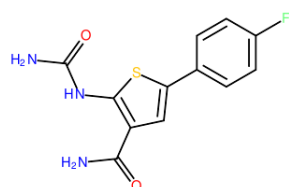
12 [10] 7,7 6 4



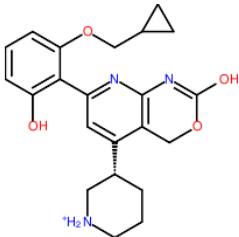
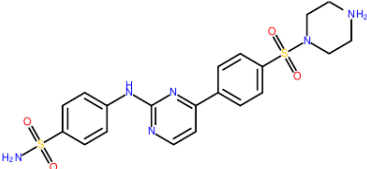
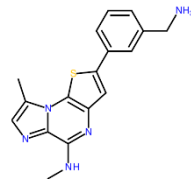
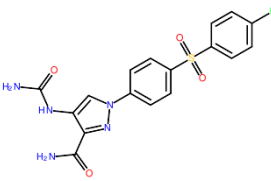
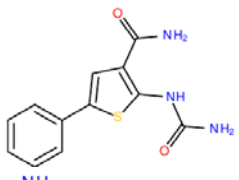
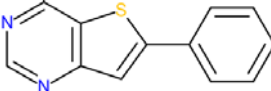
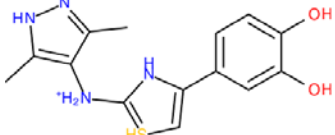
1 [10] 7,4 6 1

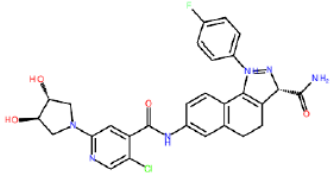
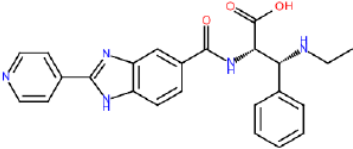
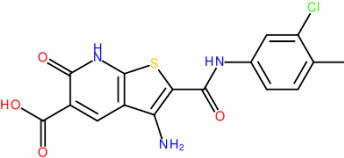
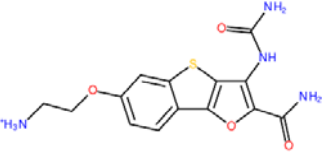
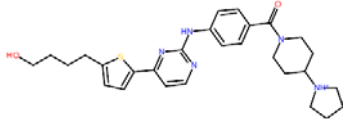


17 [10] 7,7 6 0



Ureido [9] 7,74 6 1

| | | | | |
|---|------------------------|------|---|---|
|  | Compound A [11] | 8,4 | 6 | 0 |
|  | 8u [12] | 7,96 | 8 | 0 |
|  | 22a [5] | 8,07 | 7 | 1 |
|  | 204 [13] | 7,88 | 8 | 0 |
|  | 13 [14] | 7,88 | 8 | 2 |
|  | 12 [15] | 6,01 | 8 | 2 |
|  | M7790-1103 [4] | ND | 8 | 0 |

| | | | | |
|---|-----------------------|------|----|---|
|  | PF-184 [16] | 7,43 | 6 | 0 |
|  | Aventis [4] | 7,15 | 6 | 0 |
|  | M4296-0831 [4] | ND | 10 | 1 |
|  | 16k [17] | 7,96 | 8 | 2 |
|  | 24 [18] | 7,60 | 9 | 1 |

These 36 hIKK-2 inhibitors were docked with our hIKK-2 homology model without imposing constraints that forced poses to make specific intermolecular interactions with the target. Next, the resulting hIKK-2 complexes were analyzed with the help of LigandScout to determine which complexes exhibited target-inhibitor intermolecular interactions equivalent to those described in prior studies. This knowledge-based analysis enabled us to identify at least one knowledge-based coherent pose (43 in total) for 21 out of the 36 hIKK-2 inhibitors assayed (regardless of their scoring by eHiTS; see their values in the *Knowledge-based coherent docking poses* column). By analyzing the chemical features used by each pose in its intermolecular interaction with hIKK-2, a common pharmacophore was derived that describes the mechanism of the ligand-target interaction. The *Cluster* column shows the cluster in which each molecule was classified after running a Schrödinger script that clusters molecules based on Tanimoto similarities between MOLPRINT 2D fingerprints (using the Knime v.2.0.3 module in the Schrödinger software package). The molecules distributed in these clusters are the natural

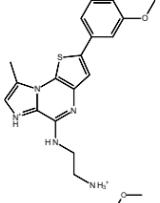
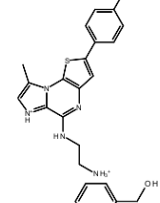
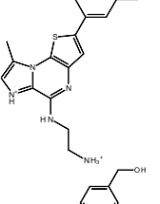
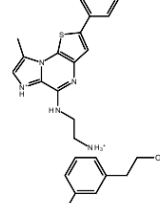
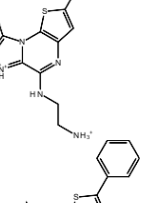
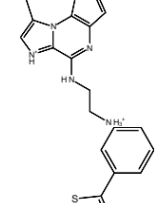
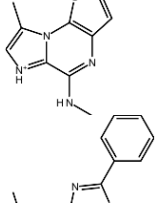
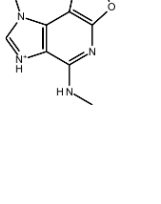
products obtained as hits in our virtual-screening protocol and all known hIKK-2 inhibitors used in the present work [either for validation (see Table S2) or for pharmacophore-generation purposes]. The pIC₅₀ values were obtained from the literature.

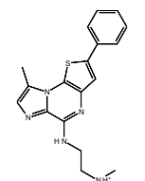
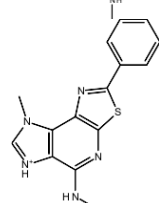
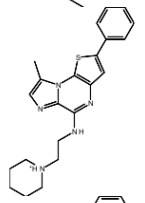
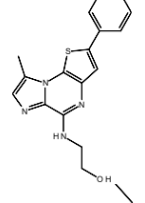
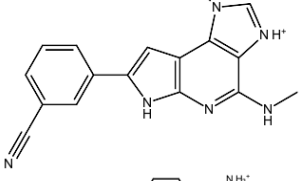
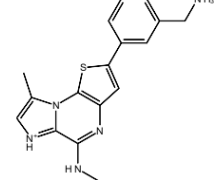
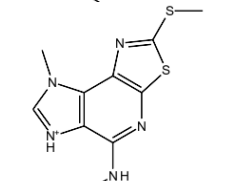
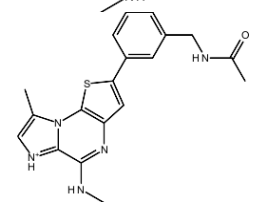
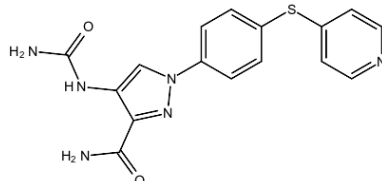
REFERENCES

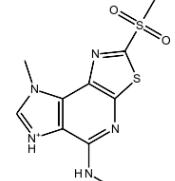
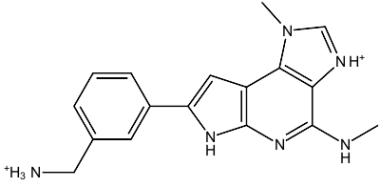
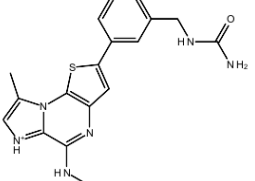
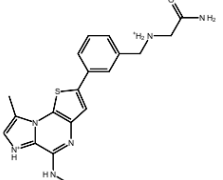
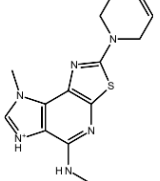
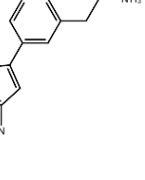
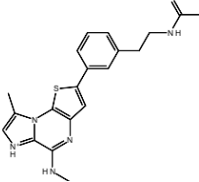
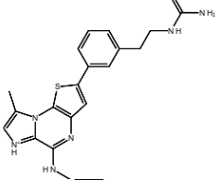
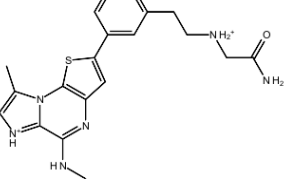
1. Beaulieu F, Ouellet C, Ruediger EH, Belema M, Qiu YP, et al. (2007) Synthesis and biological evaluation of 4-amino derivatives of benzimidazoquinoxaline, benzimidazoquinoline, and benzopyrazoloquinazoline as potent IKK inhibitors. *Bioorg Med Chem Lett* 17: 1233-1237.
2. Christopher JA, Avitabile BG, Bamborough P, Champigny AC, Cutler GJ, et al. (2007) The discovery of 2-amino-3,5-diarylbenzamide inhibitors of IKK- α and IKK- β kinases. *Bioorg Med Chem Lett* 17: 3972-3977.
3. Murata T, Shimada M, Sakakibara S, Yoshino T, Kadono H, et al. (2003) Discovery of novel and selective IKK- β serine-threonine protein kinase inhibitors. Part 1. *Bioorg Med Chem Lett* 13: 913-918.
4. Nagarajan S, Doddareddy M, Choo H, Cho YS, Oh KS, et al. (2009) IKK β inhibitors identification part I: homology model assisted structure based virtual screening. *Bioorg Med Chem Lett* 17: 2759-2766.
5. Belema M, Bunker A, Nguyen V, Beaulieu F, Ouellet C, et al. (2007) Synthesis and structure-activity relationship of imidazo(1,2-a)thieno(3,2-e)pyrazines as IKK- β inhibitors. *Bioorg Med Chem Lett* 17: 4284-4289.
6. Murata T, Shimada M, Sakakibara S, Yoshino T, Masuda T, et al. (2004) Synthesis and structure-activity relationships of novel IKK- β inhibitors. Part 3: Orally active anti-inflammatory agents. *Bioorg Med Chem Lett* 14: 4019-4022.
7. Murata T, Shimada M, Kadono H, Sakakibara S, Yoshino T, et al. (2004) Synthesis and structure-activity relationships of novel IKK- β inhibitors. Part 2: Improvement of *in vitro* activity. *Bioorg Med Chem Lett* 14: 4013-4017.
8. Bonafoux D, Bonar S, Christine L, Clare M, Donnelly A, et al. (2005) Inhibition of IKK-2 by 2-[(aminocarbonyl)amino]-5-acetylenyl-3-thiophenecarboxamides. *Bioorg Med Chem Lett* 15: 2870-2875.
9. Karin M, Yamamoto Y, Wang QM (2004) The IKK NF- κ B system: a treasure trove for drug development. *Nat Rev Drug Discov* 3: 17-26.
10. Liddle J, Bamborough P, Barker MD, Campos S, Cousins RP, et al. (2009) 4-Phenyl-7-azaindoles as potent and selective IKK2 inhibitors. *Bioorg Med Chem Lett* 19: 2504-2508.
11. Ziegelbauer K, Gantner F, Lukacs N, Berlin A, Fuchikami K, et al. (2005) A selective novel low-molecular-weight inhibitor of I κ B kinase- β (IKK- β) prevents pulmonary inflammation and shows broad anti-inflammatory activity. *Br J Pharmacol* 145: 178-192.
12. Bingham AH, Davenport RJ, Gowers L, Knight RL, Lowe C, et al. (2004) A novel series of potent and selective IKK2 inhibitors. *Bioorg Med Chem Lett* 14: 409-412.
13. Clare M, Fletcher T, Hamper BC, Hanson G, Heier RF, et al. (2005) Substituted pyrazole urea compounds for the treatment of inflammation. WO/2005/037797.

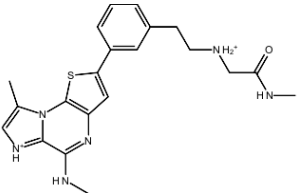
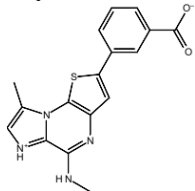
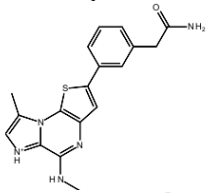
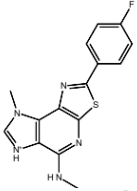
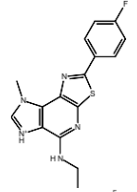
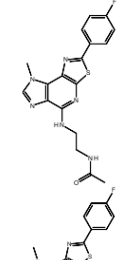
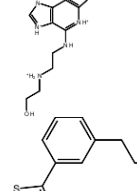
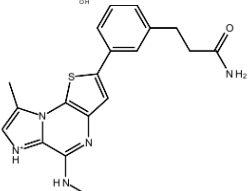
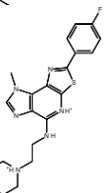
14. Baxter A, Brough S, Cooper A, Floettmann E, Foster S, et al. (2004) Hit-to-lead studies: the discovery of potent, orally active, thiophenecarboxamide IKK-2 inhibitors. *Bioorg Med Chem Lett* 14: 2817-2822.
15. Morwick T, Berry A, Brickwood J, Cardozo M, Catron K, et al. (2006) Evolution of the thienopyridine class of inhibitors of I κ B kinase- β . Part I: hit-to-lead strategies. *J Med Chem* 49: 2898-2908.
16. Sommers CD, Thompson JM, Guzova JA, Bonar SL, Rader RK, et al. (2009) Novel tight-binding inhibitory factor- κ B kinase (IKK-2) inhibitors demonstrate target-specific anti-inflammatory activities in cellular assays and following oral and local delivery in an *in vivo* model of airway inflammation. *J Pharmacol Exp Ther* 330: 377-388.
17. Sugiyama H, Yoshida M, Mori K, Kawamoto T, Sogabe S, et al. (2007) Synthesis and structure activity relationship studies of benzothieno[3,2-b]furan derivatives as a novel class of IKK β inhibitors. *Chem Pharm Bull (Tokyo)* 55: 613-624.
18. Waelchli R, Bollbuck B, Bruns C, Buhl T, Eder J, et al. (2006) Design and preparation of 2-benzamido-pyrimidines as inhibitors of IKK. *Bioorg Med Chem Lett* 16: 108-112.

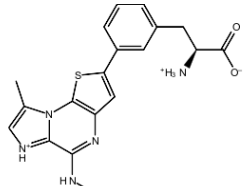
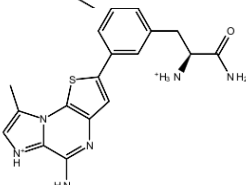
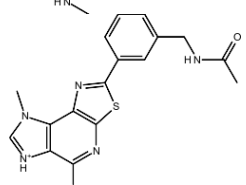
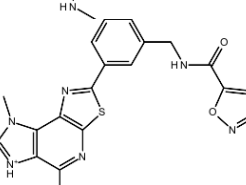
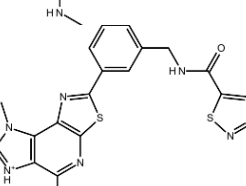
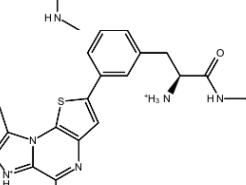
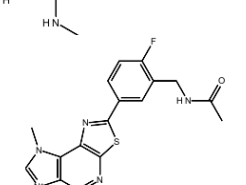
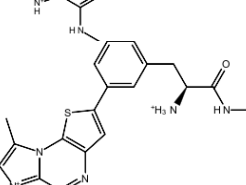
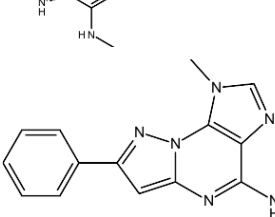
Table S2. hIKK-2 inhibitors used during the validation of the virtual-screening workflow.

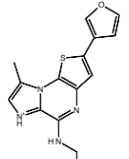
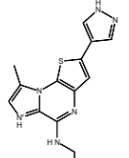
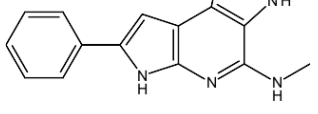
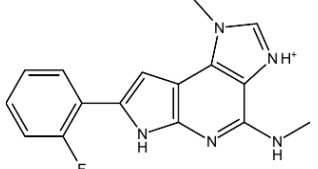
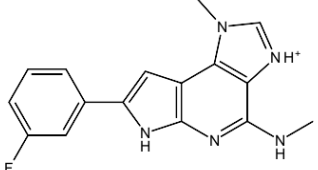
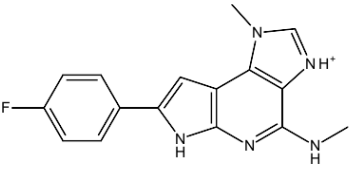
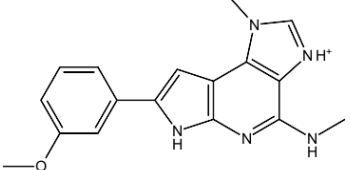
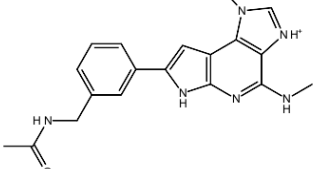
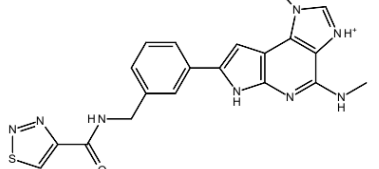
| 2D structure | Molecule name | pIC ₅₀ | cluster |
|---|---------------|-------------------|---------|
|  | 01 [1] | 6.33 | 4 |
|  | 03 [1] | 7.02 | 4 |
|  | 05 [1] | 8.09 | 4 |
|  | 06 [1] | 7.64 | 4 |
|  | 07 [1] | 7.68 | 4 |
|  | 08 [1] | 7.66 | 4 |
|  | 09 [1] | 7.57 | 4 |
|  | 12 [2] | 7.99 | 4 |

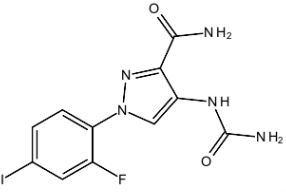
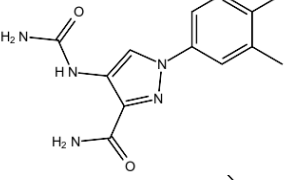
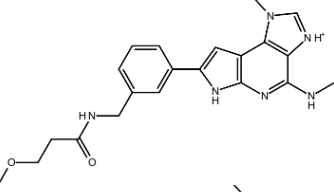
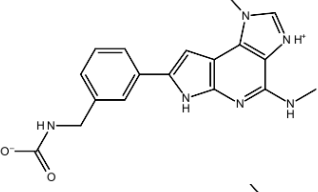
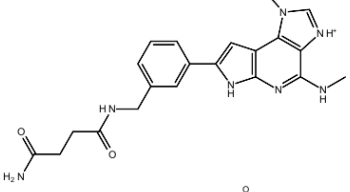
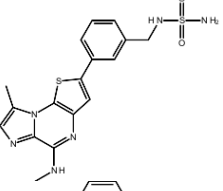
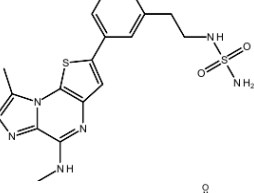
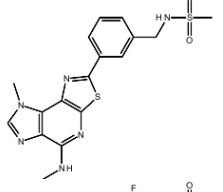
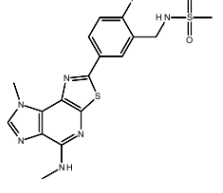
| | | | |
|---|----------------|------|---|
|  | 12 [1] | 7.02 | 4 |
|  | 14 [2] | 8.22 | 4 |
|  | 16 [1] | 7.40 | 4 |
|  | 17 [1] | 7.07 | 4 |
|  | 18 [3] | 8.15 | 4 |
|  | 18 [1] | 8.07 | 4 |
|  | 19 [2] | 7.55 | 4 |
|  | 19 [1] | 7.42 | 4 |
|  | 206 [4] | 7.29 | 8 |

| | | | |
|---|----------------|------|---|
|  | 20 [2] | 7.09 | 4 |
|  | 20 [3] | 7.96 | 4 |
|  | 20 [1] | 7.35 | 4 |
|  | 22 [1] | 7.64 | 4 |
|  | 23d [2] | 7.48 | 4 |
|  | 23 [1] | 7.35 | 4 |
|  | 24 [1] | 7.68 | 4 |
|  | 25 [1] | 7.74 | 4 |
|  | 27 [1] | 7.88 | 4 |

| | | | |
|---|----------------|------|---|
|  | 28 [1] | 7.29 | 4 |
|  | 29 [1] | 6.72 | 4 |
|  | 31 [1] | 8.09 | 4 |
|  | 32a [2] | 7.58 | 4 |
|  | 32b [2] | 7.41 | 4 |
|  | 32d [2] | 7.48 | 4 |
|  | 32e [2] | 7.12 | 4 |
|  | 32 [1] | 7.72 | 4 |
|  | 32f [2] | 7.25 | 4 |

| | | | |
|---|----------------|------|---|
|  | 33 [1] | 7.88 | 4 |
|  | 34 [1] | 7.57 | 4 |
|  | 36a [2] | 7.66 | 4 |
|  | 36c [2] | 7.60 | 4 |
|  | 36d [2] | 7.80 | 4 |
|  | 36 [1] | 7.22 | 4 |
|  | 37a [2] | 7.27 | 4 |
|  | 37 [1] | 7.44 | 4 |
|  | 3b [2] | 7.47 | 4 |

| | | | |
|---|---------------|------|---|
|  | 40 [1] | 7.54 | 4 |
|  | 43 [1] | 8 | 4 |
|  | 4a [3] | 7.40 | 4 |
|  | 4b [3] | 7.36 | 4 |
|  | 4c [3] | 8.15 | 4 |
|  | 4d [3] | 7.36 | 4 |
|  | 4e [3] | 8.09 | 4 |
|  | 4f [3] | 7.85 | 4 |
|  | 4h [3] | 7.41 | 4 |

| | | | |
|---|----------------|------|---|
|  | 210 [4] | 7.02 | 8 |
|  | 8 [4] | 7.17 | 8 |
|  | 4k [3] | 7.92 | 4 |
|  | 4g [3] | 7.60 | 4 |
|  | 4j [3] | 7.66 | 4 |
|  | 21 [1] | 8.52 | 4 |
|  | 26 [1] | 8.15 | 4 |
|  | 36b [2] | 8.30 | 4 |
|  | 37b [2] | 7.50 | 4 |

These 62 hIKK-2 inhibitors (different from the 36 used during the structure-based pharmacophore generation; see Table S1) were used to test the ability of the virtual-screening workflow to identify hIKK-2 inhibitors in a database of molecules. The *Cluster* column shows the cluster into which each molecule was classified after running a Schrödinger script that clusters molecules based on Tanimoto similarities between MOLPRINT 2D fingerprints (using the Knime v.2.0.3 module in the Schrödinger software package). The molecules distributed in these clusters are the natural products obtained as hits in our virtual-screening protocol and all known hIKK-2 inhibitors used in the present work (either for validation or for pharmacophore-generation purposes). The pIC₅₀ values were obtained from the literature.

REFERENCES

1. Wei Long PL, Xinru Li, Yang Xu, Jie Yu, Shitang Ma, Lingling Yu, Zhongmei Zou (2009) QSAR studies on imidazothienopyrazines as IKK- β inhibitors: from 2D to 3D. *Journal of Chemometrics* 23: 304-314.
2. Kempson J, Spergel SH, Guo J, Quesnelle C, Gill P, et al. (2009) Novel tricyclic inhibitors of I κ B kinase. *J Med Chem* 52: 1994-2005.
3. Kempson J, Guo J, Das J, Moquin RV, Spergel SH, et al. (2009) Synthesis, initial SAR and biological evaluation of 1,6-dihydroimidazo[4,5-d]pyrrolo[2,3-b]pyridin-4-amine derived inhibitors of I κ B kinase. *Bioorg Med Chem Lett* 19: 2646-2649.
4. Clare M, Fletcher T, Hamper BC, Hanson G, Heier RF, et al. (2005) Substituted pyrazole urea compounds for the treatment of inflammation. WO/2005/037797.

Table S3. Scaffold-hopping candidates for hIKK-2 inhibition predicted by our study.

| Cluster 1 hits | | | | Cluster 2 hits | | | | Cluster 3 hits | | | |
|----------------|-------|-------|--------|----------------------|--------------|--------------|--------------|----------------|-------|-------|--------|
| PB | Shape | Combo | hIKK-2 | PB | Shape | Combo | hIKK-2 | PB | Shape | Combo | hIKK-2 |
| ZINC12410246 | 0.699 | 0.528 | 1.227 | ZINC08298038 | 0.597 | 0.349 | 0.946 | ZINC00048544 | 0.723 | 0.480 | 1.203 |
| ZINC04222236 | 0.727 | 0.498 | 1.224 | ZINC08299000 | 0.633 | 0.308 | 0.941 | ZINC03838566 | 0.704 | 0.475 | 1.179 |
| ZINC08299593 | 0.723 | 0.462 | 1.185 | ZINC04260500 | 0.609 | 0.325 | 0.934 | ZINC12853223 | 0.485 | 0.426 | 0.911 |
| ZINC03000807 | 0.606 | 0.560 | 1.166 | ZINC03051724 | 0.520 | 0.414 | 0.933 | ZINC00034693 | 0.585 | 0.606 | 0.910 |
| ZINC08300706 | 0.697 | 0.460 | 1.157 | ZINC005415683 | 0.585 | 0.445 | 0.929 | ZINC09231861 | 0.424 | 0.474 | 0.898 |
| ZINC05739081 | 0.682 | 0.451 | 1.134 | ZINC08635779 | 0.579 | 0.349 | 0.928 | ZINC03871771 | 0.627 | 0.264 | 0.891 |
| ZINC08298517 | 0.645 | 0.489 | 1.134 | ZINC08298591 | 0.516 | 0.412 | 0.928 | ZINC13404337 | 0.341 | 0.662 | 0.890 |
| ZINC04235841 | 0.626 | 0.498 | 1.124 | ZINC08298996 | 0.497 | 0.429 | 0.926 | ZINC05224529 | 0.511 | 0.376 | 0.888 |
| ZINC05433728 | 0.638 | 0.467 | 1.105 | ZINC08298463 | 0.580 | 0.343 | 0.924 | ZINC00517451 | 0.381 | 0.504 | 0.885 |
| ZINC08298452 | 0.716 | 0.382 | 1.098 | ZINC00058208 | 0.261 | 0.661 | 0.923 | ZINC01317724 | 0.224 | 0.660 | 0.884 |
| ZINC08635299 | 0.677 | 0.414 | 1.091 | ZINC08298131 | 0.538 | 0.378 | 0.917 | ZINC038762138 | 0.497 | 0.284 | 0.883 |
| ZINC08299522 | 0.632 | 0.458 | 1.091 | ZINC08299131 | 0.538 | 0.378 | 0.917 | ZINC03786448 | 0.250 | 0.633 | 0.883 |
| ZINC12604815 | 0.542 | 0.547 | 1.089 | ZINC08297955 | 0.444 | 0.472 | 0.917 | ZINC0224397 | 0.553 | 0.329 | 0.881 |
| ZINC08298232 | 0.618 | 0.457 | 1.075 | ZINC08300143 | 0.629 | 0.287 | 0.916 | ZINC00039884 | 0.342 | 0.536 | 0.878 |
| ZINC08298456 | 0.627 | 0.441 | 1.068 | ZINC03841142 | 0.575 | 0.341 | 0.916 | ZINC0020210 | 0.388 | 0.555 | 0.943 |
| ZINC05433715 | 0.601 | 0.466 | 1.068 | ZINC00195221 | 0.491 | 0.424 | 0.915 | ZINC00526037 | 0.456 | 0.483 | 0.939 |
| ZINC03841282 | 0.686 | 0.380 | 1.066 | ZINC02023740 | 0.175 | 0.739 | 0.914 | ZINC00518532 | 0.397 | 0.537 | 0.934 |
| ZINC04029488 | 0.633 | 0.429 | 1.062 | ZINC01204686 | 0.644 | 0.267 | 0.911 | ZINC00519161 | 0.339 | 0.593 | 0.932 |
| ZINC08296915 | 0.695 | 0.361 | 1.056 | ZINC08298449 | 0.588 | 0.321 | 0.910 | ZINC04280948 | 0.342 | 0.581 | 0.923 |
| ZINC04222674 | 0.515 | 0.540 | 1.055 | ZINC04082733 | 0.528 | 0.381 | 0.910 | ZINC13404329 | 0.282 | 0.638 | 0.920 |
| ZINC08298611 | 0.629 | 0.425 | 1.054 | ZINC01669260 | 0.512 | 0.397 | 0.909 | ZINC13110434 | 0.322 | 0.588 | 0.910 |
| ZINC04235840 | 0.629 | 0.420 | 1.049 | ZINC04222254 | 0.483 | 0.426 | 0.909 | ZINC02136236 | 0.621 | 0.288 | 0.909 |
| ZINC08300121 | 0.641 | 0.407 | 1.048 | ZINC08298970 | 0.595 | 0.313 | 0.907 | ZINC03842020 | 0.556 | 0.346 | 0.901 |
| ZINC03841201 | 0.608 | 0.439 | 1.048 | ZINC03841141 | 0.595 | 0.307 | 0.902 | ZINC04261011 | 0.481 | 0.420 | 0.901 |
| ZINC08297952 | 0.648 | 0.394 | 1.043 | ZINC08298553 | 0.563 | 0.337 | 0.901 | ZINC00519153 | 0.270 | 0.628 | 0.898 |
| ZINC05433681 | 0.637 | 0.403 | 1.040 | ZINC08298686 | 0.327 | 0.573 | 0.900 | ZINC00518538 | 0.193 | 0.705 | 0.898 |
| ZINC04235837 | 0.603 | 0.424 | 1.027 | ZINC08298004 | 0.441 | 0.458 | 0.899 | ZINC123441518 | 0.336 | 0.561 | 0.896 |
| ZINC08298759 | 0.530 | 0.490 | 1.020 | ZINC05397696 | 0.467 | 0.432 | 0.898 | ZINC00517656 | 0.338 | 0.556 | 0.894 |
| ZINC04260104 | 0.648 | 0.370 | 1.018 | ZINC012604970 | 0.593 | 0.303 | 0.896 | ZINC00525794 | 0.591 | 0.302 | 0.893 |
| ZINC08298937 | 0.698 | 0.319 | 1.015 | ZINC00485710 | 0.468 | 0.427 | 0.896 | ZINC00189849 | 0.360 | 0.529 | 0.890 |
| ZINC08298033 | 0.662 | 0.346 | 1.009 | ZINC04235871 | 0.415 | 0.481 | 0.896 | ZINC03842034 | 0.515 | 0.369 | 0.884 |
| ZINC00460033 | 0.504 | 0.500 | 1.004 | ZINC04236053 | 0.519 | 0.373 | 0.892 | ZINC03844523 | 0.284 | 0.599 | 0.883 |
| ZINC04235839 | 0.618 | 0.379 | 0.997 | ZINC08298689 | 0.415 | 0.476 | 0.891 | ZINC13402109 | 0.425 | 0.449 | 0.875 |
| ZINC04222682 | 0.591 | 0.403 | 0.993 | ZINC0056225 | 0.231 | 0.660 | 0.891 | ZINC04504120 | 0.448 | 0.418 | 0.866 |
| ZINC08254138 | 0.661 | 0.331 | 0.992 | ZINC05434223 | 0.476 | 0.412 | 0.888 | ZINC00460974 | 0.553 | 0.310 | 0.863 |
| ZINC0035451 | 0.679 | 0.313 | 0.991 | ZINC08254330 | 0.641 | 0.243 | 0.883 | ZINC00178183 | 0.239 | 0.622 | 0.861 |
| ZINC05398170 | 0.589 | 0.402 | 0.991 | ZINC08298765 | 0.364 | 0.518 | 0.882 | ZINC03850520 | 0.518 | 0.342 | 0.860 |
| ZINC08254141 | 0.678 | 0.312 | 0.990 | ZINC08298536 | 0.519 | 0.362 | 0.881 | ZINC00518609 | 0.227 | 0.633 | 0.860 |
| ZINC03836891 | 0.598 | 0.390 | 0.989 | ZINC08298629 | 0.543 | 0.335 | 0.878 | ZINC00296237 | 0.429 | 0.424 | 0.853 |
| ZINC00485869 | 0.604 | 0.381 | 0.986 | ZINC05433729 | 0.385 | 0.492 | 0.877 | ZINC00042159 | 0.260 | 0.591 | 0.851 |
| ZINC08298569 | 0.589 | 0.396 | 0.985 | ZINC05279694 | 0.487 | 0.384 | 0.871 | | | | |
| ZINC00058275 | 0.329 | 0.657 | 0.985 | ZINC04235956 | 0.581 | 0.288 | 0.869 | | | | |
| ZINC08296946 | 0.689 | 0.295 | 0.984 | ZINC08296677 | 0.487 | 0.382 | 0.868 | | | | |
| ZINC04236050 | 0.505 | 0.479 | 0.984 | ZINC04236052 | 0.495 | 0.373 | 0.868 | | | | |
| ZINC04222681 | 0.591 | 0.391 | 0.982 | ZINC05433716 | 0.381 | 0.487 | 0.868 | | | | |
| ZINC00035452 | 0.592 | 0.388 | 0.979 | ZINC04236049 | 0.501 | 0.365 | 0.866 | | | | |
| ZINC08300479 | 0.535 | 0.443 | 0.978 | ZINC08300122 | 0.496 | 0.370 | 0.866 | | | | |
| ZINC08298635 | 0.641 | 0.336 | 0.977 | ZINC01601221 | 0.204 | 0.662 | 0.866 | | | | |
| ZINC08300733 | 0.613 | 0.363 | 0.976 | ZINC08298907 | 0.638 | 0.227 | 0.865 | | | | |
| ZINC12604292 | 0.547 | 0.424 | 0.971 | ZINC04235961 | 0.592 | 0.272 | 0.864 | | | | |
| ZINC03836874 | 0.578 | 0.390 | 0.968 | ZINC08254273 | 0.509 | 0.352 | 0.861 | | | | |
| ZINC08254375 | 0.562 | 0.405 | 0.966 | ZINC16946275 | 0.504 | 0.357 | 0.861 | | | | |
| ZINC08298030 | 0.668 | 0.296 | 0.963 | ZINC04236055 | 0.517 | 0.343 | 0.860 | | | | |
| ZINC0485709 | 0.386 | 0.565 | 0.961 | ZINC04236051 | 0.509 | 0.350 | 0.859 | | | | |
| ZINC05433719 | 0.494 | 0.462 | 0.957 | ZINC04235957 | 0.582 | 0.274 | 0.856 | | | | |
| ZINC04260127 | 0.594 | 0.360 | 0.954 | ZINC08298002 | 0.565 | 0.301 | 0.855 | | | | |
| ZINC08298652 | 0.570 | 0.382 | 0.953 | ZINC12604752 | 0.439 | 0.416 | 0.856 | | | | |
| ZINC08254451 | 0.664 | 0.287 | 0.951 | ZINC03841193 | 0.256 | 0.598 | 0.854 | | | | |
| ZINC04260095 | 0.624 | 0.327 | 0.951 | ZINC04236054 | 0.503 | 0.350 | 0.853 | | | | |
| ZINC08298659 | 0.619 | 0.328 | 0.947 | ZINC00048191 | 0.498 | 0.355 | 0.853 | | | | |
| | | | | ZINC08298638 | 0.616 | 0.235 | 0.851 | | | | |

The ZINC codes for the 246 hit molecules predicted to inhibit hIKK-2 and belonging to clusters consisting exclusively of natural products. For each hit molecule, the best results of the shape and electrostatic-potential comparisons with 43 poses from 21 known hIKK-2 inhibitors (see Table S1) is shown. Thus, the Tanimoto values for the comparison between the electrostatic potentials of the molecules (using an outer dielectric of 80) are shown in the *PB* columns, whereas the values for the comparison between shapes are shown in the *Shape* columns. The sum of the *PB* and *Shape* values is reported in the *Combo* columns. Hits from each cluster are sorted according their decreasing combo value. All of these hit molecules are scaffold-hopping candidates for hIKK-2 inhibition because the Tanimoto similarities between their MOLPRINT 2D fingerprints and those from the hIKK-2 inhibitors in Tables S1 and S2 are quite low. ZINC00058225, ZINC01669260 and ZINC16946275 from Cluster 1, ZINC03683886 from Cluster 2, and ZINC03871389 from Cluster 3 were selected to experimentally test the success rate of our predictions using an *in vitro* assay (in bold in Table S3). The results of this experiment showed that three out of the five molecules (*i.e.*, ZINC01669260 from Cluster 1, ZINC03683886 from Cluster 2 and ZINC03871389 from Cluster 3) inhibited hIKK-2, with IC₅₀ values ranging from 183.8 to 3,325 µM.

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

Manuscript 2

Identification of IKK-2 Inhibitors of Natural Origin
(part II): In silico prediction of IKK-2 inhibitors in
natural extracts with known anti-inflammatory
activity.

Esther Sala, Laura Guasch, Justyna Iwaszkiewicz, Miquel Mulero, Maria-
Josepa Salvadó, Cinta Bladé, Vincent Zoete, Aurélien Grosdidier, Santiago
Garcia-Vallvé, Olivier Michielin and Gerard Pujadas

Submitted to European Journal of Medicinal Chemistry
(3 February 2011)

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

Identification of Human IKK-2 Inhibitors of Natural Origin (Part II): *In Silico* Prediction of IKK-2 Inhibitors in Natural Extracts with Known Anti-Inflammatory Activity.

Esther Sala^a, Laura Guasch^a, Justyna Iwaskiewicz^b, Miquel Mulero^a, Maria-Josepa Salvadó^a, Cinta Bladé^a, Meritxell Ceballos^c, Cristina Valls^a, Vincent Zoete^b, Aurélien Grosdidier^b, Santiago Garcia-Vallvé^{a,c}, Olivier Michielin^b, Gerard Pujadas^{a,c}

^a Grup de Recerca en Nutrigenòmica, Departament de Bioquímica i Biotecnologia, Universitat Rovira i Virgili, Campus de Sescelades, C/ Marcel·lí Domingo s/n, 43007 Tarragona, Catalonia, Spain

^b Molecular Modeling Group. Swiss Institute of Bioinformatics, Quartier UNIL-Sorge. Bâtiment Génopode. CH-1015 Lausanne, Switzerland

^c Centre Tecnològic de Nutrició i Salut (CTNS), TECNIO, CEICS, Camí de Valls, 81-87, 43204 Reus, Catalonia, Spain

Corresponding author: Gerard Pujadas; e-mail: gerard.pujadas@urv.cat; Tel: +34977 559565; Fax: +34977558232; Grup de Recerca en Nutrigenòmica, Departament de Bioquímica i Biotecnologia, Universitat Rovira i Virgili, Campus de Sescelades, C/ Marcel·lí Domingo s/n, 43007 Tarragona, Catalonia, Spain

Short title: hIKK-2 Inhibitors in Anti-Inflammatory Extracts

Abstract

Human inhibitor NF- κ B kinase 2 (hIKK-2) is the primary component responsible for activating NF- κ B in response to various inflammatory stimuli. Thus, synthetic ATP-competitive inhibitors for hIKK-2 have been developed as anti-inflammatory compounds. We recently reported a virtual screening protocol (doi:10.1371/journal.pone.0016903) that is able to identify hIKK-2 inhibitors that are not structurally related to any known molecule that inhibits hIKK-2 and that have never been reported to have anti-inflammatory activity. In this study, a stricter version of this protocol was applied to an in-house database of 29,779 natural products annotated with their natural source. The search identified 274 molecules (isolated from 453 different natural extracts) predicted to inhibit hIKK-2. An exhaustive bibliographic search revealed that anti-inflammatory activity has been previously described for: **(a)** 36 out of these 453 extracts; and **(b)** 17 out of 30 virtual screening hits present in these 36 extracts. Only one of the remaining 13 hit molecules in these extracts shows chemical similarity with known synthetic hIKK-2 inhibitors. Therefore, it is plausible that a significant portion of the remaining 12 hit molecules are lead-hopping candidates for the development of new hIKK-2 inhibitors.

Keywords

IKK- β , chronic inflammation, homology model, pharmacognosy, scaffold hopping, virtual screening.

Abbreviations

hIKK-2, human inhibitor nuclear-factor κ B kinase 2; NF- κ B, nuclear-factor κ B transcription factor; NPs, natural products; VS, virtual screening

1. Introduction

For centuries, the only effective way to relieve pain and cure diseases has been the use (either by oral or topical administration) of natural products (NPs; mainly of vegetal origin) [1]. Thus, the first records of using plants to treat illnesses come from Mesopotamia and date from about 2600 BC. Later, older western civilizations such as the Egyptians (*e.g.*, the *Ebers Papyrus* dating from 1500 BC), the Greeks (*e.g.*, the *History of Plants* from Theophrastus dating from 300 BC) and the Romans (*e.g.*, the *De materia medica* from Pedanius Dioscorides written between 65 and 75 AD) documented naturally occurring drugs. Oriental civilizations such as those in China and India have also extensively documented the medicinal use of endemic plants since ancient times. For example, Wu Shi Er Bing Fang reported 52 prescriptions around 1100 BC, and the Indian Ayurvedic system dates from about 1000 BC. These reports are the basis of traditional medicines that are extensively used today.

One of the most notable effects of certain medicinal plants is their anti-inflammatory activity. For example, the *De materia medica* book from Pedanius Dioscorides describes plants that can be used to treat inflammatory disorders such as rheumatic or arthritic pains, joint swelling, snake bites, and fever [2]. Inflammation is a complex non-specific immune response triggered by damage to living tissues that protects higher organisms from infection and injury [3]. There are two types of inflammation, and their effects can be either beneficial (defense against agents interfering with homeostasis; *i.e.*, acute inflammation) or harmful (causing damage to cells and tissues, *i.e.*, chronic inflammation) [4]. Chronic inflammation is directly involved in the combination of metabolic risk factors (*i.e.*, abdominal obesity, insulin resistance, hypertension, prothrombotic state, and atherogenic dyslipidemia) that is known as metabolic syndrome [5-8]. Thus, identifying and targeting the central molecules involved in integrating metabolic responses with immune/inflammatory responses will hopefully lead to a successful therapy for metabolic syndrome [9]. One target of interest is human inhibitor nuclear-factor κ B kinase 2 (hIKK-2). hIKK-2 is a serine-threonine protein kinase belonging to the IKK complex and is the primary component responsible for activating nuclear-factor κ B transcription factor (NF- κ B) in response to inflammatory stimuli. Indeed, the importance of the NF- κ B pathway for regulating the expression of genes controlling cellular immune and inflammatory responses has motivated research groups in both academia and the pharmaceutical industry to devote increasing efforts toward developing synthetic ATP-competitive inhibitors for hIKK-2 [10-13], which could be of therapeutic use in patients affected by chronic inflammatory diseases [14].

Recently, we successfully used a virtual screening (VS) protocol to screen the ZINC Natural

Products Database (http://wiki.compbio.ucsf.edu/wiki/index.php/Natural_products_database) [15] for hIKK-2 inhibitors of natural origin that compete with ATP [16]. We predicted that 1,061 out of the 89,425 molecules present in the studied database would inhibit hIKK-2 with good ADME/Tox properties. Notably, when these 1,061 molecules were merged with the 98 synthetic hIKK-2 inhibitors used in that study (either for validation or for pharmacophore-generation purposes) and the resulting set was classified into ten clusters according to chemical similarity, there were three clusters that contained only natural products. Five molecules from these three clusters (for which no anti-inflammatory activity was previously described) were then selected for *in vitro* activity testing, and three out of the five molecules were found to inhibit hIKK-2 (one per cluster).

Hence, the goals of the present work were to find **(a)** natural extracts with known anti-inflammatory activity that contain at least one molecule that we predicted to be an hIKK-2 inhibitor and **(b)** hits from our VS protocol that have not been yet described as hIKK-2 inhibitors and that are present in such anti-inflammatory extracts. Thus, in order to achieve these goals, we **(a)** applied a modified version of the VS protocol described above (on that uses stricter conditions during the last VS step) to an in-house database of 29,779 NPs annotated with their corresponding natural source(s) of origin and **(b)** determined which of the VS hits are present in extracts with known anti-inflammatory activity.

2. Results and discussion

2.1 Virtual-screening workflow: validation and application to our in-house NP database

The ability of our modified VS workflow to distinguish between hIKK-2 inhibitors and molecules that do not inhibit hIKK-2 was evaluated by applying it to the same set of 62 known hIKK-2 inhibitors and 10,036 kinase decoys that was used to validate our previous VS workflow [16]. We decided to use stricter conditions in the EON filter (*i.e.*, $ET_{pb} \geq 0.3$, $ET_{shape} \geq 0.5$ and $ET_{coulb} \geq 0.3$ instead of $ET_{combo} \geq 0.850$) to guarantee that the similarity between the hits and the active poses used during this last step of the screening had significant contributions from both shape and electrostatic components. Thus, the new conditions in the last VS step increase the enrichment factor for the EON filter from 4.5 to 8.6 and that for the full VS process from 39.3 to 74.6 (see Figure 1). The new conditions significantly enhance the ability of our VS protocol to distinguish molecules that inhibit hIKK-2 from those that do not affect its activity. Subsequently, it identified 274 out of the 29,779 molecules in our in-house NP database as potential hIKK-2 inhibitors (see Figure 1).

2.2 Virtual-screening hits in natural extracts with known anti-inflammatory activity

According to the information available in our in-house NP database, the 274 molecules that were predicted by the VS workflow as potential hIKK-2 inhibitors have been isolated from 453 different natural sources. Interestingly, a systematic bibliographic search of PubMed (<http://www.pubmed.org>) revealed that anti-inflammatory activity has been described for extracts from 36 out of these 453 different natural sources (see Table 1). These 36 extracts contain 30 VS hit molecules (see Table S1), and therefore it is plausible that, through their role as hIKK-2 inhibitors, these molecules contribute to the observed anti-inflammatory activity of their corresponding extract. In fact, a search with SciFinder (<http://www.cas.org/products/sfacad>) revealed that anti-inflammatory activity has been described for 17 out of these 30 VS hits (see Table S1).

Analysis of these 17 VS hits revealed that some of them are polyphenols (which are secondary metabolic products in plants that are abundant in our diet) whose roles as antioxidants and in the prevention of degenerative diseases, cancer and cardiovascular diseases have been well described [4, 100-103]. The most significant phenolic compounds found in these 36 anti-inflammatory extracts are as follows:

- The biflavonoid ayanin (CAS: 572-32-7; see Table S1), which is found in seven extracts (see Table 1) and has been claimed to be of use in treating allergic asthma [20].
- The isoflavone iristectorigenin A (CAS: 39012-01-6; see Table S1) and the anthocyanin malvidin (CAS: 643-84-5; see Table S1), which show strong inhibitory activity for NO production [44, 49]. Iristectorigenin A is found in *Iris germanica* extracts (see Table 1), whereas malvidin is found in *Lavandula angustifolia* and *Vaccinium myrtillus* extracts (see Table 1).
- The flavone baicalein (CAS: 491-67-8; see Table S1) [56], which is found in *Oroxylum indicum*, *Scutellaria baicalensis* and *Terminalia arjuna* extracts (see Table 1).
- The epicatechin derivate 4 β -(2-aminoethylthiol)epicatechin [97], which is found in *Vitis vinifera* extracts (see Table 1).

In addition to polyphenols, other molecules were found to be among the 17 VS hits present in known anti-inflammatory extracts, such as:

- Aristilactam FI (CAS: 112501-42-5; see Table S1), which is found in *Piper longum* extracts (see Table 1) and has been shown to inhibit NO generation in RAW264.7 macrophages stimulated with lipopolysaccharide [73].

- Harmol (CAS: 487-03-6; see Table S1), which is found in *Passiflora edulis* [62, 63] extracts (see Table 1) and belongs to the β -carbolines (a group of methylated natural product derivatives identified as hIKK-2 inhibitors several times) [104-106].
- 2-tert-butyl-hydroquinone (CAS: 1948-33-0; see Table S1), which is found in *Paeonia suffruticosa* extracts (see Table 1) and is a component of experimental diets which, in synergy with other compounds, reduces atherosclerotic plaque in the aortic sinus [59].

Therefore, it is reasonable to suspect that a substantial portion of the remaining 13 VS hits also act as anti-inflammatory agents by inhibiting hIKK-2 and thus contribute, at least partially, to the observed anti-inflammatory activity of the various extracts that contain them. Moreover, when we merged these 13 hits with 377 synthetic hIKK-2 inhibitors and the resulting set was classified according to structural similarity, there was a single molecule (*i.e.*, CAS: 38940-60-2; see Table S1) that shares a cluster with the synthetic molecules (results not shown). Moreover, 4 out of the remaining 12 molecules were grouped into two clusters (CAS: 93973-22-9 and gamm-cotonefuran in one cluster and methyl β -D-apiofuranosyl-(1->6)- β -D-glucopyranoside and 1,4-dideoxy-1,4-imino-(2-O- β -D-glucopyranosyl)-D-arabinitol in the other), and the rest have unrelated chemical structures. Thus, these 12 predicted hIKK-2 inhibitors show 10 different chemical scaffolds that are very different from those presently known as synthetic hIKK-2 inhibitors.

3. Conclusions

In a previous study [16], we developed a VS workflow that was able to successfully discern between molecules that inhibit hIKK-2 and molecules that do not inhibit this enzyme. We experimentally proved that our VS protocol was able to identify hIKK-2 inhibitors that **(a)** were not structurally related to any known synthetic molecule that inhibits hIKK-2 and **(b)** have never been reported to have anti-inflammatory activity. In the present work we applied the same VS workflow using more strict conditions in the last filter step to an in-house database of 29,779 NPs annotated with their corresponding natural source(s). According to the validation results (see Figure 1), the more restrictive conditions of the current EON filter increased the success rate of our predictions. Interestingly, our screen predicted 30 hIKK-2 inhibitors that are present in natural extracts with known anti-inflammatory activity, and 17 of these molecules have been described as anti-inflammatory. Among the remaining 13 NP hits, 12 show no chemical similarity either between themselves or with 377 synthetic hIKK-2 inhibitors. Thus, it can be concluded that a significant portion of these 12 NPs could be lead-hopping candidates for the development of new hIKK-2 inhibitors.

Additionally, we predicted that there are 414 other extracts with undescribed anti-inflammatory activity that contain at least one out of 274 VS hits. Consequently, our work opens the door to the discovery of new anti-inflammatory extracts of natural origin that could be of use, for example, in the design of functional foods aimed at preventing diseases that are the result of chronic inflammation. Therefore, the characterization of such extracts merits further attention, and current work in this regard is underway.

4. Experimental section

4.1 Virtual-screening workflow

The VS workflow used in this work is the same as that described previously [16] except that the conditions of the last filter were more restrictive.

The first step of the VS protocol consisted of an ADME/Tox filter performed with the FAF-Drugs2 tool [107], which aimed to discard molecules that either had poor ADME properties or were potentially toxic. Thus, the drug-like properties of a compound were evaluated using the Lipinski rule of five [108], and only one violation of this rule was allowed. Then, molecules containing toxic groups were filtered using the 204 substructures for “warhead” chelators, frequent hitters, promiscuous inhibitors and other undesirable functional groups available in FAF-Drugs2 tool [107]. Molecules with appropriate ADME/Tox properties were then incorporated into a Phase database, and conformations and sites were generated for them.

The next step of the VS protocol used a *structure-based* pharmacophore that was obtained by docking 36 known selective ATP-competitive inhibitors for hIKK-2 from different chemical families into a hIKK-2 homology model. The docking was done with eHiTS v2009 (SimBioSys Inc., Toronto, Canada; <http://www.simbiosys.ca/ehits>) [109] by considering the receptor to be a rigid body and the ligands as flexible (*i.e.*, free rotation was allowed around the single bonds of the ligand). Default docking conditions were selected except for the size of the sides of the cubic box encompassing the ATP-binding pocket, which was increased from 10 Å to 15 Å. Then, the protein-ligand interactions of the poses obtained were analyzed with LigandScout v2.03 (Inteligand GmbH, Vienna, Austria, <http://www.inteligand.com/ligandscout>) [110], and the poses that satisfied the generic binding features of ATP-competitive kinase inhibitors were selected as correct (regardless of their eHiTS scores). As a result of this analysis, 43 poses belonging to 21 hIKK-2 inhibitors were identified as correct, and their corresponding sites (*i.e.*, functional groups used by the poses in their intermolecular interactions with the hIKK-2 kinase domain) were used to derive

a common structure-based pharmacophore that contains one hydrogen-bond acceptor, two hydrogen-bond donors and one hydrophobic site. Moreover, the pharmacophore was completed with a shell of excluded volumes that schematically represent the location of the residues forming the ATP-binding site in the hIKK-2 homology model. This pharmacophore was used by Phase v3.1 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) [111] to screen molecules by (a) allowing reorientation of the conformers to determine if they match the pharmacophore or not, (b) matching at least three out of the four sites of the structure-based common pharmacophore, (c) not considering any site as mandatory, (e) not preferring partial matches involving more sites and (f) using the excluded volumes from the structure-based common pharmacophore. The other options and parameter values used during the search were the default values. Those ligands with at least one hit in this Phase search were then docked to our hIKK-2 homology model with eHiTS v2009 using the same docking conditions described above. Then, the resulting ligand poses were again filtered with Phase through the structure-based common pharmacophore using the same filtering conditions as in the first Phase run with the exception that no reorientation of the poses was allowed during the search (*i.e.*, the *score in place* option was used) to find docking poses compatible with the pharmacophore.

Finally, in the last step of the VS protocol, the poses that were hits in this second pharmacophore screen were submitted to a shape and electrostatic potential comparison with 43 poses from 21 known hIKK-2 inhibitors that (a) were also obtained by docking with our hIKK-2 homology model and (b) also match the structure-based common pharmacophore without reorientation. This comparison was done with EON v2.0.1 (OpenEye Scientific Software, Inc., Santa Fe, New Mexico, USA; <http://www.eyesopen.com>), and only molecules with $ET_{pb} \geq 0.3$, $ET_{shape} \geq 0.5$ and $ET_{coulb} \geq 0.3$ were considered as predicted hIKK-2 inhibitors.

4.2 Clustering of molecules according to their chemical similarity

Clustering of the set of molecules formed by (a) the 13 NP hits present in extracts with known anti-inflammatory activity but whose effect on inflammation has not been described and (b) the 377 known synthetic inhibitors obtained from the BindingDB database (<http://www.bindingdb.org>) [112] was performed with Canvas v1.2.211 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>). MOLPRINT2D binary fingerprints were calculated for each molecule followed by a hierarchical clustering based on the Tanimoto similarity metric. Finally, the Kelly criterion was used to split the molecules into clusters.

Acknowledgements

This manuscript has been edited for English-language fluency by American Journal Experts. This study was supported by Grant Number AGL2008-00387/ALI from the Ministerio de Educación y Ciencia of the Spanish Government, Grant Number ASTF 61-2009 from the EMBO Short-Term Fellowship program and the ACCIÓ (TECCT10-1-0008) program (Generalitat de Catalunya). The authors wish to thank the Servei de Disseny de Fàrmacs (Drug Design Service) of the Catalonia Supercomputer Center (CESCA) for providing access to Schrödinger software and the Protein Modeling Facility of Lausanne University for help during the homology modeling of hIKK-2.

References

- [1] History of Medicine. U.S. National Library of Medicine. National Institutes of Health (<http://www.nlm.nih.gov/hmd>)
- [2] Rollinger, J. M.; Haupt, S.; Stuppner, H.; Langer, T. J. *Chem. Inf. Comput. Sci.* 44 (2004) 480-488.
- [3] Egner, M. S. a. U. *Med. Chem.* 6 (2007) 5-17.
- [4] David A Evans, J. B. H. a. S. D. J. *Sci. Food Agric.* 86 (2006) 2503-2509.
- [5] Mathieu, P.; Lemieux, I.; Després, J. P. *Clin. Pharmacol. Ther.* 87 (2010) 407-416.
- [6] Reilly, M.; Rader, D. *Circulation* 108 (2003) 1546-1551.
- [7] Hotamisligil, G. S.; Erbay, E. *Nat. Rev. Immunol.* 8 (2008) 923-934.
- [8] Roche, H.; Phillips, C.; Gibney, M. *Proc. Nutr. Soc.* 64 (2005) 371-377.
- [9] Wamil, M.; Seckl, J. R. *Drug Discov. Today* 12 (2007) 504-520.
- [10] Coish, P. D. G.; Wickens, P. L.; Lowinger, T. B. *Expert Opin. Ther. Pat.* 16 (2006) 1-12.
- [11] Israël, A. *Perspect. Biol.* 2 (2010) a000158.
- [12] Karin, M.; Yamamoto, Y.; Wang, Q. M. *Nat. Rev. Drug Discov.* 3 (2004) 17-26.
- [13] Schmid, J. A.; Birbach, A. *Cytokine Growth Factor Rev.* 19 (2008) 157-165.
- [14] Sala, E.; Guasch, L.; Vaqué, M.; Mateo-Sanz, J. M.; Blay, M.; Bladé, C.; Garcia-Vallvé, S.; Pujadas, G. *QSAR Comb. Sci.* 28 (2009) 678-695.
- [15] Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Model.* 45 (2005) 177-182.
- [16] Sala, E.; Guasch, L.; Iwaszkiewicz, J.; Mulero, M.; Salvadó, M. J.; Pinent, M.; Zoete, V.; Grosdidier, A.; Garcia-Vallvé, S.; Michielin, O.; Pujadas, G. *PLoS One* 2011, doi10.1371/journal.pone.0016903.
- [17] Llópiz, N.; Puiggròs, F.; Céspedes, E.; Arola, L.; Ardèvol, A.; Bladé, C.; Salvadó, M. J. *Agric. Food Chem.* 52 (2004) 1083-1087.
- [18] Del Bas, J. M.; Fernández-Larrea, J.; Blay, M.; Ardèvol, A.; Salvadó, M. J.; Arola, L.; Bladé, C. *FASEB J.* 19 (2005) 479-481.
- [19] Dohadwala, M. M.; Vita, J. A. *J. Nutr.* 139 (2009) 1788S-1793S.
- [20] Puiggròs, F.; Sala, E.; Vaque, M.; Ardevol, A.; Blay, M.; Fernandez-Larrea, J.; Arola, L.; Blade, C.; Pujadas, G.; Salvado, M. J. *J. Agric. Food Chem.* 57 (2009) 3934-3942.
- [21] Lee, F.-P.; Shih, C.-M.; Shen, H.-Y.; Chen, C.-M.; Chen, C.-M.; Ko, W.-C. *Eur. J. Pharmacol.* 635 (2010) 198-203.
- [22] Kim, H. K.; Cheon, B. S.; Kim, Y. H.; Kim, S. Y.; Kim, H. P. *Biochem. Pharmacol.* 58 (1999)

759-765.

[23] Wang, J.; Mazza, G. J. *Agric. Food Chem.* 50 (2002) 850-857.

[24] Huang, Y.; Tsang, S.-Y.; Yao, X.; Chen, Z.-Y. *Curr. Drug Targets Cardiovasc. Haematol. Disord.* 5 (2005) 177-184.

[25] Ugartondo, V.; Mitjans, M.; Lozano, C.; Torres, J. L.; Vinardell, M. P. *J. Agric. Food Chem.* 54 (2006) 6945-6950.

[26] Yu-Hsuan Lan, Y.-C. C., Fang-Rong Chang, Tsong-Long Hwang, Chih-Chaung Liaw, Yang-Chang Wu. *Helv. Chim. Acta* 88 (2005) 905-909.

[27] Dasse, O.; Parrott, J. A.; Putman, D.; Adam, J. World Intellectual Property Organization Patent WO2008100977, 2008.

[28] Lutomski, J.; Malek, B.; Rybacka, L. *Planta Med.* 27 (1975) 112-121.

[29] Castro, A. C.; Dang, L. C.; Soucy, F.; Grenier, L.; Mazdiyasni, H.; Hottelet, M.; Parent, L.; Pien, C.; Palombella, V.; Adams, J. *Bioorg. Med. Chem. Lett.* 13 (2003) 2419-2422.

[30] Newton, R.; Holden, N. S.; Catley, M. C.; Oyelusi, W.; Leigh, R.; Proud, D.; Barnes, P. J. *J. Pharmacol. Exp. Ther.* 321 (2007) 734-742.

[31] Wen, D.; Nong, Y.; Morgan, J. G.; Gangurde, P.; Bielecki, A.; Dasilva, J.; Keaveney, M.; Cheng, H.; Fraser, C.; Schopf, L.; Hepperle, M.; Harriman, G.; Jaffee, B. D.; Ocain, T. D.; Xu, Y. *J. Pharmacol. Exp. Ther.* 317 (2006) 989-1001.

[32] Xia, M.; Ling, W. H.; Ma, J.; Kitts, D. D.; Zawistowski, J. *J. Nutr.* 133 (2003) 744-751.

[33] Jin, M. H.; Yook, J.; Lee, E.; Lin, C. X.; Quan, Z.; Son, K. H.; Bae, K. H.; Kim, H. P.; Kang, S. S.; Chang, H. W. *Biol. Pharm. Bull.* 29 (2006) 884-888.

[34] Ohmoto, T.; Kazuo, K.; Yohko, S. *Chem. Pharm. Bull. (Tokyo)* 29 (1981) 390-395.

[35] Ruppelt, B. M.; Pereira, E. F.; Goncalves, L. C.; Pereira, N. A. *Mem. Inst. Oswaldo Cruz*, 86 Suppl 2 (1991) 203-205.

[36] Braz Filho, R.; Gottlieb, O. R. *Phytochemistry* 10 (1971) 2433-2450.

[37] Kubota, H.; Kojima-Yuasa, A.; Morii, R.; Huang, X.; Norikura, T.; Rho, S.-N.; Matsui-Yuasa, I. *Am. J. Chin. Med.* 37 (2009) 843-854.

[38] Osaki, N.; Koyano, T.; Kowithayakorn, T.; Hayashi, M.; Komiyama, K.; Ishibashi, M. *J. Nat. Prod.* 68 (2005) 447-449.

[39] Lin, L.-W.; Chen, H.-Y.; Wu, C.-R.; Liao, P.-M.; Lin, Y.-T.; Hsieh, M.-T.; Ching, H. *Biosci. Biotechnol. Biochem.* 72 (2008) 2377-2384.

[40] Lee, D.; Bhat, K. P.; Fong, H. H.; Farnsworth, N. R.; Pezzuto, J. M.; Kinghorn, A. D. *J. Nat. Prod.* 64 (2001) 1286-1293.

[41] Ahrens, J.; Demir, R.; Leuwer, M.; de la Roche, J.; Krampfl, K.; Foadi, N.; Karst, M.; Haeseler, G. *Pharmacology* 83 (2009) 217-222.

- [42] Appendino, G.; Fiebich, B. L.; Grassi, G.; Munoz Blanco, E. World Intellectual Property Organization Patent WO2009043836, 2009.
- [43] Crombei, L.; Crombei, W. M. L. *J. Chem. Soc., Perkin Trans. 1* (1982) 1455-1466.
- [44] Nsonde Ntandou, G. F.; Banzouzi, J. T.; Mbatchi, B.; Elion-Itou, R. D. G.; Etou-Ossibi, A. W.; Ramos, S.; Benoit-Vical, F.; Abena, A. A.; Ouamba, J. M. *J. Ethnopharmacol.* 127 (2010) 108-111.
- [45] Ahn, B. Z.; Zymalkowski, F. *Tetrahedron Lett.* 17 (1976) 821-824.
- [46] McGaw, L. J.; Rabe, T.; Sparg, S. G.; Jager, A. K.; Eloff, J. N.; van Staden, J. J. *Ethnopharmacol.* 75 (2001) 45-50.
- [47] Malan, E.; Swinny, E. *Phytochemistry* 34 (1993) 1139-1142.
- [48] Tadic, V. M.; Dobric, S.; Markovic, G. M.; Dordevic, S. M.; Arsic, I. A.; Menkovic, N. R.; Stevic, T. *J. Agric. Food Chem.* 56 (2008) 7700-7709.
- [49] Kokubun, T.; Harborne, J. B. *Phytochemistry* 40 (1995) 1649-1654.
- [50] Dhandapani, S.; Subramanian, V. R.; Rajagopal, S.; Namasivayam, N. *Pharmacol. Res.* 46 (2002) 251-255.
- [51] Takayanagi, T.; Ishikawa, T.; Kitajima, J. *Phytochemistry* 63 (2003) 479-484.
- [52] Chiou, W. F.; Ko, H. C.; Chen, C. F.; Chou, C. J. *J. Pharm. Pharmacol.* 54 (2002) 1399-1405.
- [53] Heefner, D. L.; Zepp, C. M.; Gao, Y.; Jones, S. W. World Intellectual Property Organization Patent WO1999031267, 1999.
- [54] Yu, L. L.; Ho, L. K.; Liao, J. F.; Chen, C. F. *Planta Med.* 63 (1997) 471-472.
- [55] Zhang, Y.-N.; Zhong, X.-G.; Zheng, Z.-P.; Hu, X.-D.; Zuo, J.-P.; Hu, L.-H. *Bioorg. Med. Chem.* 15 (2007) 988-996.
- [56] Schinella, G. R.; Tournier, H. A.; Prieto, J. M.; Mordujovich de Buschiazso, P.; Rios, J. L. *Life Sci.* 70 (2002) 1023-1033.
- [57] Musa, H. A. Z.; Emad, M. H.; Salim, S. S.; Wolfgang, V.; Klaus-Peter, Z. *J. Nat. Prod.* 61 (1998) 798-800.
- [58] Rahman, A.-u.; Nasim, S.; Baig, I.; Jalil, S.; Orhan, I.; Sener, B.; Choudhary, M. I. *J. Ethnopharmacol.* 86 (2003) 177-80.
- [59] Pailer; Franke. *Monatsh. Chem.* 104 (1973) 1394-1408.
- [60] Sae-wong, C.; Tansakul, P.; Tewtrakul, S. *J. Ethnopharmacol.* 124 (2009) 576-580.
- [61] Azuma, T.; Tanaka, Y.; Kikuzaki, H. *Phytochemistry* 69 (2008) 2743-2748.
- [62] Hajhashemi, V.; Ghannadi, A.; Sharif, B. *J. Ethnopharmacol.* 89 (2003) 67-71.
- [63] Banthorpe, D. V.; Bilyard, H. J.; Watson, D. G. *Phytochemistry* 24 (1985) 2677-2680.

- [64] Choi, E.-M.; Hwang, J.-K. *Fitoterapia* 76 (2005) 608-613.
- [65] Asano, N.; Yamashita, T.; Yasuda, K.; Ikeda, K.; Kizu, H.; Kameda, Y.; Kato, A.; Nash, R. J.; Lee, H. S.; Ryu, K. S. *J. Agric. Food Chem.* 49 (2001) 4208-4213.
- [66] Al-Ghamdi, M. S. *J. Ethnopharmacol.* 76 (2001) 45-48.
- [67] Atta-ur-Rahman; Malik, S.; Hasan, S. S.; Choudhary, M. I.; Ni, C.-Z.; Clardy, J. *Tetrahedron Lett.* 36 (1995) 1993-1996.
- [68] Laupattarakasem, P.; Houghton, P. J.; Hoult, J. R. S.; Itharat, A. *J. Ethnopharmacol.* 85 (2003) 207-215.
- [69] Hari Babu, T.; Manjulatha, K.; Suresh Kumar, G.; Hymavathi, A.; Tiwari, A. K.; Purohit, M.; Madhusudana Rao, J.; Suresh Babu, K. *Bioorg. Med. Chem. Lett.* 20 (2010) 117-120.
- [70] Hong, M. H.; Kim, J.-H.; Na, S. H.; Bae, H.; Shin, Y.-C.; Kim, S.-H.; Ko, S.-G. *Biosci. Biotechnol. Biochem.* 74 (2010) 1152-1156.
- [71] Okubo, T.; Nagai, F.; Seto, T.; Satoh, K.; Ushiyama, K.; Kano, I. *Biol. Pharm. Bull.* 23 (2000) 199-203.
- [72] Montanher, A. B.; Zucolotto, S. M.; Schenkel, E. P.; Frode, T. S. *J. Ethnopharmacol.* 109 (2007) 281-288.
- [73] Bremner, P.; Rivera, D.; Calzado, M. A.; Obon, C.; Inocencio, C.; Beckwith, C.; Fiebich, B. L.; Munoz, E.; Heinrich, M. *J. Ethnopharmacol.* 124 (2009) 295-305.
- [74] Pracharova, L.; Okenkova, K.; Lojek, A.; Ciz, M. *Life Sci.* 86 (2010) 518-523.
- [75] McKenzie, E.; Nettleship, L.; M, S. *Phytochemistry* 14 (1975) 273-275.
- [76] Berlin, J.; Ruegenhagen, C.; Greidziak, N.; Kuzovkina, I. N.; Witte, L.; Wray, V. *Phytochemistry* 33 (1993) 593-597.
- [77] Khan, M. A.; Khan, H.; Khan, S.; Mahmood, T.; Khan, P. M.; Jabar, A. *J. Enzyme Inhib. Med. Chem.* 24 (2009) 632-637.
- [78] Mohini, G.; Bagchi, A.; Sinha, S. C.; Sahai, M.; Ray, A. B. *J. Indian Chem. Soc.* 67 (1990) 597-599.
- [79] Jiao, W.-H.; Gao, H.; Li, C.-Y.; Zhao, F.; Jiang, R.-W.; Wang, Y.; Zhou, G.-X.; Yao, X.-S. *J. Nat. Prod.* 73 (2010) 167-171.
- [80] Ohmoto, T.; Koike, K.; Higuchi, T.; Ikeda, K. *Chem. Pharm. Bull. (Tokyo)* 33 (1985) 3356-3360.
- [81] Singh, N.; Kumar, S.; Singh, P.; Raj, H. G.; Prasad, A. K.; Parmar, V. S.; Ghosh, B. *Phytomedicine* 15 (2008) 284-291.
- [82] Desai, S. J.; Prabhu, B. R.; Mulchandani, N. B. 27 (1988) 1511-1515.
- [83] Mary, N. K.; Babu, B. H.; Padikkala, J. *Phytomedicine* 10 (2003) 474-482.

- [84] Dinda, B.; Das, S. K.; Hajra, A. K.; Bhattacharya, A.; De, K.; Chel, G.; Achari, B. *Indian J. Chem. B* 38 (1999) 577-582.
- [85] Srinivasan, K.; Muruganandan, S.; Lal, J.; Chandra, S.; Tandan, S. K.; Prakash, V. R. J. *Ethnopharmacol.* 78 (2001) 151-157.
- [86] Li, L.; Li, X.; Shi, C.; Deng, Z.; Fu, H.; Proksch, P.; Lin, W., Pongamone A-E. *Phytochemistry* 67 (2006) 1347-1352.
- [87] Deciga-Campos, M.; Palacios-Espinosa, J. F.; Reyes-Ramirez, A.; Mata, R. J. *Ethnopharmacol.* 114 (2007) 161-8.
- [88] Bordat, P.; Tarroux, R.; Charveron, M.; David, B. World Intellectual Property Organization Patent WO2006089881, 2006.
- [89] Estrada, S.; Rojas, A.; Mathison, Y.; Israel, A.; Mata, R. *Planta Med.* 65 (1999) 109-114.
- [90] Freire, S. M.; Torres, L. M.; Roque, N. F.; Souccar, C.; Lapa, A. J. *Mem. Inst. Oswaldo Cruz* 86 Suppl 2 (1991) 149-51.
- [91] Boerger, L.; Daum, W. US20100119582, 2010.
- [92] Phan, M. G.; Phan, T. S.; Matsunami, K.; Otsuka, H. *Chem. Pharm. Bull. (Tokyo)* 54 (2006) 546-549.
- [93] Kim, E. H.; Shim, B.; Kang, S.; Jeong, G.; Lee, J.-s.; Yu, Y.-B.; Chun, M. J. *Ethnopharmacol.* 126 (2009) 320-331.
- [94] Jia, Q.; Zhao, Y. World Intellectual Property Organization Patent WO2006099217, 2006.
- [95] Kim, J. Y.; Lim, H. J.; Ryu, J.-H. *Bioorg. Med. Chem. Lett.* 18 (2008) 1511-1514.
- [96] Mascolo, N.; Autore, G.; Capasso, F. J. *Ethnopharmacol.* 19 (1987) 81-84.
- [97] Rethy, B.; Kovacs, A.; Zupko, I.; Forgo, P.; Vasas, A.; Falkay, G.; Hohmann, J. *Planta Med.* 72 (2006) 767-770.
- [98] Malik, N.; Dhawan, V.; Bahl, A.; Kaul, D. *Platelets* 20 (2009) 183-190.
- [99] Bhuyan, R.; Saikia, C. N. *Bioresour Technol.* 96 (2005) 363-372.
- [100] Chen, J.; Uto, T.; Tanigawa, S.; Kumamoto, T.; Fujii, M.; Hou, D.-X. *Nutr. Cancer* 60 Suppl 1 (2008) 43-50.
- [101] Katsube, N.; Iwashita, K.; Tsushida, T.; Yamaki, K.; Kobori, M. J. *Agric. Food Chem.* 51 (2003) 68-75.
- [102] Jacobo-Herrera, N. J.; Vartiainen, N.; Bremner, P.; Gibbons, S.; Koistinaho, J.; Heinrich, M. *Phytother. Res.* 20 (2006) 917-919.
- [103] Torssell, K.; Wahlberg, K. *Acta Chem. Scand.* 21 (1967) 53-62.
- [104] Kosar, M.; Kupeli, E.; Malyer, H.; Uylaser, V.; Turkben, C.; Baser, K. H. C. *J. Agric. Food Chem.* 55 (2007) 4596-4603.

- [105] Torres, J. L.; Bobet, R. J. *Agric. Food Chem.* 49 (2001) 4627-4634.
- [106] Chen, J.-J.; Wang, T.-Y.; Hwang, T.-L. *J. Nat. Prod.* 71 (2008) 212-217.
- [107] Lagorce, D.; Sperandio, O.; Galons, H.; Miteva, M. A.; Villoutreix, B. O. *BMC Bioinformatics* 2008, (9), doi:10.1186/1471-2105-9-396.
- [108] Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* 46 (2001) 3-26.
- [109] Zsoldos, Z.; Reid, D.; Simon, A.; Sadjad, B. S.; Johnson, A. P. *Curr. Protein Pept. Sci.* 7 (2006) 421-435.
- [110] Wolber, G.; Langer, T. J. *Chem. Inf. Model.* 45 (2005) 160-169.
- [111] Dixon, S. L.; Smondyrev, A. M.; Knoll, E. H.; Rao, S. N.; Shaw, D. E.; Friesner, R. A. J. *Comput. Aided Mol. Des.* 20 (2006) 647-671.
- [112] Liu, T. Q.; Lin, Y. M.; Wen, X.; Jorissen, R. N.; Gilson, M. K. *Nucleic Acids Res.* 35 (2007) D198-D201.

Table 1. Extracts with described anti-inflammatory activity that contain at least one molecule predicted to inhibit hIKK-2 by our virtual screening protocol.

| Source | Molecule | References |
|---------------------------------|---|--------------------------|
| <i>Ailanthus altissima</i> | 38940-60-2 | [17] - [18] |
| <i>Apuleia leiocarpa</i> | 572-32-7 | [19] [20] [21] |
| <i>Blumea balsamifera</i> | 572-32-7 | [22] [20] [23] |
| <i>Broussonetia papyrifera</i> | broussonin F | [24] - [25] |
| <i>Cannabis sativa</i> | 64052-90-0 | [26] [27] [28] |
| <i>Cassia Siamea</i> | 60352-12-7 | [29] - [30] |
| <i>Combretum apiculatum</i> | 39499-93-9 | [31] - [32] |
| <i>Crataegus monogyna</i> | 93973-22-9 γ -cotonefuran | [33] - [34] |
| <i>Cuminum cyminum</i> | methyl β -D-apiofuranosyl-(1->6)- β -D-glucopyranoside | [35] - [36] |
| <i>Evodia rutaecarpa</i> | 20315-68-8 | [37] [38] [39] |
| <i>Fissistigma oldhamii</i> | oldhamactam aristolactam B III | [40] [40] [40] |
| <i>Inula viscosa</i> | 572-32-7 | [41] [20] [42] |
| <i>Iris germanica</i> | 39012-01-6 | [43] [44] [45] |
| <i>Kaempferia parviflora</i> | 572-32-7 | [46] [20] [47] |
| <i>Lavandula angustifolia</i> | 643-84-5 | [48] [49] [50] |
| <i>Morus alba</i> | 1,4-dideoxy-1,4-imino-(2-O- β -D-glucopyranosyl)-D-arabinitol | [51] - [52] |
| <i>Nigella sativa</i> | nigellidine | [53] - [54] |
| <i>Oroxylum indicum</i> | 491-67-8 | [55] [56] [57] |
| <i>Paeonia suffruticosa</i> | 1948-33-0 | [58] [59] [60] |
| <i>Passiflora edulis</i> | 487-03-6 | [61] [62] [63] |
| <i>Peganum harmala</i> | 50-67-9 83789-01-9 | [64] [65] [66] - [67] |
| <i>Physalis minima</i> | 572-32-7 | [68] [20] [69] |
| <i>Picrasma quassioides</i> | 100234-63-7 | [70] - [71] |
| <i>Piper longum</i> | 112501-42-5 | [72] [73] [74] |
| <i>Plumbago indica</i> | 572-32-7 | [75] [20] [76] |
| <i>Pongamia pinnata</i> | 5-hydroxyfuran[7,6:4",5"]flavone | [77] - [78] |
| <i>Scaphyglottis livida</i> | 118169-17-8 | [79] [80] [81] |
| <i>Scoparia dulcis</i> | 51-43-4 | [82] [83] [84] |
| <i>Scutellaria baicalensis</i> | 491-67-8 | [85] [56] [86] |
| <i>Siegesbeckia glabrescens</i> | 572-32-7 | [87] [20] [87] |
| <i>Tamus communis</i> | 108909-02-0 118169-17-8 | [88] [80] [89] |
| <i>Terminalia arjuna</i> | 491-67-8 | [90] [56] [91] |
| <i>Vaccinium myrtillus</i> | 643-84-5 | [92] [49] [93] |
| <i>Valeriana officinalis</i> | 14140-13-7 | [94] - [95] |
| <i>Vitis vinifera</i> | 4 β -(2-aminoethylthio)epicatechin | [96] [97] [98] |
| <i>Zanthoxylum avicennae</i> | 2255-50-7 | [99] [99] [99] |

Our VS workflow predicted that 274 molecules from our in-house NP database (isolated from 453 different natural sources) are potential hIKK-2 inhibitors. A bibliographic search in PubMed revealed that anti-inflammatory activity has been described for 36 extracts from among these 453 different natural sources. These 36 extracts contain 30 NP hits from our VS, and a search in SciFinder revealed that anti-inflammatory activity has been described for 17 of them. Table 1 shows these 36 extracts and the name of the VS hits that have been purified from them (identified, when available, by their corresponding CAS number). The bibliographic references for each extract are split into three columns in which **(a)** the first column reports papers that describe the anti-inflammatory activity of the corresponding extract; **(b)** the second column reports papers that describe the anti-inflammatory activity of the corresponding molecule (when available), and **(c)** the third column reports papers that describe the purification of each molecule from the corresponding extract. The 2D structures of these virtual screening hits are described in Table S1 of the supporting information.

Figure 1. The VS workflow used in the present work.

| Workflow steps | Validation | | | Virtual Screening |
|---|------------|--------|-----|-------------------|
| | Actives | Decoys | EF | In-house database |
| Starting Database | 62 | 10036 | - | 29779 |
| ADME/Toxicity analysis FaF-Drug2 | 58 | 6816 | 1.4 | 11656 |
| Structure-based Pharmacophore Phase v3.1 Rigid Docking eHiTs v2009 Structure-based Pharmacophore Phase v3.1 | 26 | 462 | 6.3 | 4376 |
| Similarity & Electrostatic analysis EON v2.0.1 | 11 | 24 | 8.6 | 274 |

The data beside each VS step show the number of molecules that survived it. The *ACTIVES* and *DECOYS* columns refer to known hIKK-2 and kinase decoys used during VS validation, respectively. The *In-house* column refers to data obtained with our in-house database of NPs. Enrichment factors were calculated during the validation of each step of the VS protocol as the quotient between the fraction of actives in the sample that survived the VS step and the fraction of actives in the sample before the VS step.

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

SUPPORTING INFORMATION

Manuscript 2

Identification of IKK-2 Inhibitors of Natural Origin (part II): In silico prediction of IKK-2 inhibitors in natural extracts with known anti-inflammatory activity.

Esther Sala, Laura Guasch, Justyna Iwaszkiewicz, Miquel Mulero, Maria-Josepa Salvadó, Cinta Bladé, Vincent Zoete, Aurélien Grosdidier, Santiago Garcia-Vallvé, Olivier Michielin and Gerard Pujadas

Submitted to European Journal of Medicinal Chemistry
(3 February 2011)

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

SUPPORTING INFORMATION

Identification of Human IKK-2 Inhibitors of Natural Origin (Part II): *In Silico* Prediction of IKK-2 Inhibitors in Natural Extracts with Known Anti-Inflammatory Activity.

Esther Sala^a, Laura Guasch^a, Justyna Iwaszkiewicz^b, Miquel Mulero^a, Maria-Josepa Salvadó^a, Cinta Bladé^a, Meritxell Ceballos^c, Cristina Valls^a, Vincent Zoete^b, Aurélien Grosdidier^b, Santiago Garcia-Vallvé^{a,c}, Olivier Michielin^b, Gerard Pujadas^{a,c*}

^a Grup de Recerca en Nutrigenòmica, Departament de Bioquímica i Biotecnologia, Universitat Rovira i Virgili, Campus de Sescelades, C/ Marcel·lí Domingo s/n, 43007 Tarragona, Catalonia, Spain

^b Molecular Modeling Group. Swiss Institute of Bioinformatics, Quartier UNIL-Sorge. Bâtiment Génopode. CH-1015 Lausanne, Switzerland

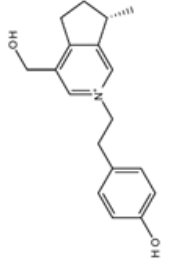
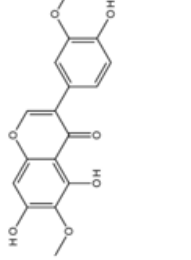
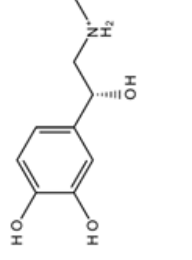
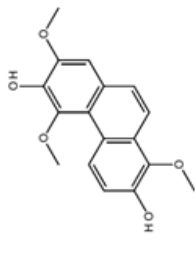
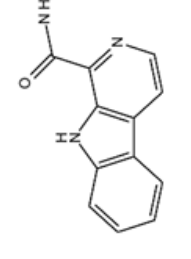
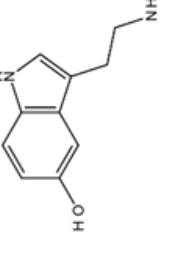
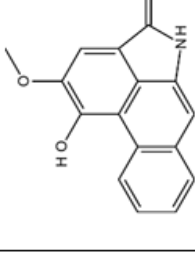
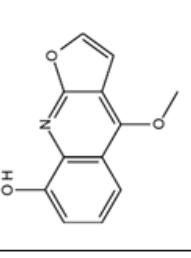
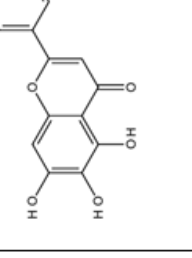
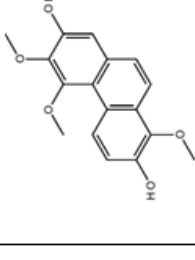
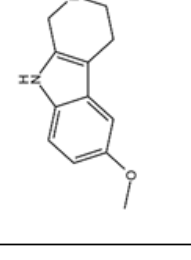
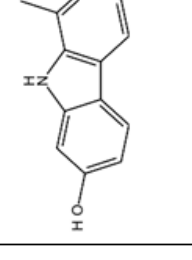
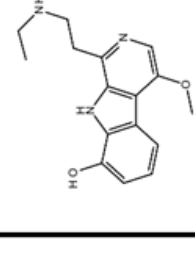
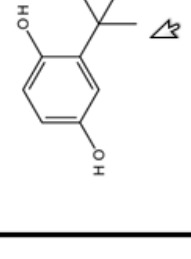
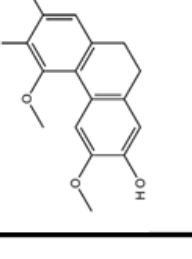
^c Centre Tecnològic de Nutrició i Salut (CTNS), TECNIO, CEICS, Camí de Valls, 81-87, 43204 Reus, Catalonia, Spain

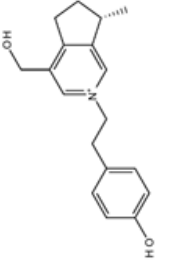
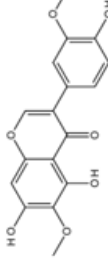
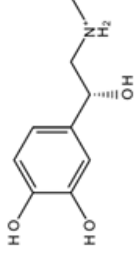
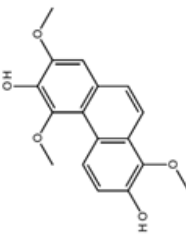
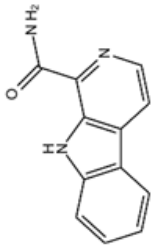
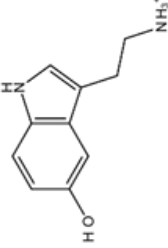
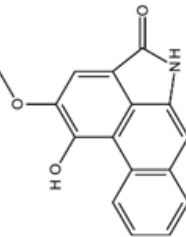
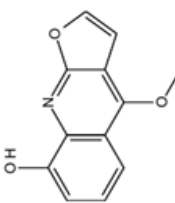
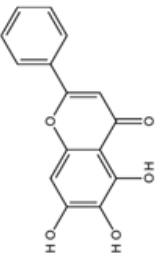
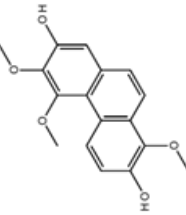
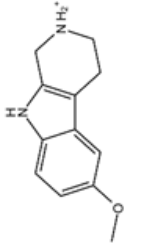
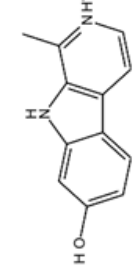
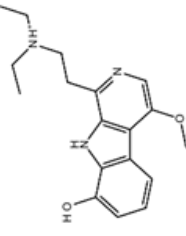
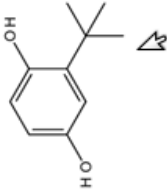
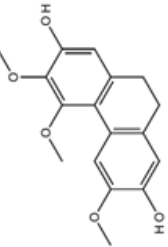
*e-mail: gerard.pujadas@urv.cat; Tel: +34977 559565; Fax: +34977558232

CONTENTS

- 2D structures of the virtual screening hits present in the 36 extracts with described anti-inflammatory activity **S1**

Table S1. 2D structures of the virtual screening hits present in the 36 extracts with known anti-inflammatory activity. Molecules are identified and sorted by their CAS number (when available)

| | |
|--|-------------|
|  | 14140-13-7 |
|  | 39012-01-6 |
|  | 5143-4 |
|  | 118169-17-8 |
|  | 38940-60-2 |
|  | 50-67-9 |
|  | 112501-42-5 |
|  | 2255-50-7 |
|  | 491-67-8 |
|  | 108909-02-0 |
|  | 20315-68-8 |
|  | 487-03-6 |
|  | 100234-63-7 |
|  | 1948-33-0 |
|  | 39499-93-9 |

| | | | | | |
|---|--------------------|---|-------------------|--|-------------------|
|  | 14140-13-7 |  | 39012-01-6 |  | 5143-4 |
|  | 118169-17-8 |  | 38940-60-2 |  | 50-67-9 |
|  | 112501-42-5 |  | 2255-50-7 |  | 491-67-8 |
|  | 108909-02-0 |  | 20315-68-8 |  | 487-03-6 |
|  | 100234-63-7 |  | 1948-33-0 |  | 39499-93-9 |

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

Manuscript 3

3D-QSAR study of Pyridine Derivates as IKK-2

Esther Sala, Laura Guasch, Montserrat Vaqué, Josep Maria Mateo-Sanz,
Mayte Blay, Cinta Bladé, Santiago Garcia-Vallvé and Gerard Pujadas

QSAR & Combinatorial Science; 28 (6-7), 678–695, July 2009

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

3D-QSAR Study of Pyridine Derivates as IKK-2 Inhibitors

Esther Sala,^a Laura Guasch,^a Montserrat Vaqué,^a Josep Maria Mateo-Sanz,^b Mayte Blay,^a Cinta Bladé,^a Santiago Garcia-Vallvé,^a Gerard Pujadas^{a*}

^a Grup de Recerca en Nutrigenòmica, Departament de Bioquímica i Biotecnologia, Universitat Rovira i Virgili, Campus de Sescelades, C/ Marcel·lí Domingo s/n, 43007 Tarragona, Catalonia, Spain

^b Grup Crises, Departament d'Enginyeria Química, Universitat Rovira i Virgili, Campus de Sescelades, Av. Països Catalans 26, 43007 Tarragona, Catalonia, Spain

*e-mail: gerard.pujadas@urv.cat; Tel: +34977559565; Fax: +34977558232

Keywords: Phase, Pharmacophore, NF- κ B pathway, Inflammation, Structure-activity relationships

Received: December 3, 2008; Accepted: February 11, 2009

DOI: 10.1002/qsar.200860167

Abstract

I κ B kinase β (IKK-2) is a serine-threonine protein kinase that is critically involved in activating the transcription factor NF- κ B in response to various inflammatory stimuli (hence, the interest in synthesizing molecules that can inhibit IKK-2). Some pyridine derivates have been shown to be among the most potent IKK-2 inhibitors known at present. This study reports how pyridine derivates have been used to develop a 3D-QSAR model for rationalizing results from previous structure-activity relationship studies on IKK-2 inhibition by these molecules; and predicting the IC_{50} of other pyridine-derivate IKK-2 inhibitors whose activity is unknown. Therefore, 3D-QSAR models were built with PHASE and a training set of 21 molecules with IC_{50} values that span five orders of magnitude (ranging from 0.004 to more than 20 μ M) and in which each order of magnitude had from two to seven molecules. The resulting models were validated with a test set made up of 16 molecules of known IC_{50} . Only 3 out of 1245 3D-QSAR models had statistical parameter values in the following ranges: a) R^2 , Q^2 and r -Pearson value ≥ 0.8000 ; b) RMSE/ SD in the [0.90–1.05] interval; c) $Q^2/R^2 \approx 1$; and d) $P < 0.05$ (statistical significance). Moreover, a second test set (consisting of 63 molecules whose activity has been reported in the form of the IC_{50} interval to which it belongs) was used for selecting which of the three remaining 3D-QSAR models makes the best predictions. The best 3D-QSAR model obtained is associated with a pharmacophore that contains all the characteristics that previous structural-activity studies have reported to be essential if IKK-2 is to be inhibited by these ligands. Therefore, it is concluded that this 3D-QSAR model can be used to reliably predict how other pyridine derivates inhibit IKK-2.

1 Introduction

The NF- κ B pathway is important in regulating the expression of cellular genes that are involved in controlling the immune and inflammatory response [1]. The activation of NF- κ B, then, induces the expression of more than 150 genes such as cytokines (e.g. TNF- β , IL-1, IL-6), chemokines (e.g. IL-8, MCP-1), cell adhesion molecules (e.g. ICAM-1, VCAM-1) and proteases [1]. NF- κ B exists in the cytoplasm in an inactive form associated with regulatory proteins called inhibitors of κ B or I κ B. In order to activate NF- κ B, I κ B has to be phosphorylated by IKK. This kinase is a complex that consists of: a) two catalytic subunits (i.e. IKK- α or IKK-1 and IKK- β or IKK-2); b) one regulatory subunit called IKK- γ or NEMO; c) NIK; and d) various other known and unknown proteins [2]. Although IKK-1

and IKK-2 are highly homologous and contain similar structural domains, it is IKK-2 that plays the major role in activating NF- κ B via the canonical pathway [3]. Thus, the importance of the IKK complex in regulating the NF- κ B pathway has led to pharmaceutical companies taking considerable interest in developing synthetic IKK-2 inhibitors from several chemical families (e.g. pyrimidines, quinazolines, thiophenecarboxamides, pyridines, benzimidazoles, indoles, and carboline derivatives [3]) that can be used, for instance, in patients affected by chronic-based inflammation diseases.

A large number of pyridine derivatives that are IKK-2 inhibitors have been synthesized by the Bayer Research Center in Kyoto [4–8]. These compounds are potent and selective IKK-2 inhibitors that are anti-inflammatory both in vitro and in vivo. In fact, it has been claimed that they

are a novel class of antiasthma and anti-inflammatory drugs [9]. 3D-QSAR studies for IKK- β inhibitors are only available for thienopyridine [10] and thiophenecarboxamide derivatives [11]. Therefore, the present study aims to contribute to knowledge of the structural-activity relationships of IKK-2 inhibitors by using Bayer's pyridine derivatives [4-8] for developing a 3D-QSAR model that: a) correlates the experimental activity with the structures of the inhibitors used to develop the model; and b) is able to predict the IC_{50} of a large set of other pyridine derivatives not used during the development of the model.

2 Materials and Methods

2.1 IKK-2 Inhibitors Used in the Present Study

We are working with the 252 pyridine derivatives in Figures A to D in the Appendix (henceforth, figures in the Appendix will be distinguished from figures in the other sections by using letters instead of numbers) [4-8]. According to the literature, these derivatives act as IKK-2 inhibitors by competing with ATP for the binding site [9]. The IC_{50} values (i.e. the concentration needed to inhibit the enzyme activity by 50%) of 37 of these 252 molecules were obtained in the same experimental conditions by the same scientific team (see Figures A and B, Appendix) [5-8]. Their chemical structure is also unequivocally known (i.e. either there are no chiral atoms in their structure or the chirality of the molecules is known), their IC_{50} values span five orders of magnitude (from 0.004 to more than 20 μ M) and each order of magnitude is represented by several compounds (see Figures A and B, Appendix). These 37 molecules have been assigned to either the training set from which the 3D-QSAR models are obtained (21 molecules; see Fig. A, Appendix) or the test set, which is used to validate these models (16 molecules; see Fig. B, Appendix). The activity of 63 other pyridine derivatives of unambiguous structure was obtained under the same conditions as the 37 molecules in the training and test sets but it is reported in the literature as the IC_{50} interval to which the corresponding IC_{50} value belongs (see Fig. C, Appendix) [4]. Therefore, the 63 derivatives were used as an external validation set to help to choose the best predictive 3D-QSAR model. The remaining 152 pyridine derivatives (see Fig. D, Appendix) have ambiguities in their structure (i.e. they have at least one chiral atom) and their measured IC_{50} or their IC_{50} range corresponds to an undetermined mixture of the stereoisomers [4, 6-8]. Consequently, they were used neither to derive the 3D-QSAR models nor to validate them but their individual activity was predicted with the best 3D-QSAR model that we obtained.

2.2 Generation of the 3D Structures of the Ligands and Further Set-up for Using Them with PHASE

The 3D-structure of all ligands in Figures A to D (Appendix) was built and minimized with ChemDraw Ultra v11.0 [12] (CambridgeSoft Corporation, Cambridge, MA, USA; <http://www.cambridgesoft.com/>). These 3D structures were then incorporated into the LigPrep module of the Schrödinger 2007 suite (<http://www.schrodinger.com>) and improved by *cleaning*. The cleaning process was carried out with default parameter values except for *Generate possible states at target pH*, which was set to 7.6 (i.e. the pH used to measure IC_{50} values [4-8]) and the *Stereoisomers* checkbox, in which the *retain specified chiralities* option was selected.

2.3 Generation of Conformers for Training and Test Set Ligands

Once the ligands were ready for use with the PHASE [13] module of the Schrödinger 2007 suite, a conformational search was run in order to generate sets of representative conformers for each ligand in the training and test sets [this is done because: a) the 3D structure of IKK-2 is unknown so protein-ligand docking cannot be used to find the bioactive conformation of each ligand; and b) this conformer generation is expected to find the bioactive conformation for each ligand]. The search method that we used was the ConfGen conformational utility from MacroModel through the Ligand Torsion Search interface. Briefly, ConfGen carefully examines the structure of each ligand to discover where local minima are expected as a function of rotations about rotatable bonds. Then, it systematically generates the conformations that arise from various combinations of these local minima and provides a broad and fairly uniform coverage of the available conformational space. The non-default parameter values used during this conformational search were: a) Force field: MMFFs; b) Dielectric constant: 1.0; and c) non-selection of the suppress hydrogen bonding electrostatics checkbox in order to allow intramolecular hydrogen bonds during the conformational search (it has been reported that the most active IKK-2 inhibitors used in the present study have an internal hydrogen bond [4-8] so this bond should be possible in our conformations).

2.4 Generation of 3D-QSAR models

The 3D-QSAR study was carried out with the PHASE v2.5 module [13] built within the Schrödinger 2007 suite. Initially, the pIC_{50} values were obtained for each ligand in the training and test sets (see Figures A and B, Appendix) by calculating $-\log IC_{50} + 6$ (where the IC_{50} units are μ M) because PHASE expects that the higher the activity of the ligand is, the higher the value that quantifies it. After that, pharmacophore sites (i.e. a set of points in the 3D space,

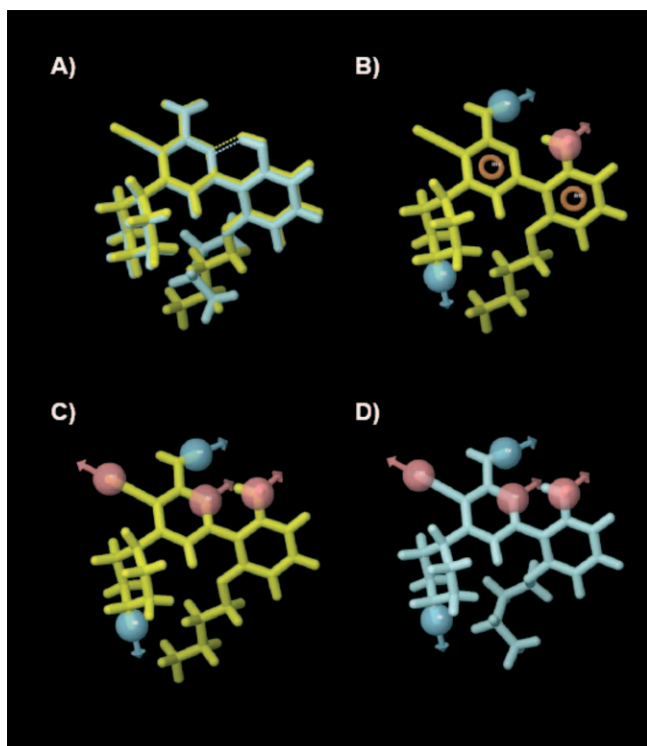


Figure 1. The conformations of the reference ligand and the associated sites (i.e. **A** for hydrogen-bond acceptors, **D** for hydrogen-bond donors and **R** for aromatic ring) that have been used to build the three hypotheses with best statistical parameters are shown (see Table 1). The reference ligand is **4e** [8] for the three hypotheses. Hypotheses ADDRR.383 and AADDD.2750 share the same conformation of **4e** (in yellow) whereas the conformation used by hypothesis AADDD.1950 is shown in turquoise. The orientation is the same in all panels so that they can be easily compared. Figure 1A compares the two conformations of **4e** used by the three hypotheses and also shows the dashed lines that correspond to the intramolecular hydrogen bond that has been described as crucial for the activity of this kind of ligand [6]. Figures 1B, 1C, and 1D show the conformation of the reference ligand and sites that have been used to build hypotheses ADDRR.383, AADDD.2750, and AADDD.1950, respectively.

which are coincident with chemical features that can facilitate non-covalent binding between the ligand and its target receptor) were obtained for the different conformations of these ligands by using the default site definitions built within PHASE. Thus, the program classifies the chemical features as hydrogen bond acceptors (labeled **A**), hydrogen bond donors (labeled **D**), hydrophobic groups (labeled **H**), negatively charged groups (labeled **N**), positively charged groups (labeled **P**) and aromatic rings (labeled **R**).

PHASE then uses the most active molecules in the training set to build the pharmacophores (i.e. the program assumes that the most active ligands have the strongest binding interaction with IKK-2 or contain the highest

number of pharmacophore features that are involved in the binding to the protein target). The user, therefore, needs to set a pIC_{50} threshold that helps to select which IKK-2 inhibitors have to be used during pharmacophore development. Thus, we considered to be all those IKK-2 inhibitors with a pIC_{50} higher than 7.620 (7 out of 37 molecules). Only five of these seven molecules belong to the training set, so only these were used to develop the pharmacophore [the other two molecules belong to the test set and were kept for validation purposes (see Figures A and B, Appendix)]. At this point it is worth mentioning that the active molecules we selected were diverse enough in structure and activity to ensure the reliability of the pharmacophore models that were subsequently obtained with PHASE.

The next step in generating the pharmacophore was the search for the so-called *common pharmacophores* or *hypotheses*. In this step, pharmacophores from the different conformations of all five active ligands in the training set were examined, and those conformations with identical sets of five pharmacophore sites that were similarly spatially arranged were grouped together. Once the common pharmacophores had been identified, they were scored. Initially, this scoring was based on how well the active ligands superimpose when they were aligned on the chemical features associated with each hypothesis. The quality of each alignment was measured as the result of three different contributions: 1) the alignment score (i.e. the RMSD of the site-point positions); 2) the vector score [i.e. the average cosine of the angles formed by the corresponding pairs of vector features (acceptors, donors, and aromatic rings) in the aligned structures]; and 3) a volume score that is based on the overlap of the van der Waals models of the nonhydrogen atoms in each pair of structures. The weights that measure the contribution of these three scores to the final hypothesis score were by default. Moreover, we use the inactive compounds to perform an additional scoring to penalize those hypotheses that could not distinguish between active and inactive compounds. During this second scoring, we considered those molecules in the training set whose pIC_{50} was lower than 4.699 as *inactive* IKK-2 inhibitors (see Fig. A, Appendix). After this step, the top 10% of the scored hypotheses (i.e. 1245) were kept to develop the corresponding 3D-QSAR models and to understand how the structural differences between IKK-2 inhibitors affect the activity of the enzyme. We generated 3D-QSAR models by using the experimental pIC_{50} from all the IKK-2 inhibitors in the training set (see Fig. A, Appendix) and aligning them to the reference ligand associated with each of the 1245 hypotheses. At this point, it is worth pointing out that the training set includes the five active molecules that were used to build the common pharmacophores. All the parameters used during the building of the 3D-QSAR models had default values except for the maximum number of PLS factors (i.e. 3), the 3D-QSAR model type (i.e. atom-based) and the percent

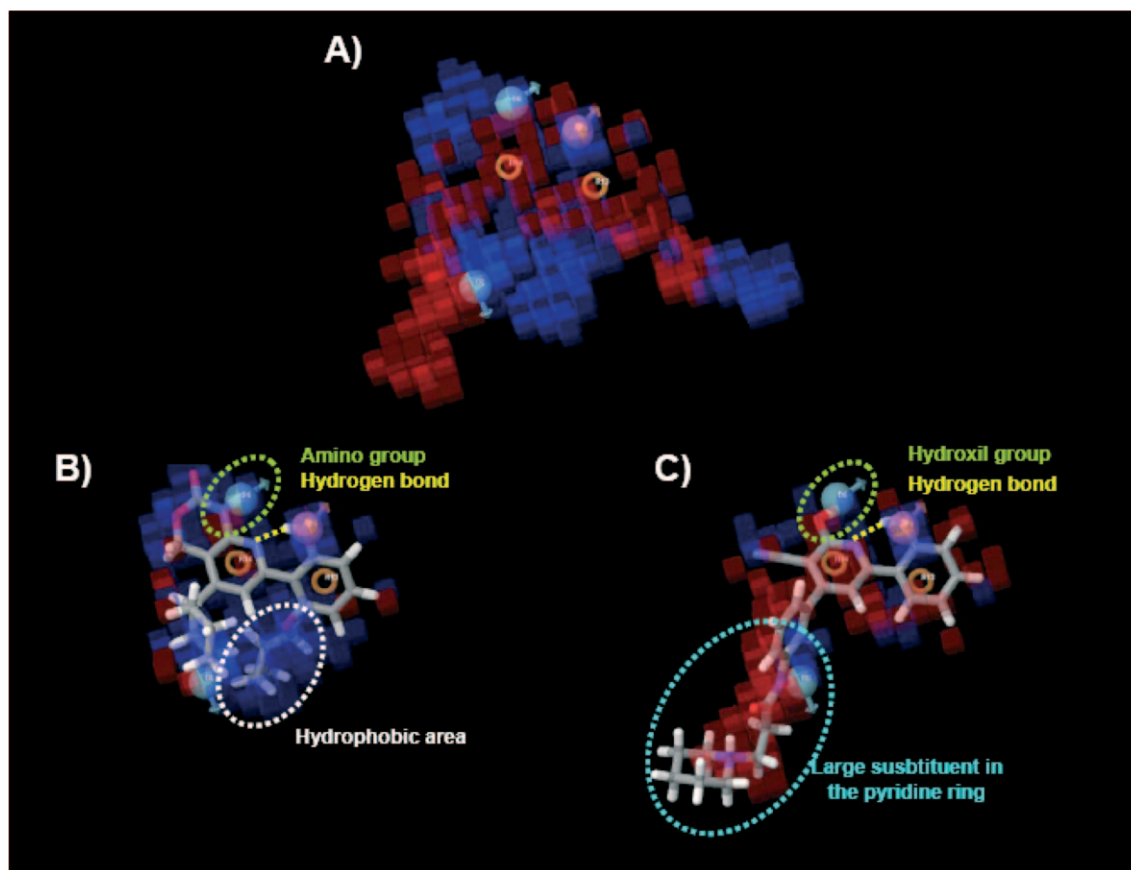


Figure 2. The ADDRR.383 model in the context of the most active ligand and one of the least active ones. The orientation is the same in all panels so that they can be easily compared. The complete ADDRR.383 model is shown in panel A. Panels B and C show only the cubic volume elements of ADDRR.383 that are occupied either by the most active compound (i.e. **S_patent** [5, 9]; panel B) or by one of the least active compounds (i.e. **41** [6]; panel C). The blue cubes in ADDRR.383 indicate regions that are favorable for activity (for instance, the intramolecular hydrogen bond, the amino donor group and the hydrophobic area from **S_patent**). The red cubes from ADDRR.383 indicate regions that are unfavorable for activity (for instance, the large substituent in the pyridine ring and the hydroxyl group from **41**). Finally, areas from ADDRR.383 without blue and red regions suggest that they are not important for the activity. The thresholds used for viewing the blue and red cubes are 4.000×10^{-3} and -4.000×10^{-3} , respectively.

of ligands assigned to the training and test sets (i.e. 57 and 43%; respectively).

After all the 1245 3D-QSAR models had been built, they were used to predict the pIC_{50} for all the IKK-2 inhibitors from the training and the test sets. Then, from the comparison between the experimental and predicted pIC_{50} , we derived statistics to describe how good each 3D-QSAR model is for predicting the pIC_{50} for pyridine derivatives. The statistical parameters that were used to evaluate the training set prediction were: a) R^2 (i.e. coefficient of determination); b) SD (standard deviation of regression); c) F statistic (i.e. overall significance of model); and d) statistical significance (i.e. the probability that the correlation could occur by chance). The parameters used to evaluate the test set predictions were: a) Q^2 ; b) root-mean-squared error; and c) Pearson r value (i.e. Pearson correlation coefficient). Finally, the 3D-QSAR models with the best predictive power were considered to be those that simultaneously met all the following criteria: a) R^2 , Q^2 and Pearson r

value ≥ 0.8000 ; b) $RMSE/SD$ in the interval [0.90–1.05]; c) $Q^2/R^2 \approx 1$ (because the 3D-QSAR needs to predict equally well molecules in the training and the test sets); and d) $P < 0.05$ (statistical significance). The selection was made with the help of a Filemaker database (<http://www.filemaker.com/>) to which all the statistical results from the 1245 hypotheses had been exported. With this criterion and the help of the database, we found that 3 of 1245 hypotheses (see Table 1) had enough predictive power to be used to predict the IC_{50} of pyridine derivatives other than the ones in the training and test sets.

2.5 Find Matches to Hypothesis

We created two different 3D databases by using the Manage 3D Database application of PHASE. The first contained 63 pyridine derivatives of unambiguous structure and known activity range (see Fig. C, Appendix) [4]. This database was used to help us choose which of the remain-

ing three 3D-QSAR models had the best predictive power. The second database was built with 152 pyridine derivatives of unknown stereochemistry (i.e. they have at least one chiral carbon) (see Fig. D, Appendix) and, therefore, for these molecules, the measured IC_{50} or their IC_{50} range corresponds to an undetermined mixture of the different stereoisomers [4–8]. Consequently, their individual activity was predicted with the best 3D-QSAR model that we obtained.

To generate the conformation of all the ligands in these two databases, we used the same conformational search method and parameter values as were previously used to generate the conformers for the training and test set ligands. Then, these conformations were imported into the two corresponding databases and sites were created for each conformer. Finally, the activity for these pyridine derivatives was predicted with the *Search for Matches* application in the *Build QSAR Model step* from Develop Pharmacophore Model wizard of PHASE.

3 Results and Discussion

3.1 3D-QSAR Model Selection

Three of the 1245 3D-QSAR models built by PHASE can: a) describe the structural-activity relationships from the training set molecules (see Fig. A, Appendix); and b) make good predictions for the activity of other pyridine derivatives not used to build the models (see Fig. B, Appendix). Table 1 shows the most relevant statistical parameters obtained when the experimental pIC_{50} for the training and test set molecules are compared with the pIC_{50} values predicted by these three 3D-QSAR models. We made some additional studies and analyses to choose which of these three hypotheses: a) has the best predictive power; and b) best describes the most important characteristics of the pyridine IKK inhibitors that have been described in previous structural-activity studies [4–8].

Table 1 and Figure 1 show that all three hypotheses have been built with the same reference ligand (i.e. **4e** [8]; see Fig. A, Appendix) but with two different conformations. Hypotheses ADDRR.383 and AAADD.2750 share the same conformation of **4e** (in yellow in Fig. 1) whereas hypothesis AAADD.1950 uses a slightly different conformation (in turquoise in Figure 1). Moreover the kind, number and location of the sites (i.e. the chemical features) used to define hypotheses AAADD.2750 and AAADD.1950 are the same (i.e. three hydrogen-bond acceptors and two donors) but they differ in the conformation of **4e** that they use (see Figs. 1A, 1C and 1D). On the other hand, ADDRR.383 and AAADD.2750 share the same conformation of **4e** but use different kinds of sites to align the training set molecules during the building of the 3D-QSAR model (see Figs. 1A, B and C).

The intramolecular hydrogen bond between the nitrogen of the pyridine ring and the hydroxyl group in the phenolic moiety is important for the activity of this kind of inhibitors [6]. The three models that we have selected from the initial 1245 sample have this bond (see Fig. 1A). At this point, it should be pointed out that PHASE generates the 3D-QSAR models by superimposing the structures of the molecules from the training set onto the reference ligand of the hypothesis and using its sites to lead the superimposition. Therefore, all pharmacophores that have to be used to derive 3D-QSAR models should contain the two aromatic rings from **4e** because: a) they are shared by all the molecules in the training set (see Fig. A, Appendix); and b) once the position of the rings has been fixed, the intramolecular hydrogen bond would be formed and this guarantees that the functional groups hanging from these two rings remain in a position that is suitable for interacting with IKK-2. Of the three hypotheses selected, ADDRR.383 is the only one to satisfy all these conditions.

Table 2 shows how well the three selected hypotheses can predict the activities of the molecules in the training and test sets. Thus, if pIC_{50} values are classified on an activity scale [i.e. ‘++++’ for highly active molecules ($pIC_{50} \geq 8.0$); ‘+++’ for moderately active molecules ($8.0 > pIC_{50} \geq 7.0$); ‘++’ for moderately inactive molecules ($7.0 > pIC_{50} \geq 5.0$); and ‘+’ for inactive compounds ($pIC_{50} < 5.0$)], then the correct category is predicted in 14 out of 21 molecules (67%) for ADDRR.383, 16 out of 21 molecules (76%) for AAADD.2750 and 17 out of 21 molecules (81%) for AAADD.1950. If a similar analysis is made with the molecules in the test set, all hypotheses correctly predicted the activity scale in 11 out of 16 cases (69%) [although all the 3D-QSAR models are more accurate at predicting the moderately active (i.e. +++) and the moderately inactive compounds (i.e. ++) than predicting the inactive compounds (i.e. +)]. Therefore, from the results in the Table 2 it is not possible to select which of the models has the best predictive power. For this reason, we analyzed how these three 3D-QSAR models perform when predicting a large number of compounds (the so-called *extended test* molecules; see Fig. C, Appendix) that are very similar to the training and test set molecules but which have not been used to build the hypotheses. The results of this analysis showed that the 3D-QSAR model that best predicted the activity range for the extended test molecules was ADDRR.383 (see Table 3).

Therefore, we can conclude that the ADDRR.383 model is the best one obtained in our study because it: a) contains two sites for the aromatic rings from which the proton donor and the acceptor of the intramolecular hydrogen bond hang (Figure 1A) and which are common to all training set molecules; and b) best predicts an external test set of molecules (with a success rate of nearly 50%; see Table 3). At this point, it should be pointed out that, although a predictive power close to 50% seems not to be very successful, we are comparing our pIC_{50} predictions

Table 1. Statistical parameter values for the best 3D-QSAR models obtained in our study. Three of the 1245 3D-QSAR models built by PHASE show good statistics for the comparison between the experimental and predicted pIC_{50} of the training and test set molecules. They are the only 3D-QSAR models that simultaneously meet all following next criteria: a) R^2 , Q^2 , and Pearson r value ≥ 0.8000 ; b) $RMSE/SD$ in the interval $[0.90-1.05]$; c) $Q^2/R^2 \approx 1$ (because the 3D-QSAR needs to predict equally well the molecules in the training and the test sets); and d) $P < 0.05$ (statistical significance). The reference ligand of each hypothesis is the same (i.e. **4e** [8]) but, depending on the hypothesis, it can use one of two different conformations (i.e. conformation 123 for AAADD.1950 and conformation 69 for ADDRR.383 and AAADD.2750; see Fig. 1).

| HYPOTHESIS | PLS factor | SD | R^2 | F | P | RMSE | Q^2 | r -Pearson | RMSE/SD | Q^2/R^2 | Reference ligand | Conformer number |
|------------|------------|--------|--------|------|----------|--------|--------|--------------|---------|-----------|------------------|------------------|
| ADDRR.383 | 2 | 0.5055 | 0.8369 | 46.2 | 8.18E-05 | 0.4600 | 0.8024 | 0.8967 | 0.91 | 0.96 | 4e [8] | 69 |
| AAADD.2750 | 2 | 0.4460 | 0.8730 | 61.9 | 8.58E-06 | 0.4091 | 0.8437 | 0.9209 | 0.92 | 0.97 | 4e [8] | 69 |
| AAADD.1950 | 2 | 0.4352 | 0.8791 | 65.5 | 5.51E-06 | 0.4178 | 0.8370 | 0.9195 | 0.96 | 0.95 | 4e [8] | 123 |

with the very narrow activity ranges that were used to classify the molecules in the patent from which the molecules in the extended test set were borrowed [4] (i.e. $A > 6.3$; $6.3 \geq B > 5.6$; $5.6 \geq C > 5$ and $D \leq 5$; see Table 3). The narrower is the activity scale, the easier it is to incorrectly predict the correct activity range. Thus, a large number of *incorrect* predictions are apparently made when most of the molecules in activity range C are classified in range B (i.e. 13 out of 15) but, interestingly, no molecules that have experimental pIC_{50} in the A range, are predicted to belong to the D range and vice versa. Moreover, when the results that were used to validate PHASE versus Catalyst [13] are compared to the predictions made by ADDRR.383, our model compares favorably. Hence, when the 45 molecules in ranges B and C are considered (i.e. the narrowest ranges used to classify the molecules from the extended test set [4]), then the activity range of 29 of them (i.e. 64.4%) is correctly predicted by ADDRR.383. In contrast, only 8 of the 19 human dihydrofolate reductase inhibitors (i.e. 42.1%) with experimental pIC_{50} values in the same activity ranges were correctly classified when validating PHASE versus Catalyst [13]. Therefore, we can conclude that ADDRR.383 can also perform well at low pIC_{50} values at which it has been reported that PHASE has more problems to correctly predict the activity [13].

3.2 3D-QSAR Model Description

We analyzed the 3D-QSAR model resulting from the ADDRR.383 hypothesis in order to understand how the different moieties from the inhibitor structures contribute, either positively or negatively, to the activity. Figure 2 not only shows the complete three-dimensional aspect of this 3D-QSAR model (see Figure 2A) but also how it fits in the context of: a) the most active compound (i.e. **S_patent** [5, 9]; see Figure 2B); and b) one of the least active compounds (i.e. **41** [6]; see Figure 2C). The blue cubes indicate regions that are favorable for activity; the red cubes indicate regions that are unfavorable for activity; and the areas with no blue or red suggest that they do not affect activity. Moreover, in Figures 2B and 2C, the 3D-QSAR model is

represented only by the cubic volume elements that are occupied by the corresponding ligand.

When the most active compound in our ligand set (i.e. **S_patent** [5, 9]; see Figure 2B) is considered in the context of ADDRR.383, we see: a) an excellent fit with the blue areas of the model; and b) that only a few red cubes are present. Figure 2B also shows that when the intramolecular hydrogen bond is present, there is a hydrophobic region (occupied by the $-OCH_2$ -cyclopropyl substituent from **S_patent**) that favors activity. This is in strong agreement with structure-activity relationship studies on thiophenecarboxamide derivatives that have shown that some hydrophobic substituents can also increase their IKK-2 inhibitory power because they interact with a lipophilic pocket in the ATP binding site of the enzyme [14]. All of this suggests, then, that the $-OCH_2$ -cyclopropyl substituent from **S_patent** can interact also with this pocket.

Figure 2C shows one of the most inactive compounds in our dataset (i.e. **41** [6]) in the context of ADDRR.383. Although the conformation of the molecule that is shown in Figure 2C has the intramolecular hydrogen bond that is present in the most active ligands, in this case the ADDRR.383 model suggests that its inactivity could be caused by the large substituent in the 4 position of the pyridine ring (see Fig. A, Appendix). Thus, when **41** adopts the conformation shown in Figure 2C, this substituent may sterically interact with the receptor. Another important factor for the inactivity of **41** is the lack of an amino substituent in the 2 position of the pyridine ring. When its activity is compared with that of **26** [6] which only differs from **41** because there is an amino and not a hydroxyl substituent at the 2 position of the pyridine ring (see Fig. A, Appendix), the IC_{50} value of **41** is at least two orders of magnitude higher than for **26** (> 20 versus $0.6 \mu M$; see Fig. A, Appendix). In Figure 2C, this is reflected by red cubes over this hydroxyl group. Another important factor for the low activity of **41** is that this molecule does not have a hydrophobic substituent in a location equivalent to the one in the $-OCH_2$ -cyclopropyl group in **S_patent** (see Figs. 2B and 2C) and, therefore, a putative interaction with the lipophilic pocket in the ATP binding site of the enzyme [14] would not be possible.

Table 2. Comparison between the experimental pIC_{50} values of the molecules in the training and test sets and the predicted values that are obtained when the three best 3D-QSAR models obtained by our study are applied. Comparison between the experimental pIC_{50} values of the molecules in the training and test sets (i.e. pIC_{50} Exp. column [4–8]) and the predicted values that are obtained when the three best 3D-QSAR models obtained by our study are applied (i.e. pIC_{50} Pred. columns). The different pIC_{50} values have been classified according to the following activity scale: ‘++++’ for highly active molecules (i.e. $pIC_{50} \geq 8.0$); ‘+++’ for moderately active molecules (i.e. $8.0 > pIC_{50} \geq 7.0$); ‘++’ for moderately inactive molecules (i.e. $7.0 > pIC_{50} \geq 5.0$); and ‘+’ for inactive compounds (i.e. $pIC_{50} < 5.0$). The residual column is computed as the difference between the experimental pIC_{50} and the estimated pIC_{50} and a negative value indicates that the estimated pIC_{50} is higher than the corresponding experimental value.

| Training set | | | | | | | | | | | |
|-----------------------|-----------------|----------------|------------------|----------------|----------|------------------|----------------|----------|------------------|----------------|----------|
| Compound Name | pIC_{50} Exp. | activity scale | ADDRR.383 | | | AAADD.2750 | | | AAADD.1950 | | |
| | | | pIC_{50} Pred. | activity scale | residual | pIC_{50} Pred. | activity scale | residual | pIC_{50} Pred. | activity scale | residual |
| S_patent [5,9] | 8.40 | ++++ | 7.71 | +++ | -0.688 | 8.00 | ++++ | -0.40 | 8.03 | ++++ | -0.37 |
| 4j [8] | 8.07 | ++++ | 7.74 | +++ | -0.331 | 7.62 | +++ | -0.45 | 7.63 | +++ | -0.44 |
| 4e [8] | 7.82 | +++ | 7.78 | +++ | -0.044 | 7.63 | +++ | -0.19 | 7.58 | +++ | -0.24 |
| 4f [8] | 7.70 | +++ | 7.78 | +++ | 0.081 | 7.64 | +++ | -0.06 | 7.59 | +++ | -0.11 |
| 4d [8] | 7.62 | +++ | 7.69 | +++ | 0.07 | 7.54 | +++ | -0.08 | 7.47 | +++ | -0.15 |
| 4h [8] | 7.30 | +++ | 7.38 | +++ | 0.079 | 7.47 | +++ | 0.17 | 7.54 | +++ | 0.24 |
| R_patent [5,9] | 7.23 | +++ | 7.30 | +++ | 0.071 | 7.53 | +++ | 0.30 | 7.55 | +++ | 0.32 |
| 4l [8] | 6.96 | ++ | 7.47 | +++ | 0.511 | 7.23 | +++ | 0.27 | 7.27 | +++ | 0.31 |
| 4b [8] | 6.57 | ++ | 6.94 | ++ | 0.371 | 6.61 | ++ | 0.04 | 6.70 | ++ | 0.13 |
| 22 [6] | 6.30 | ++ | 5.97 | ++ | -0.331 | 6.42 | ++ | 0.12 | 6.38 | ++ | 0.08 |
| 26 [6] | 6.22 | ++ | 4.89 | + | -1.332 | 4.95 | + | -1.27 | 5.02 | ++ | -1.20 |
| 30 [6] | 6.22 | ++ | 6.00 | ++ | -0.222 | 6.28 | ++ | 0.06 | 6.36 | ++ | 0.14 |
| 12 [6] | 6.10 | ++ | 5.84 | ++ | -0.257 | 6.15 | ++ | 0.05 | 6.07 | ++ | -0.03 |
| 27 [6] | 6.00 | ++ | 6.00 | ++ | 0 | 6.36 | ++ | 0.36 | 6.37 | ++ | 0.37 |
| 29 [6] | 5.75 | ++ | 5.33 | ++ | -0.415 | 5.16 | ++ | -0.59 | 5.14 | ++ | -0.61 |
| 43 [6] | 5.60 | ++ | 5.63 | ++ | 0.028 | 6.04 | ++ | 0.44 | 6.01 | ++ | 0.41 |
| 5d [7] | 5.21 | ++ | 5.57 | ++ | 0.362 | 5.03 | ++ | -0.18 | 5.04 | ++ | -0.17 |
| 17 [6] | 4.70 | + | 5.39 | ++ | 0.691 | 5.20 | ++ | 0.50 | 5.18 | ++ | 0.48 |
| 41 [6] | 4.70 | + | 4.92 | + | 0.221 | 4.88 | + | 0.18 | 4.86 | + | 0.16 |
| 40 [6] | 4.70 | + | 5.00 | ++ | 0.301 | 4.86 | + | 0.16 | 4.89 | + | 0.19 |
| 3 [6] | 4.70 | + | 5.53 | ++ | 0.831 | 5.27 | ++ | 0.57 | 5.20 | ++ | 0.50 |
| Test set | | | | | | | | | | | |
| Compound name | pIC_{50} Exp. | Activity scale | ADDRR.383 | | | AAADD.2750 | | | AAADD.1950 | | |
| | | | pIC_{50} Pred. | activity scale | residual | pIC_{50} Pred. | activity scale | residual | pIC_{50} Pred. | activity scale | residual |
| 4k [8] | 7.921 | +++ | 7.73 | +++ | -0.191 | 7.51 | +++ | -0.411 | 7.45 | +++ | -0.471 |
| 4i [8] | 7.824 | +++ | 7.74 | +++ | -0.084 | 7.55 | +++ | -0.274 | 7.56 | +++ | -0.264 |
| 4g [8] | 7.602 | +++ | 7.79 | +++ | 0.188 | 7.63 | +++ | 0.028 | 7.53 | +++ | -0.072 |
| 4a [7] | 7.155 | +++ | 6.60 | ++ | -0.555 | 6.65 | ++ | -0.505 | 6.56 | ++ | -0.595 |
| 4c [8] | 6.921 | ++ | 7.54 | +++ | 0.619 | 7.31 | +++ | 0.389 | 7.27 | +++ | 0.349 |
| 4a [8] | 6.523 | ++ | 6.70 | ++ | 0.177 | 6.47 | ++ | -0.053 | 6.52 | ++ | -0.003 |
| 28 [6] | 6.301 | ++ | 6.25 | ++ | -0.051 | 6.34 | ++ | 0.039 | 6.35 | ++ | 0.049 |
| 20 [6] | 6.155 | ++ | 5.62 | ++ | -0.535 | 5.57 | ++ | -0.585 | 5.49 | ++ | -0.665 |
| 13 [6] | 6.000 | ++ | 5.88 | ++ | -0.12 | 6.07 | ++ | 0.070 | 6.09 | ++ | 0.090 |
| 21 [6] | 5.854 | ++ | 5.23 | ++ | -0.624 | 5.55 | ++ | -0.304 | 5.56 | ++ | -0.294 |
| 23 [6] | 5.854 | ++ | 5.39 | ++ | -0.464 | 5.13 | ++ | -0.724 | 5.09 | ++ | -0.764 |
| 19 [6] | 5.602 | ++ | 5.60 | ++ | -0.002 | 5.59 | ++ | -0.012 | 5.50 | ++ | -0.102 |
| 16 [6] | 5.523 | ++ | 5.22 | ++ | -0.303 | 5.02 | ++ | -0.503 | 5.10 | ++ | -0.423 |
| 42 [6] | 4.824 | + | 5.35 | ++ | 0.526 | 5.22 | ++ | 0.396 | 5.21 | ++ | 0.386 |
| 9 [6] | 4.699 | + | 5.39 | ++ | 0.691 | 5.01 | ++ | 0.311 | 5.05 | ++ | 0.351 |
| 4 [6] | 4.699 | + | 5.67 | ++ | 0.971 | 5.57 | ++ | 0.871 | 5.47 | ++ | 0.771 |

In order to analyze the relative contributions of the blue regions in the ADDR.383 model (see Fig. 2A) from a qualitative point of view, we visualized the most active li-

gand in our set (i.e. **S_patent** [5, 9]; see Fig. A, Appendix) in the context of four decreasing threshold values for favorable interactions (see Fig. 3). Thus, Figure 3A shows

Table 3. Comparison of the experimental activity of the molecules in the extended test set with the activities that are predicted for the same molecules by the best three 3D-QSAR models of our study. Results of the comparison between the experimental activity of the molecules in the extended test set (Exp. activity range columns) and the activities that are predicted for the same molecules by the best three 3D-QSAR models of our study. pIC_{50} values have been classified by using the same ranges used in patent WO02044153 [4] for distributing the activities of the molecules in our extended test set (i.e. $A > 6.3$; $6.3 \geq B > 5.6$; $5.6 \geq C > 5$ and $D \leq 5$). For predicted activities, both the pIC_{50} and its corresponding range are shown.

| Molecule Name | Exp. activity range | Prediction by AAADD. 2750 | Prediction by ADDR. 383 | Prediction by AAADD. 1950 | Molecule Name | Exp. activity range | Prediction by AAADD. 2750 | Prediction by ADDR. 383 | Prediction by AAADD. 1950 |
|---------------|---------------------|---------------------------|-------------------------|---------------------------|---------------|---------------------|---------------------------|-------------------------|---------------------------|
| 1-23 [4] | A | 5.6 (C) | 6.0 (B) | 5.6 (C) | 4-22 [4] | B | 5.5 (C) | 5.8 (B) | 5.5 (C) |
| 2-13 [4] | A | 5.6 (C) | 6.0 (B) | 5.6 (C) | 4-23 [4] | B | 5.5 (C) | 5.9 (B) | 5.5 (C) |
| 2-15 [4] | A | 5.6 (C) | 6.0 (B) | 5.6 (C) | 4-24 [4] | B | 5.7 (B) | 5.9 (B) | 5.7 (B) |
| 2-16 [4] | A | 6.0 (B) | 5.9 (B) | 5.9 (B) | 4-26 [4] | B | 5.6 (C) | 5.9 (B) | 5.6 (C) |
| 3-03 [4] | A | 6.0 (B) | 6.1 (B) | 6.0 (B) | 4-30 [4] | B | 5.6 (C) | 5.6 (C) | 5.6 (C) |
| 3-04 [4] | A | 5.6 (C) | 6.1 (B) | 5.6 (C) | 4-36 [4] | B | 5.5 (C) | 5.8 (B) | 5.5 (C) |
| 3-07 [4] | A | 5.9 (B) | 6.1 (B) | 5.9 (B) | 5-02 [4] | B | 5.7 (B) | 5.7 (B) | 5.7 (B) |
| 3-08 [4] | A | 6.0 (B) | 6.2 (B) | 6.0 (B) | 5-03 [4] | B | 5.8 (B) | 5.7 (B) | 5.8 (B) |
| 4-18 [4] | A | 5.7 (B) | 5.9 (B) | 5.7 (B) | 5-05 [4] | B | 5.7 (B) | 5.8 (B) | 5.7 (B) |
| 4-20 [4] | A | 5.7 (B) | 5.6 (C) | 5.7 (B) | 6-02 [4] | B | 6.0 (B) | 6.0 (B) | 6.1 (B) |
| 4-25 [4] | A | 5.5 (C) | 5.8 (B) | 5.8 (B) | 1-06 [4] | C | 6.0 (B) | 6.1 (B) | 6.0 (B) |
| 4-31 [4] | A | 6.2 (B) | 6.2 (B) | 6.2 (B) | 1-07 [4] | C | 5.6 (C) | 6.0 (B) | 5.6 (C) |
| 6-03 [4] | A | 6.4 (A) | 6.8 (A) | 6.4 (A) | 2-07 [4] | C | 5.5 (C) | 5.7 (B) | 5.6 (C) |
| 8-04 [4] | A | 5.7 (B) | 6.0 (B) | 5.9 (B) | 2-08 [4] | C | 5.7 (B) | 5.9 (B) | 5.7 (B) |
| 2-09 [4] | B | 5.5 (C) | 5.7 (B) | 5.5 (C) | 2-12 [4] | C | 5.6 (C) | 6.0 (B) | 5.6 (C) |
| 2-10 [4] | B | 5.1 (C) | 5.2 (C) | 5.1 (C) | 3-09 [4] | C | 5.1 (C) | 5.4 (C) | 5.2 (C) |
| 2-14 [4] | B | 5.6 (C) | 6.2 (B) | 5.7 (B) | 3-13 [4] | C | 6.1 (B) | 6.2 (B) | 6.1 (B) |
| 2-17 [4] | B | 5.6 (C) | 6.1 (B) | 5.6 (C) | 4-03 [4] | C | 6.1 (B) | 5.8 (B) | 6.2 (B) |
| 2-18 [4] | B | 5.6 (C) | 5.8 (B) | 5.4 (C) | 4-05 [4] | C | 5.6 (C) | 5.7 (B) | 5.6 (C) |
| 3-05 [4] | B | 6.0 (B) | 6.2 (B) | 6.0 (B) | 4-09 [4] | C | 6.1 (B) | 5.7 (B) | 6.0 (B) |
| 3-06 [4] | B | 5.5 (C) | 6.1 (B) | 5.5 (C) | 4-11 [4] | C | 5.7 (B) | 5.9 (B) | 5.7 (B) |
| 4-02 [4] | B | 6.2 (B) | 6.0 (B) | 6.1 (B) | 4-28 [4] | C | 5.6 (C) | 5.9 (B) | 5.6 (C) |
| 4-04 [4] | B | 5.7 (B) | 5.8 (B) | 5.8 (B) | 4-32 [4] | C | 6.1 (B) | 6.0 (B) | 6.0 (B) |
| 4-06 [4] | B | 5.7 (B) | 5.8 (B) | 5.7 (B) | 4-39 [4] | C | 5.6 (C) | 5.6 (C) | 5.7 (B) |
| 4-07 [4] | B | 5.9 (B) | 5.9 (B) | 5.9 (B) | 5-04 [4] | C | 5.9 (B) | 5.7 (B) | 5.7 (B) |
| 4-08 [4] | B | 5.9 (B) | 6.2 (B) | 5.8 (B) | 1-14 [4] | D | 5.5 (C) | 5.9 (B) | 5.5 (C) |
| 4-10 [4] | B | 6.1 (B) | 5.8 (B) | 6.0 (B) | 2-19 [4] | D | 5.6 (C) | 6.0 (B) | 5.6 (C) |
| 4-12 [4] | B | 5.8 (B) | 5.9 (B) | 5.7 (B) | 4-33 [4] | D | 6.1 (B) | 6.0 (B) | 6.0 (B) |
| 4-13 [4] | B | 5.8 (B) | 5.9 (B) | 5.7 (B) | 4-34 [4] | D | 6.1 (B) | 5.9 (B) | 6.1 (B) |
| 4-14 [4] | B | 5.7 (B) | 5.9 (B) | 5.8 (B) | 4-35 [4] | D | 5.8 (C) | 6.1 (B) | 5.8 (B) |
| 4-17 [4] | B | 5.7 (B) | 6.1 (B) | 5.7 (B) | | | | | |
| 4-19 [4] | B | 5.7 (B) | 5.9 (B) | 5.7 (B) | | | | | |
| 4-21 [4] | B | 5.7 (B) | 5.9 (B) | 5.6 (C) | | | | | |

that the main contribution to the activity comes from the oxygen in the oxazin ring (a group that is only present in **S_patent**). The next most important region for the activity is the $-OCH_2$ -cyclopropyl group (see Fig. 3B). It should be pointed out that, in the same location as $-OCH_2$ -cyclopropyl in **S_patent**, the other most active ligands, such as **4j**, **4k**, **4e**, and **4i** [8] always have an $-O$ linked hydrophobic group (see Figs. A and B, Appendix). Figure 3C shows that the next largest contributions to activity arise from the: a) hydroxyl group in the phenolic moiety (basic for making the intramolecular hydrogen bond between the two aromatic rings from **S_patent**); b) amino and carbonyl groups in the oxazin ring; and c) nitrogen in the piperidin ring. Finally, Figure 3D shows that the last contribution to activity is made by the nitrogen from the pyridine ring

(which is the counterpart to the hydroxyl group in the phenyl moiety in the intramolecular hydrogen bond). It should be pointed out that, although this hydrogen bond has been reported to be essential to activity (because it fixes the relative location of the two aromatic rings and consequently the location of the hanging substituents that interact with IKK-2; [6]), it is not detected by the ADDR.383 model as the most important feature for activity. We suggest that this is because the same hydrogen bond is also possible for inactive molecules such as **41** [6] (in these cases, inactivity should be the result of steric hindrance with IKK-2; see Fig. 2C) and, therefore, PHASE does not consider it to have an essential characteristic of the most active pyridine derivatives.

Table 4. Prediction of the pIC_{50} of pyridine derivatives of unknown activity by means of the ADDRR.383 3D-QSAR model. The 3D-QSAR model ADDRR.383 has been used to predict the pIC_{50} value for other pyridine derivatives of unknown activity (see molecules in Figure D, Appendix). The chirality of each molecule is indicated besides its name. For molecules with more than one chiral carbon, see Figure D, Appendix, to see where they are located in the molecular structure; and how they are correlated with the label that describes the chirality of the molecule.

| Molecule name | Estimated activity | Molecule name | Estimated activity | Molecule name | Estimated activity | Molecule name | Estimated activity | Molecule name | Estimated activity | Molecule name | Estimated activity |
|-------------------|--------------------|--------------------|--------------------|-----------------------|--------------------|-----------------------|--------------------|-----------------------|--------------------|-----------------------|--------------------|
| 3q (R) [8] | 7.5 | 5e (RS) [7] | 6.8 | 3c (S) [8] | 6.5 | 1-12 (R) [4] | 6.2 | 3-02 (R) [4] | 6.0 | 32 (R) [6] | 5.8 |
| 3p (R) [8] | 7.5 | 3h (S) [8] | 6.8 | 5e (RR) [7] | 6.5 | 31 (R) [6] | 6.2 | 4-16 (RRS) [4] | 6.0 | 4-38 (R) [4] | 5.8 |
| 3p (S) [8] | 7.5 | 3e (R) [8] | 6.8 | 38 (S) [6] | 6.5 | 3-16 (RS) [4] | 6.2 | 4-37 (S) [4] | 6.0 | 39 (R) [6] | 5.8 |
| 3n (R) [8] | 7.5 | 3c (R) [8] | 6.8 | 3d (R) [7] | 6.5 | 3-11 (RS) [4] | 6.2 | 4-15 (SRR) [4] | 6.0 | 4-16 (SSR) [4] | 5.8 |
| 3k (S) [8] | 7.5 | 38 (R) [6] | 6.8 | 4b (R) [7] | 6.5 | 3a (S) [7] | 6.2 | 25 (R) [6] | 6.0 | 33 (R) [6] | 5.8 |
| 3k (R) [8] | 7.5 | 3d (R) [8] | 6.8 | 5c (S) [7] | 6.5 | 3-11 (SS) [4] | 6.2 | 24 (R) [6] | 6.0 | 33 (S) [6] | 5.8 |
| 3r (R) [8] | 7.5 | 3g (S) [8] | 6.8 | 5h (S) [7] | 6.5 | 3-16 (RR) [4] | 6.2 | 4-29 (RR) [4] | 6.0 | 4-16 (RSS) [4] | 5.7 |
| 3j (R) [8] | 7.4 | 5f (S) [7] | 6.7 | 3f (S) [8] | 6.4 | 5i (R) [7] | 6.2 | 34 (S) [6] | 6.0 | 3-14 (S) [4] | 5.7 |
| 3o (R) [8] | 7.4 | 3e (S) [8] | 6.7 | 3d (S) [7] | 6.4 | 3-12 (R) [4] | 6.2 | 35 (S) [6] | 6.0 | 4-16 (RRR) [4] | 5.7 |
| 3q (S) [8] | 7.3 | 5c (R) [7] | 6.7 | 3b (S) [7] | 6.4 | 3-17 (S) [4] | 6.2 | 1-20 (S) [4] | 6.0 | 3-11 (SR) [4] | 5.7 |
| 3r (S) [8] | 7.2 | 5f (R) [7] | 6.7 | 1-15 (S) [4] | 6.4 | 3-16 (SR) [4] | 6.2 | 24 (S) [6] | 6.0 | 4-16 (SRS) [4] | 5.7 |
| 3m (R) [8] | 7.2 | 5h (R) [7] | 6.7 | 1-15 (R) [4] | 6.4 | 3-15 (S) [4] | 6.2 | 3-14 (R) [4] | 6.0 | 4-15 (RSR) [4] | 5.6 |
| 3n (S) [8] | 7.2 | 4c (R) [7] | 6.7 | 34 (R) [6] | 6.3 | 3-11 (RR) [4] | 6.2 | 4-15 (RRR) [4] | 6.0 | 4-15 (SSR) [4] | 5.6 |
| 3l (S) [8] | 7.2 | 5i (S) [7] | 6.7 | 3-17 (R) [4] | 6.3 | 4-15 (RSS) [4] | 6.2 | 1-12 (S) [4] | 6.0 | 2-11 (S) [4] | 5.6 |
| 3j (S) [8] | 7.1 | 3g (R) [7] | 6.7 | 25 (S) [6] | 6.3 | 1-22 (S) [4] | 6.2 | 4-27 (RS) [4] | 5.9 | 4-16 (SRR) [4] | 5.6 |
| 3i (R) [8] | 7.1 | 3a (S) [8] | 6.6 | 6 (R) [6] | 6.3 | 4b (S) [7] | 6.1 | 4-27 (RR) [4] | 5.9 | 1 (R) [6] | 5.6 |
| 3i (S) [8] | 7.1 | 5g (R) [7] | 6.6 | 1-22 (R) [4] | 6.3 | 3-15 (R) [4] | 6.1 | 6 (S) [6] | 5.9 | 35 (R) [6] | 5.6 |
| 3m (S) [8] | 7.0 | 3b (R) [8] | 6.6 | 1-05 (S) [4] | 6.3 | 3-12 (S) [4] | 6.1 | 4-29 (SS) [4] | 5.9 | 1 (S) [6] | 5.6 |
| 3o (S) [8] | 7.0 | 3a (R) [8] | 6.6 | 36 (R) [6] | 6.3 | 4-37 (R) [4] | 6.1 | 4-27 (SR) [4] | 5.9 | 4-16 (SSS) [4] | 5.6 |
| 3h (R) [8] | 6.9 | 4c (S) [7] | 6.6 | 1-05 (R) [4] | 6.3 | 2-11 (R) [4] | 6.1 | 18 [6] | 5.9 | 4-27 (SS) [4] | 5.5 |
| 3b (R) [7] | 6.9 | 3b (S) [8] | 6.6 | 1-11 (S) [4] | 6.3 | 37 (S) [6] | 6.1 | 4-15 (SRS) [4] | 5.9 | 39 (S) [6] | 5.3 |
| 3f (R) [8] | 6.9 | 5e (SS) [7] | 6.6 | 3f (S) [7] | 6.3 | 3-16 (SS) [4] | 6.1 | 36 (S) [6] | 5.9 | 4-15 (SSS) [4] | 5.3 |
| 3g (R) [8] | 6.9 | 3d (S) [8] | 6.6 | 4-16 (RSR) [4] | 6.3 | 3-02 (S) [4] | 6.1 | 32 (S) [6] | 5.9 | 4-29 (RS) [4] | 5.2 |
| 3c (R) [7] | 6.9 | 5e (SR) [7] | 6.6 | 3g (S) [7] | 6.3 | 4-29 (SR) [4] | 6.1 | 4-15 (RRS) [4] | 5.9 | | |
| 3a (R) [7] | 6.9 | 5g (S) [7] | 6.6 | 37 (R) [6] | 6.3 | 4-38 (S) [4] | 6.1 | 1-20 (R) [4] | 5.8 | | |
| 3l (R) [8] | 6.9 | 3f (R) [7] | 6.5 | 3c (S) [7] | 6.3 | 1-11 (R) [4] | 6.0 | 31 (S) [6] | 5.8 | | |

3.3 IC_{50} Prediction of Pyridine-Based IKK-2 Inhibitors of Unknown Experimental Activity

Finally, the ADDRR.383 model was used to predict the activity of 152 pyridine-based IKK-2 inhibitors (see Fig. D, Appendix) that have ambiguities in their structure (i.e. they have at least one chiral atom) and whose reported IC_{50} or IC_{50} range corresponds to an undetermined mixture of their different stereoisomers [4–8]. The fact that all these molecules have been derived from the same chemical scaffold as the ones used to obtain ADDRR.383 (i.e. the molecules in the training set; see Fig. A, Appendix) suggests that the predicted pIC_{50} values are reliable. The results of this prediction are shown in Table 4 and they indicate that only some of the molecules in Figure D (Appendix) have meaningful differences between the pIC_{50} of their stereoisomers. Thus, the largest differences that are observed between the pIC_{50} of two stereoisomers are 0.9 (for **4_15** and **4_29** [4]) and 0.7 (for **3a** [7] and **4_16** [4]).

4 Conclusions

By using pyridine derivatives that have been shown to act as IKK-2 inhibitors, we have been able: a) to build a pharmacophore that contains all the characteristics that previous structural-activity studies have revealed to be essential for their activity; and b) to use this pharmacophore to build a 3D-QSAR model (i.e. ADDRR.383) that can be used to predict the pIC_{50} of other pyridine derivatives of unknown inhibitory activity versus IKK-2. Thus, the predictive power of ADDRR.383 has been proved to be higher than that of previously published 3D-QSAR models also obtained with PHASE [13]. ADDRR.383 is also expected to be of help in designing other pyridine derivatives with IKK-2 inhibitory activity. Finally, the ADDRR.383 model and its associated pharmacophore is available upon request.

5 Acknowledgements

We thank Mr. John Bates of our University's Language Service for correcting the manuscript and Dr. Lúcia Gon-

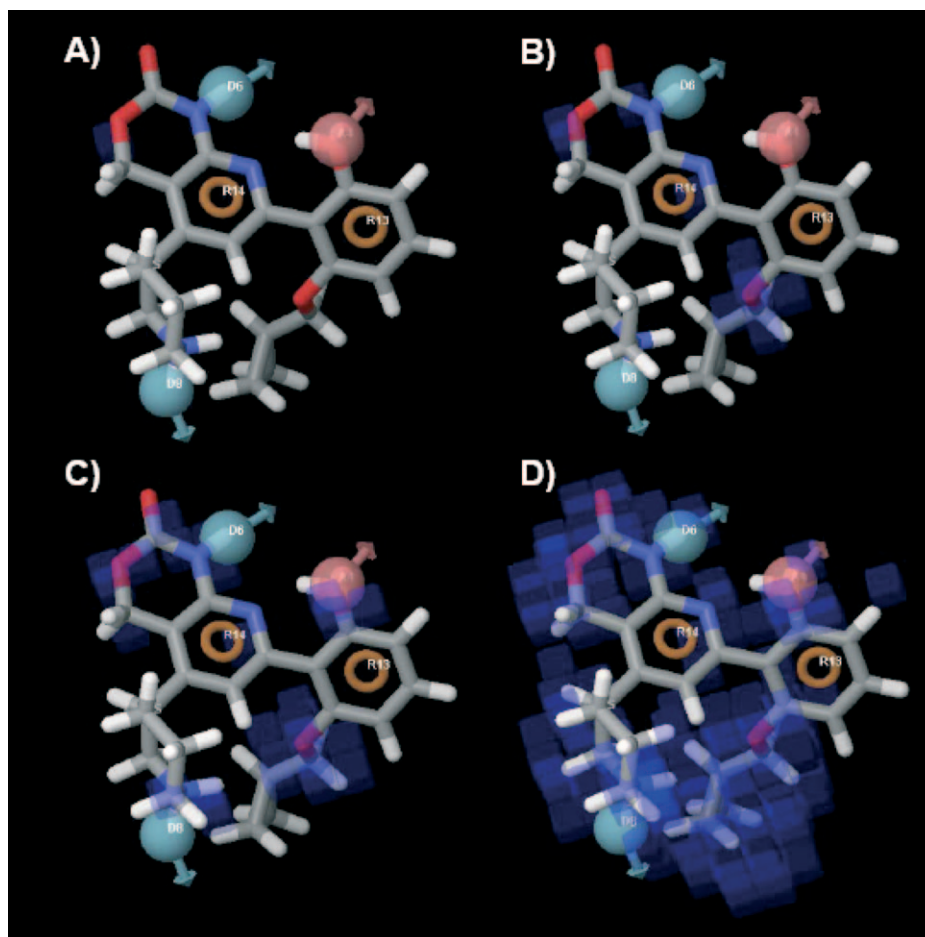


Figure 3. Relative contributions of the regions from ADDRR.383 that are favorable for activity in the context of the most active ligand. Several thresholds are used to qualitatively analyze the relative contribution of the **S_{patent}** moieties [5, 9] to the activity of this ligand. To make the analysis, **S_{patent}** is shown in the context of four decreasing thresholds for the favorable interactions in the ADDRR.383 model (i.e. $2.218e^{-02}$, $2.000e^{-02}$, $1.200e^{-02}$, and $6.000e^{-03}$ for panels A, B, C, and D, respectively). Thus, panel A emphasizes that the main contribution to the activity of this ligand is made by the oxygen in the oxazin ring. Panel B shows that the next most important region for the activity of **S_{patent}** is the $-OCH_2$ -cyclopropyl group. Panel C shows that the next largest contributions to the activity of **S_{patent}** arise from the: a) hydroxyl group in the phenyl moiety (essential for making the intramolecular hydrogen bond between the two aromatic rings); b) amino and carbonyl groups in the oxazin ring; and c) nitrogen in the piperidin ring. Finally, panel D indicates that the last contribution to the activity of **S_{patent}** is made by the nitrogen from the pyridine ring (which is the counterpart to the hydroxyl group in the phenyl moiety in the intramolecular hydrogen bond).

zález for her help in writing it. This study was supported by Grant Number AGL2008-00387/ALI from the Comisión Interministerial de Ciencia y Tecnología (CICYT) and Grant MET-DEV-FUN (CDETI; CENIT Project) from Centro para el Desarrollo Tecnológico Industrial of the Spanish Government. The authors want to thank the Servei de Disseny de Fàrmacs (Drug Design Service) of the Catalonia Supercomputer Center (CESCA) for providing access to Schrödinger software.

6 References

- [1] A. H. Bingham, R. J. Davenport, R. Fosbeary, L. Gowers, R. L. Knight, C. Lowe, D. A. Owen, D. M. Parry, W. R. Pitt, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3622.
- [2] P. P. Tak, G. S. Firestein, *J. Clin. Invest.* **2001**, *107*, 7.
- [3] P. D. G. Coish, P. L. Wickens, T. B. Lowinger, *Expert Opin. Ther. Pat.* **2006**, *16*, 1.
- [4] T. Murata, S. Sakakibara, T. Yoshino, Y. Ikegami, T. Masuda, M. Shimada, T. Shintani, M. Shimazaki, T. B. Lowinger, K. B. Ziegelbauer, K. Fuchikami, M. Umeda, H. Komura, N. Yoshida, *WO02044153* **2002**.
- [5] T. Murata, S. Sakakibara, T. Yoshino, H. Sato, Y. Koriyama, N. Nunami, M. Yamauchi, K. Fukushima, R. Grosser, K. Fuchikami, K. Bacon, T. Lowinger, *WO03076447* **2003**.
- [6] T. Murata, M. Shimada, S. Sakakibara, T. Yoshino, H. Kadono, T. Masuda, M. Shimazaki, T. Shintani, K. Fuchikami,

- K. Sakai, H. Inbe, K. Takeshita, T. Niki, M. Umeda, K. B. Bacon, K. B. Ziegelbauer, T. B. Lowinger, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 913.
- [7] T. Murata, M. Shimada, H. Kadono, S. Sakakibara, T. Yoshino, T. Masuda, M. Shimazaki, T. Shintani, K. Fuchikami, K. B. Bacon, K. B. Ziegelbauer, T. B. Lowinger, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4013.
- [8] T. Murata, M. Shimada, S. Sakakibara, T. Yoshino, T. Masuda, T. Shintani, H. Sato, Y. Koriyama, K. Fukushima, N. Nunami, M. Yamauchi, K. Fuchikami, H. Komura, A. Watanabe, K. Ziegelbauer, K. Bacon, T. Lowinger, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4019.
- [9] K. Ziegelbauer, F. Gantner, N. Lukacs, A. Berlin, K. Fuchikami, T. Niki, K. Sakai, H. Inbe, K. Takeshita, M. Ishimori, H. Komura, T. Murata, T. Lowinger, K. Bacon, *Br. J. Pharmacol.* **2005**, *145*, 178.
- [10] T. Morwick, A. Berry, J. Brickwood, M. Cardozo, K. Castron, M. DeTuri, J. Emeigh, C. Homon, M. Hrapchak, S. Jaconer, S. Jakes, P. Kaplita, T. A. Kelly, J. Ksiazek, M. Liuzzi, R. Magolda, C. Mao, D. Marshall, D. McNeil, A. Prokopowicz, 3rd, C. Sarko, E. Scouten, C. Sledziona, S. Sun, J. Watrous, J. P. Wu, C. L. Cywin, *J. Med. Chem.* **2006**, *49*, 2898.
- [11] W. Long, P. Liu, Q. Li, Y. Xu, J. Gao, *QSAR Comb. Sci.* **2008**, *27*, 1113.
- [12] Cambridge Corporation, Cambridge, MA, USA, **2007**.
- [13] S. L. Dixon, A. M. Smodyrev, E. H. Knoll, S. N. Rao, D. E. Shaw, R. A. Friesner, *J. Comput. Aid. Mol. Des.* **2006**, *20*, 647.
- [14] A. Baxter, S. Brough, A. Cooper, E. Floettmann, S. Foster, C. Harding, J. Kettle, T. McInally, C. Martin, M. Mobbs, M. Needham, P. Newham, S. Paine, S. St-Gallay, S. Salter, J. Unitt, Y. F. Xue, *Bioorg Med Chem Lett* **2004**, *14*, 2817.

Appendix

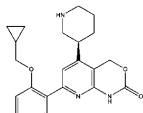
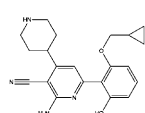
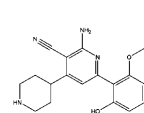
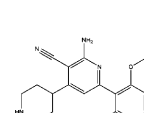
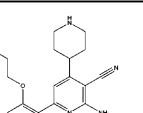
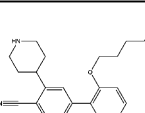
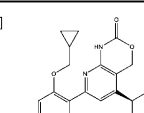
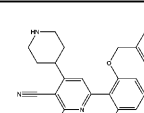
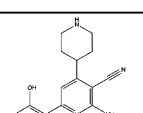
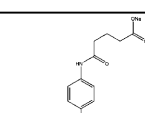
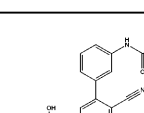
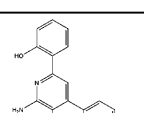
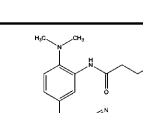
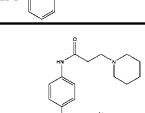
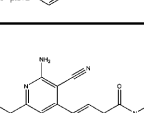
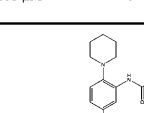
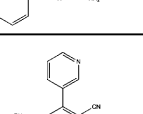
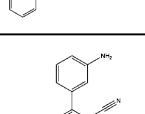
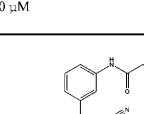
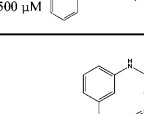
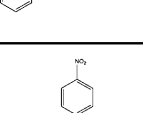
| | | | |
|--|--|---|--|
| <p>S_patent [9]</p>  <p>IC₅₀ = 0.004 μM</p> | <p>4j [8]</p>  <p>IC₅₀ = 0.0085 μM</p> | <p>4e [8]</p>  <p>IC₅₀ = 0.015 μM</p> | <p>4f [8]</p>  <p>IC₅₀ = 0.020 μM</p> |
| <p>4d [8]</p>  <p>IC₅₀ = 0.024 μM</p> | <p>4h [8]</p>  <p>IC₅₀ = 0.050 μM</p> | <p>R_patent [9]</p>  <p>IC₅₀ = 0.059 μM</p> | <p>4i [8]</p>  <p>IC₅₀ = 0.110 μM</p> |
| <p>4b [8]</p>  <p>IC₅₀ = 0.270 μM</p> | <p>22 [6]</p>  <p>IC₅₀ = 0.500 μM</p> | <p>26 [6]</p>  <p>IC₅₀ = 0.600 μM</p> | <p>30 [6]</p>  <p>IC₅₀ = 0.600 μM</p> |
| <p>12 [6]</p>  <p>IC₅₀ = 0.800 μM</p> | <p>27 [6]</p>  <p>IC₅₀ = 1 μM</p> | <p>29 [6]</p>  <p>IC₅₀ = 1.800 μM</p> | <p>43 [6]</p>  <p>IC₅₀ = 2.500 μM</p> |
| <p>5d [7]</p>  <p>IC₅₀ = 6.200 μM</p> | <p>17 [6]</p>  <p>IC₅₀ > 20 μM</p> | <p>41 [6]</p>  <p>IC₅₀ > 20 μM</p> | <p>40 [6]</p>  <p>IC₅₀ > 20 μM</p> |
| <p>3 [6]</p>  <p>IC₅₀ > 20 μM</p> | | | |

Figure A. Training set. Molecules from **S_patent** [9] to **4d** [8] were used to generate the pharmacophore. Molecules from **17** [6] to **3** [6] were used to score hypotheses based on inactive compounds.

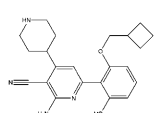
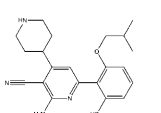
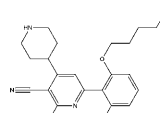
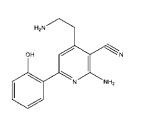
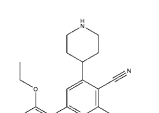
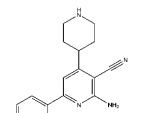
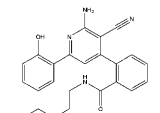
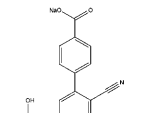
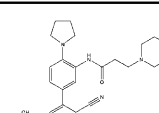
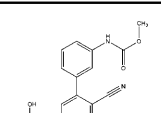
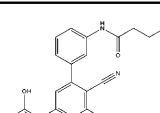
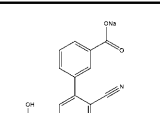
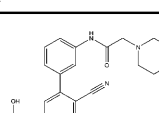
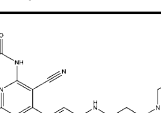
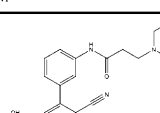
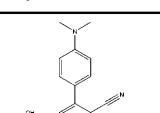
| | | | |
|--|--|---|--|
| <p>4k [8]</p>  <p>IC₅₀ = 0.012 μM</p> | <p>4i [8]</p>  <p>IC₅₀ = 0.015 μM</p> | <p>4g [8]</p>  <p>IC₅₀ = 0.025 μM</p> | <p>4a [7]</p>  <p>IC₅₀ = 0.070 μM</p> |
| <p>4c [8]</p>  <p>IC₅₀ = 0.120 μM</p> | <p>4a [8]</p>  <p>IC₅₀ = 0.300 μM</p> | <p>28 [6]</p>  <p>IC₅₀ = 0.500 μM</p> | <p>20 [6]</p>  <p>IC₅₀ = 0.700 μM</p> |
| <p>13 [6]</p>  <p>IC₅₀ = 1.000 μM</p> | <p>21 [6]</p>  <p>IC₅₀ = 1.400 μM</p> | <p>23 [6]</p>  <p>IC₅₀ = 1.400 μM</p> | <p>19 [6]</p>  <p>IC₅₀ = 2.500 μM</p> |
| <p>16 [6]</p>  <p>IC₅₀ = 3 μM</p> | <p>42 [6]</p>  <p>IC₅₀ = 15 μM</p> | <p>9 [6]</p>  <p>IC₅₀ > 20 μM</p> | <p>4 [6]</p>  <p>IC₅₀ > 20 μM</p> |

Figure B. Test set.

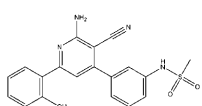
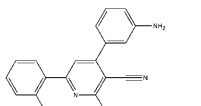
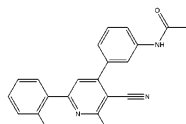
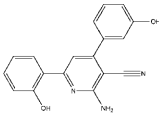
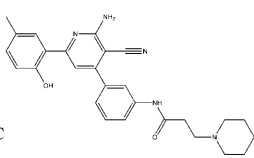
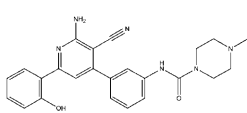
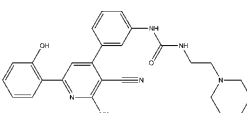
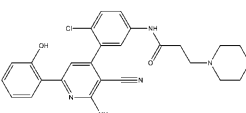
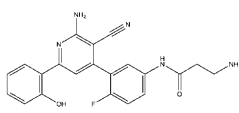
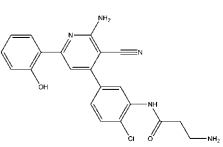
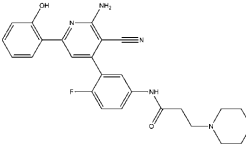
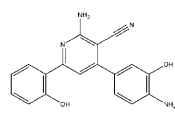
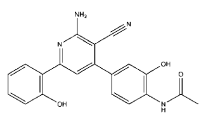
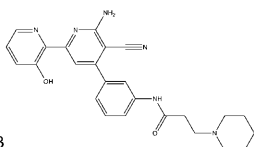
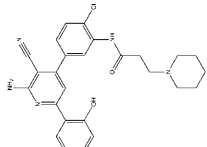
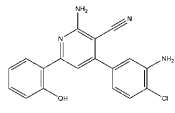
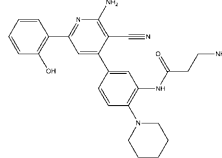
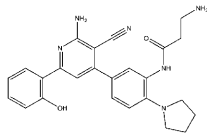
| | | |
|---|---|---|
| <p>1-06 [4]</p>  <p>Activity range: C</p> | <p>1-07 [4]</p>  <p>Activity range: C</p> | <p>1-14 [4]</p>  <p>Activity range: D</p> |
| <p>1-23 [4]</p>  <p>Activity range: A</p> | <p>2-07 [4]</p>  <p>Activity range: C</p> | <p>2-08 [4]</p>  <p>Activity range: C</p> |
| <p>2-09 [4]</p>  <p>Activity range: B</p> | <p>2-10 [4]</p>  <p>Activity range: B</p> | <p>2-12 [4]</p>  <p>Activity range: C</p> |
| <p>2-13 [4]</p>  <p>Activity range: A</p> | <p>2-14 [4]</p>  <p>Activity range: B</p> | <p>2-15 [4]</p>  <p>Activity range: A</p> |
| <p>2-16 [4]</p>  <p>Activity range: A</p> | <p>2-17 [4]</p>  <p>Activity range: B</p> | <p>2-18 [4]</p>  <p>Activity range: B</p> |
| <p>2-19 [4]</p>  <p>Activity range: D</p> | <p>3-03 [4]</p>  <p>Activity range: A</p> | <p>3-04 [4]</p>  <p>Activity range: A</p> |

Figure C. Extended test set.

| | | |
|---|---|---|
| <p>3-05 [4]</p> <p>Activity range: B</p> | <p>3-06 [4]</p> <p>Activity range: B</p> | <p>3-07 [4]</p> <p>Activity range: A</p> |
| <p>3-08 [4]</p> <p>Activity range: A</p> | <p>3-09 [4]</p> <p>Activity range: C</p> | <p>3-13 [4]</p> <p>Activity range: C</p> |
| <p>4-02 [4]</p> <p>Activity range: B</p> | <p>4-03 [4]</p> <p>Activity range: C</p> | <p>4-04 [4]</p> <p>Activity range: B</p> |
| <p>4-05 [4]</p> <p>Activity range: C</p> | <p>4-06 [4]</p> <p>Activity range: B</p> | <p>4-07 [4]</p> <p>Activity range: B</p> |
| <p>4-08 [4]</p> <p>Activity range: B</p> | <p>4-09 [4]</p> <p>Activity range: C</p> | <p>4-10 [4]</p> <p>Activity range: B</p> |
| <p>4-11 [4]</p> <p>Activity range: C</p> | <p>4-12 [4]</p> <p>Activity range: B</p> | <p>4-13 [4]</p> <p>Activity range: B</p> |

Figure C. Extended test set (continued).

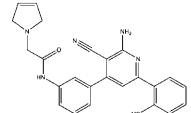
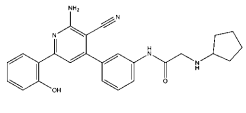
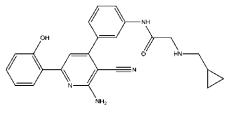
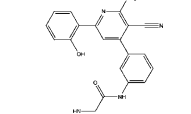
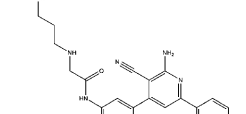
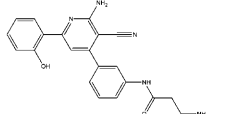
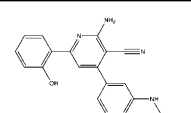
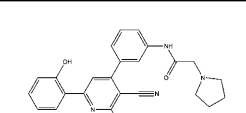
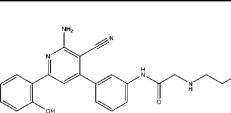
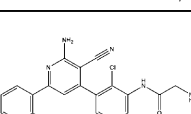
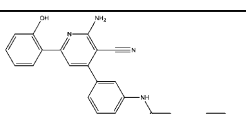
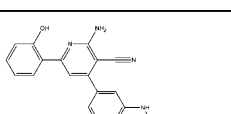
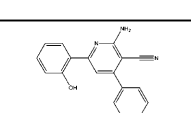
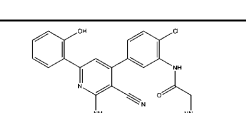
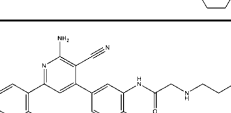
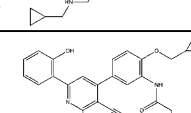
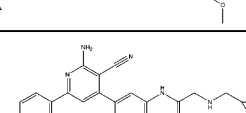
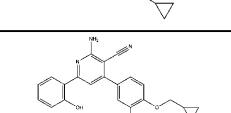
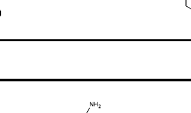
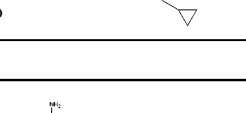
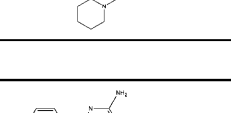
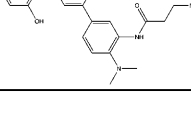
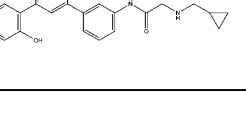
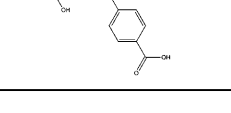
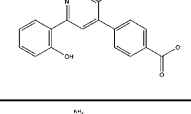
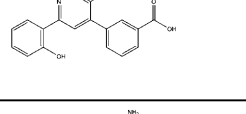
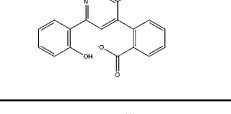
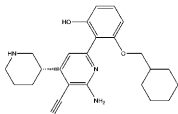
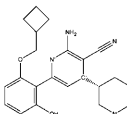
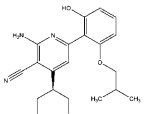
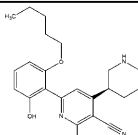
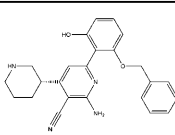
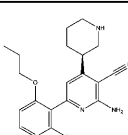
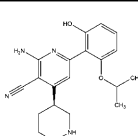
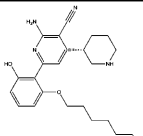
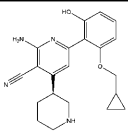
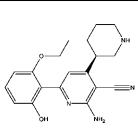
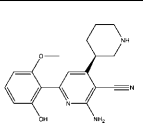
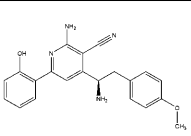
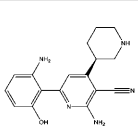
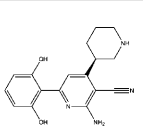
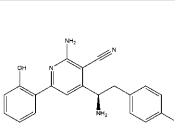
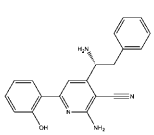
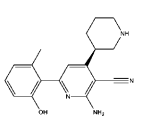
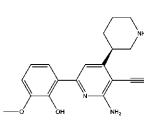
| | | |
|--|--|--|
| <p>4-14 [4]</p>  <p>Activity range: B</p> | <p>4-17 [4]</p>  <p>Activity range: B</p> | <p>4-18 [4]</p>  <p>Activity range: A</p> |
| <p>4-19 [4]</p>  <p>Activity range: B</p> | <p>4-20 [4]</p>  <p>Activity range: A</p> | <p>4-21 [4]</p>  <p>Activity range: B</p> |
| <p>4-22 [4]</p>  <p>Activity range: B</p> | <p>4-23 [4]</p>  <p>Activity range: B</p> | <p>4-24 [4]</p>  <p>Activity range: B</p> |
| <p>4-25 [4]</p>  <p>Activity range: A</p> | <p>4-26 [4]</p>  <p>Activity range: B</p> | <p>4-28 [4]</p>  <p>Activity range: C</p> |
| <p>4-30 [4]</p>  <p>Activity range: B</p> | <p>4-31 [4]</p>  <p>Activity range: A</p> | <p>4-32 [4]</p>  <p>Activity range: C</p> |
| <p>4-33 [4]</p>  <p>Activity range: D</p> | <p>4-34 [4]</p>  <p>Activity range: D</p> | <p>4-35 [4]</p>  <p>Activity range: D</p> |
| <p>4-36 [4]</p>  <p>Activity range: B</p> | <p>4-39 [4]</p>  <p>Activity range: C</p> | <p>5-02 [4]</p>  <p>Activity range: B</p> |
| <p>5-03 [4]</p>  <p>Activity range: B</p> | <p>5-04 [4]</p>  <p>Activity range: C</p> | <p>5-05 [4]</p>  <p>Activity range: B</p> |
| <p>6-02 [4]</p>  <p>Activity range: B</p> | <p>6-03 [4]</p>  <p>Activity range: A</p> | <p>8-04 [4]</p>  <p>Activity range: A</p> |

Figure C. Extended test set (continued).

| | | |
|---|---|---|
| 3q [8]  | 3p [8]  | 3n [8]  |
| 3k [8]  | 3r [8]  | 3j [8]  |
| 3m [8]  | 3l [8]  | 3o [8]  |
| 3i [8]  | 3h [8]  | 3b [7]  |
| 3f [8]  | 3g [8]  | 3c [7]  |
| 3a [7]  | 3e [8]  | 3c [8]  |

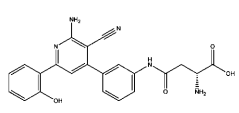
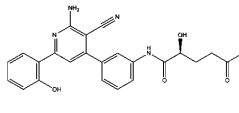
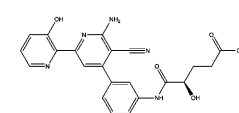
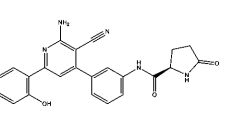
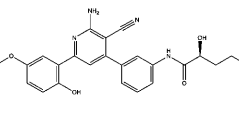
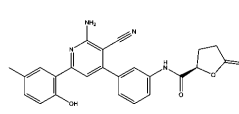
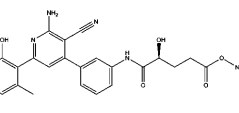
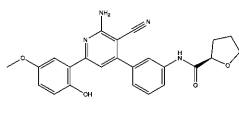
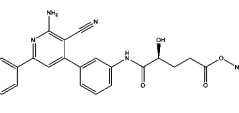
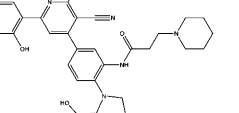
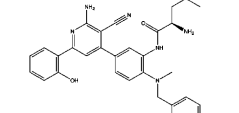
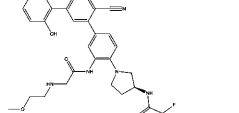
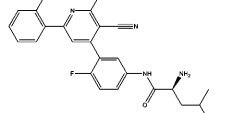
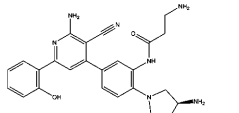
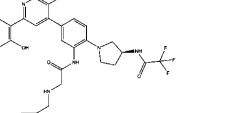
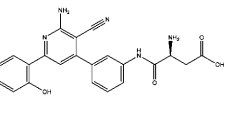
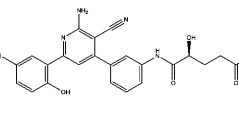
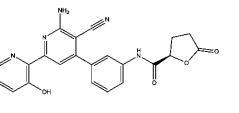
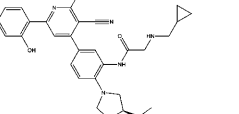
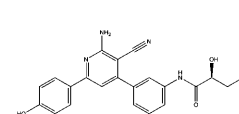
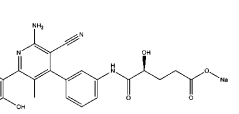
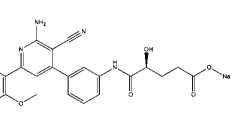
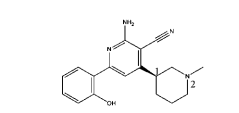
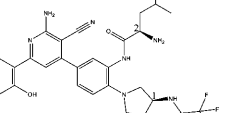
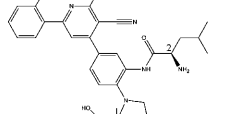
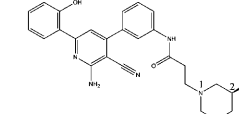
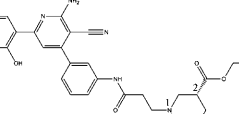
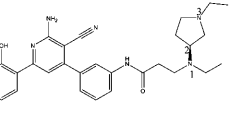
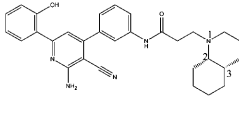
When more than one chiral carbon is present in a molecule, the numbers are used to: (a) indicate their relative location; and (b) correlate them with the label that describes their stereochemistry in Table 4. Thus, in this Table, the stereochemistry label SSR for the molecule **4-16** [4] indicates that the chiral carbons labeled 1 and 2 in this Figure are in S configuration whereas the one labeled 3 is in R.

Figure D. Molecules of unknown IC50.

| | | |
|------------|--------------|--------------|
| 38 [6] | 3d [8] | 5f [7] |
| 5c [7] | 5h [7] | 4c [7] |
| 5i [7] | 3g [7] | 3a [8] |
| 5g [7] | 3b [8] | 3f [7] |
| 3d [7] | 4b [7] | 1-15 [4] |
| 34 [6] | 3-17 [4] | 1 [6] |

When more than one chiral carbon is present in a molecule, the numbers are used to: (a) indicate their relative location; and (b) correlate them with the label that describes their stereochemistry in Table 4. Thus, in this Table, the stereochemistry label SSR for the molecule 4-16 [4] indicates that the chiral carbons labeled 1 and 2 in this Figure are in S configuration whereas the one labeled 3 is in R.

Figure D. Molecules of unknown IC₅₀ (continued).

| | | | | | |
|----------|---|----------|---|----------|---|
| 25 [6] |  | 6 [6] |  | 1-22 [4] |  |
| 1-05 [4] |  | 36 [6] |  | 1-11 [4] |  |
| 37 [6] |  | 1-12 [4] |  | 31 [6] |  |
| 3-12 [4] |  | 3-15 [4] |  | 4-37 [4] |  |
| 2-11 [4] |  | 3-02 [4] |  | 4-38 [4] |  |
| 24 [6] |  | 35 [6] |  | 1-20 [4] |  |
| 3-14 [4] |  | 32 [6] |  | 39 [6] |  |
| 33 [6] |  | 5e [7] |  | 3-16 [4] |  |
| 3-11 [4] |  | 4-29 [4] |  | 4-27 [4] |  |
| 4-16 [4] |  | 4-15 [4] |  | | |

When more than one chiral carbon is present in a molecule, the numbers are used to: (a) indicate their relative location; and (b) correlate them with the label that describes their stereochemistry in Table 4. Thus, in this Table, the stereochemistry label SSR for the molecule 4-16 [4] indicates that the chiral carbons labeled 1 and 2 in this Figure are in S configuration whereas the one labeled 3 is in R.

Figure D. Molecules of unknown IC₅₀ (continued).

11 beta hydroxysteroid dehydrogenase type 1 (11 β -HSD1)

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

Manuscript 4

A Structure-Based *In Silico* Methodology for Predicting 11beta-HSD2 Inhibitors

Esther Sala, Laura Guasch, Miquel Mulero, Cristina Valls, Maria-Josepa
Salvadó, Anna Arola-Arnal, Santiago Garcia-Vallvé and Gerard Pujadas

Submitted needing revision to Journal of Computer-Aided Molecular Design

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

A Structure-Based *In Silico* Methodology for Predicting 11beta-HSD2 Inhibitors

Esther Sala[§], Laura Guasch[§], Miquel Mulero[§], Cristina Valls[§], Maria-Josepa Salvadó[§], Anna Arola-Arnal[§], Santiago Garcia-Vallvé^{§,†}, Gerard Pujadas^{§,†*}.

[§]Grup de Recerca en Nutrigenòmica, Departament de Bioquímica i Biotecnologia, Universitat Rovira i Virgili, Campus de Sescelades, Tarragona, Catalonia, Spain

[†]Centre Tecnològic de Nutrició i Salut, Reus, Catalonia, Spain

*e-mail: gerard.pujadas@urv.cat; Tel: +34977 559565; Fax: +34977558232

Abstract

Glucocorticoids are ubiquitous hormones that play a key role in modulating immune and inflammatory responses and regulating energy metabolism, cardiovascular homeostasis and the body's responses to stress. The principal glucocorticoid is cortisol, which has a concentration in target tissues that is modulated by 11 β -HSD1 (which reduces cortisone to cortisol) and by 11 β -HSD2 (which converts cortisol to cortisone). A growing body of evidence suggests that increased 11 β -HSD1 activity within target tissues may promote insulin resistance, obesity, hypertension and dyslipidemia. Therefore, the search for 11 β -HSD1 inhibitors has been an active area of study during recent years. Unfortunately, there are many examples of 11 β -HSD1 inhibitors that cause side effects, such as sodium retention and hypertension, because they also inhibit 11 β -HSD2. Thus, one important task to consider when finding new 11 β -HSD1 inhibitors consists of evaluating, by means of *in vitro* assays, their lack of activity on 11 β -HSD2. Thus, we have developed a robust, easy-to-use and accurate *in silico* approach that is able to discriminate between molecules that inhibit 11 β -HSD2 and those that do not inhibit that target in a molecular database. This virtual screening filter can be of use either to significantly reduce the expenses needed to find potent and selective 11 β -HSD1 inhibitors (because it could lower the number of molecules that are experimentally tested and increase the chance of finding molecules that inhibit 11 β -HSD1 but not 11 β -HSD2) or to find 11 β -HSD2 inhibitors (of current interest, for instance, as anti-cancer drugs).

Keywords

Metabolic syndrome, homology model, protein-ligand docking, similarity search, 11beta-hydroxysteroid dehydrogenase type 1, 11beta-hydroxysteroid dehydrogenase type 2.

Abbreviations

11 β -HSD: 11 β -hydroxysteroid dehydrogenase; 11 β -HSD1: 11 β -hydroxysteroid dehydrogenase type 1; 11 β -HSD2: 11 β -hydroxysteroid dehydrogenase type 2; ET_combo: the sum of ET_pb and ST; ET_pb: electrostatic Tanimoto using full Poisson-Boltzmann electrostatics; GA: glycyrrhetic acid; GCs: glucocorticoids; IFD: Induced Fit Docking; PDB: Protein Data Bank; SDR: short-chain dehydrogenase/reductase; ST: shape Tanimoto; VS: virtual screening.

Introduction

Glucocorticoids (GCs) are well-known ubiquitous hormones that play a key role in modulating immune and inflammatory responses and regulating energy metabolism, cardiovascular homeostasis and the body's responses to stress [1]. In fact, GCs are thought to modulate the expression of up to 20% of the genes in the mammalian genome [2]. Normalization of GC levels can reverse metabolic syndrome by improving obesity, insulin resistance, hypertension and lipid and lipoprotein profiles [3]. The principal GC is cortisol, whose concentration in target tissues — together with the activation of GC receptors— is tightly regulated by tissue-specific enzymes called 11 β -hydroxysteroid dehydrogenases (11 β -HSDs). 11 β -hydroxysteroid dehydrogenase type 1 (11 β -HSD1) is a NADPH-dependent enzyme that is predominantly expressed in liver, adipose tissue and skeletal muscles, where it increases intracellular GC action by catalyzing the reduction of inactive 11-ketoglucocorticoids (*i.e.*, cortisone in humans) to active 11 β -hydroxyglucocorticoids (*i.e.*, cortisol in humans). On the other hand, corticosteroid 11 β -hydroxysteroid dehydrogenase type 2 (11 β -HSD2) is a NAD⁺-dependent enzyme that is highly expressed in classical aldosterone-selective target tissues (such as the distal nephron, colon, sweat glands and placenta), where it catalyzes the reaction opposite to that of 11 β -HSD1 and thus prevents the binding of cortisol to the mineralocorticoid receptor and subsequently protects it from GCs [4].

A growing body of evidence suggests that increased 11 β -HSD1 activity within target tissues may promote insulin resistance, obesity, hypertension and dyslipidemia. Therefore, the search for 11 β -HSD1 inhibitors has been an active area of study during recent years [1,5-16]. Unfortunately, there are many examples of 11 β -HSD1 inhibitors that cause side effects, such as sodium retention and hypertension, because they also inhibit 11 β -HSD2 [1,17-20]. Thus, one important task to consider when finding new 11 β -HSD1 inhibitors consists of evaluating, by means of *in vitro* assays, their lack of activity on 11 β -HSD2 [10,21-25].

The increasing interest in detecting specific 11 β -HSD1 inhibitors for the therapeutic treatment of GC-dependent diseases has encouraged us to develop a structure-based *in silico* methodology that can be used to discriminate between molecules that inhibit 11 β -HSD2 and those that do not inhibit that target in a molecular database. To achieve this goal, we have: (a) developed a homology model of 11 β -HSD2, (b) docked into this model a set of actives (*i.e.*, molecules that are known to inhibit 11 β -HSD2 irrespective of their activity on 11 β -HSD1) and a set of non-actives (*i.e.*, molecules that are known to not inhibit 11 β -HSD2 molecules), (c) used each active pose as the query molecule in an electrostatic potential and shape comparison with the poses of the rest of the active and non-active molecules and quantified the corresponding similarities with different

Tanimoto similarity scores, (d) evaluated the performance of each active pose to discern between actives and non-actives, and (e) identified which combination of query molecule, docking methodology and Tanimoto similarity score is best for discerning between actives and non-actives.

Finally, all of the tools developed in the current study (*i.e.*, the homology model for 11 β -HSD2 and the query pose that best discerns between active and non-active molecules) are available to interested readers upon request.

Materials and computational methods

Homology model building

11 β -HSD2 belongs to the classical short-chain dehydrogenase/reductase (SDR) family [26], has a sequence length of 405 residues and has no experimental 3D structure available from the Protein Data Bank (<http://www.pdb.org>; PDB) [27]. To build a homology model for the human protein, its sequence was retrieved from the UniProt database (<http://www.uniprot.org>), where its accession number is P80365.

Suitable templates were searched from the PDB with the *Sequence Search* tool available at the PDB website. Thus, four templates from different SDR families were selected: (a) 17 β -HSD1 (PDB code: 1A27; resolution: 1.90 Å; and R-value: 0.210) with 28% sequence identity and 41% positive residues in the pairwise alignment (according to the similarity matrix BLOSUM62); (b) clavulanic acid dehydrogenase (PDB code: 2JAP; resolution: 2.10 Å; and R-value: 0.186) with 27% and 45% identity and positive residues in the pairwise alignment, respectively; (c) actinorhodin ketoreductase (PDB code: 2RHC; resolution: 2.10 Å; and R-value: 0.186) with 27% and 36% identity and positive residues in the pairwise alignment, respectively; and (d) β -keto acyl carrier protein reductase (PDB code: 1EDO; resolution: 2.30 Å; and R-value: 0.190) with 23% and 42% identity and positive residues in the pairwise alignment, respectively.

Then, a homology model was built for the 11 β -HSD2 segment between Val80 and Ile350 with Prime v2.1 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) in a five-step workflow. Thus, in the first step, a sequence multi-alignment was built with the sequences of the human 11 β -HSD2 and the sequences of four templates (*i.e.*, 1A27, 2JAP, 2RHC and 1EDO) by using —with default parameter values— the *multi sequence viewer* tool available from the Schrödinger suite 2009. During the second step, the multi-alignment was exported to the Prime's *align structures* window, and structural alignment was performed by superimposing the binding

sites of the four templates. In the third step, different regions from each template were predominantly used to build the initial homology model for 11 β -HSD2: (a) 2JAP was mainly used to model the N-terminal region (where the NAD⁺ cofactor is located); (b) 1A27 was mainly used to model the C-terminal region; and (c) 2RHC was mainly used to model the ligand region because 2RHC's ligand —*i.e.*, 3-methyl-1,6,8-trihydroxyanthraquinone— binds in a manner similar to that expected for the 11 β -HSD2 substrate and therefore gave us an idea of the possible position of cortisol on the substrate binding site. Once the initial model was finished, cortisol was introduced in its substrate binding site (replacing the former 2RHC ligand), and the model was submitted to an *Induced Fit Docking* protocol (IFD; Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) [28]. During this IFD run, the side chains of the 11 β -HSD2 catalytic triad (*i.e.*, Ser219, Tyr232 and Lys236) were made flexible during the *Initial Glide Docking* step, and extra precision docking (XP) was used during the *Glide Redocking* step (the rest of the running conditions were by default). In the fourth step, the homology model was submitted to a loop refinement with Prime's *Refinement* window. Thus, the parameters used in that refinement were different sampling methods that were chosen depending on the loop length (*i.e.*, the default recommended a parameter for loops of 5 or fewer residues and an extended low parameter for loops within 6 to 11 residues). The minimum distance between atoms as a fraction of their van der Waals radii was the default value (*i.e.*, 0.70; meaning that the distance between atoms must be at least 70% of their ideal van der Waals separation). After the refinement, an energy minimization was performed with the help of the *Protein Preparation Wizard* from Schrödinger suite 2009 with default conditions. Finally, in the fifth step, the quality of the resulting homology model was evaluated by analyzing the: (a) outliers in the Ramachandran plot and (b) interactions between cortisol and the homology model binding site and comparing them with information described in the literature [22,29-33].

Ligand set up

Experimental data (*i.e.*, IC₅₀ or % of inhibition) were collected from the literature for 258 11 β -HSD1 inhibitors [7,22,23,29,34-42]. Then, these molecules were further split into three sets (*i.e.*, *active*, *non-active* and *other* sets) by considering their experimental activity on 11 β -HSD2. Hence, molecules were included in the *active* set only if they had either a percentage of inhibition $\geq 80\%$ or IC₅₀ ≤ 20 nM; molecules were included in the *non-active* set only if they had either a percentage of inhibition $\leq 30\%$ or IC₅₀ ≥ 100 nM; the remaining molecules were included in the *other* set and were not further considered for training purposes. At this point, it is important to remember that the goal of the present work was to develop an *in silico* methodology that can be used for removing putative 11 β -HSD1 inhibitors that also have a good chance of inhibiting 11 β -HSD2.

Thus, this justifies choosing the decision criterion used to select a molecule as a member of the *active* set to be based, exclusively, on its capacity as a potent 11 β -HSD2 inhibitor (irrespective of its activity as an 11 β -HSD1 inhibitor). Consequently, a molecule was classified as a member of the *non-active* set based, exclusively, on its low capacity for inhibiting 11 β -HSD2 (irrespective of its activity as an 11 β -HSD1 inhibitor). As a result of this classification, 29 and 177 molecules were assigned to the *active* and *non-active* sets, respectively.

The 3D-structure for all active and non-active molecules was built using ChemDraw Ultra v11.0 (CambridgeSoft Corporation, Cambridge, MA, USA; <http://www.cambridgesoft.com/>). Then, the structures were cleaned and minimized with LigPrep v2.3 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>). During this cleaning process, the running conditions were as follows: (a) use the OPLS_2005 force field, (b) generate with Ionizer all possible states at pH 7.00, (c) do not generate tautomers, and (d) determine the chiralities from the 3D structure.

To avoid any bias of the enrichment metrics calculations by the existence of very similar structures inside each ligand set, *Extended Connected Fingerprints* (ECFP4) were calculated for all of the molecules in the *active* set, followed by a *Diversity Analysis* (where the soergel distance was taken as a metric and the diversity selection method was the sphere exclusion method with a value of 0.5 for the exclusion sphere size). These calculations were done with Canvas v1.2.211 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) and were also performed for the *non-active* set. Thus, the numbers of molecules in the representative *active* and *non-active* sets were 20 and 140, respectively (see supplementary tables S1 and S2). The molecules in these two sets were used for subsequent calculations.

Poses for ligands in the representative *active* and *non-active* sets were obtained by docking all of them in the ligand-binding site of the 11 β -HSD2 homology model (by using this model as a rigid body and only allowing rotation around the ligand's single bonds). Two types of docking runs were performed with Glide v5.5 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) [43]: (a) SP (Standard Precision) and (b) XP (Extra Precision). All conditions used during the docking calculations were by default, with the exception of the maximum number of poses per ligand, which was set up to 32.

Pose similarity analysis

Comparisons between poses were performed with EON v2.0.1 (OpenEye Scientific Software, Inc., Santa Fe, New Mexico, USA; <http://www.eyesopen.com>) [44]. During such comparisons, each

pose from the *active* set was used as a *query* and compared with all of the poses from: (a) the rest of the *active* set members (*i.e.*, the remaining poses of the query molecule were not included in this calculation) and (b) the molecules in the *non-active* set. The similarity between poses was quantified by EON with different kinds of Tanimoto similarity scores: (a) the electrostatic Tanimoto using full Poisson-Boltzmann electrostatics with an outer dielectric of 80 (*i.e.*, *ET_pb*), which compares the electrostatic potential of two small molecules (where 1 corresponds to identical potentials and negative values correspond to an overlap of positive and negative charges) [45]; (b) the Tanimoto shape (*i.e.*, *ST*), which is a quantitative measure of three-dimensional overlap [where 1 corresponds to a perfect overlap (*i.e.*, the same shape)] [46]; and (c) the sum of *ST* and *ET_pb* (*i.e.*, *ET_combo*). When comparing the query pose with all of the poses of any other molecule (either from the *active* or *non-active* set), only the highest Tanimoto similarity scores (*i.e.*, *ET_pb*, *ST* and *ET_combo*) were kept.

Selection of the conditions that best discriminate between active and non-active molecules

The running conditions that best discriminate between active and non-active molecules were identified with the help of the python script *enrichment.py* (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>), which allows for the calculation of different enrichment metrics [47], including: (a) the BEDROC (**B**oltzmann-**E**nhanced **D**iscrimination of **R**eceiver-**O**perating **C**haracteristic) area under the curve, (b) the ROC (**R**eceiver **O**perator **C**haracteristic) area under the curve, (c) the AUC (Area **U**nder the receiver-operating **C**haracteristic curve), (d) the RIE (**R**obust **I**nitial **E**nhancement), (e) different kinds of enrichment factors and (f) the percentage of actives at 1%, 2%, 5%, 10% and 20% of the database.

Results and Discussion

Homology model description

Figures 1A and 1B show the homology model derived for 11 β -HSD2 with bound cortisol and NAD⁺. Although the model does not include the first 79 residues of the enzyme, which are not conserved among SDR family members, this had no influence on the subsequent protein-ligand docking results that used this model because this sequence segment is not part of the ligand-binding site [30]. The multi-alignment that was used to guide the homology modeling process is shown in Figure 1C, where it can be seen that 11 β -HSD2 is well aligned with 17 β -HSD1 (although 17 β -HSD1 has some loops that are longer than their equivalent in 11 β -HSD2). It is also well

known that members of the SDR family share a Ser-Tyr-Lys triad in the catalytic center (see Figures 1B and 1C) [48], and this information was taken into account while building the homology model. Figure 2 shows that only 2% of the non-glycine residues (*i.e.*, Arg131, Thr159, Glu268, Arg271 and Phe349) are at disallowed regions of the Ramachandran plot. Remarkably, none of these residues are located on the substrate- or cofactor-binding sites, and therefore, their conformations did not have any influence on subsequent protein-ligand docking calculations that used our 11 β -HSD2 homology model.

11 β -HSD2 is a NAD⁺-dependent enzyme, and this cofactor is one of the main promoters in the beginning of proton movement between key residues of the catalytic site and the substrate. Thus, NAD⁺ accepts the hydrogen liberated from the cortisol's hydroxyl onto position 4 of the nicotinamide ring. Accordingly, Figure 1B shows that our model is coherent with previous studies that have shown that NAD⁺ is hydrogen-bonded to Ser92, Phe94 and Lys236 (which, through a hydrogen bond with the 2'-hydroxyl group of the ribose, stabilizes the position of the nicotinamide moiety) [30]. Interestingly, the positively charged side chain of Lys236 also induces an unusually low pKa for Tyr232, which deprotonates its phenolic group, and allows this latter residue to act as a catalytic base that extracts the hydrogen from the cortisol's hydroxyl group and oxidizes it to become the cortisone's ketone [26,49,50]. In agreement with other 11 β -HSD2 homology models [30,51], Figure 1B shows that, in our model, Tyr232 is also able to form the above-mentioned intermolecular interaction with the cortisol's hydroxyl group. In contrast with other studies that propose a role in the stabilization of the cortisol's hydroxyl group [26,50,52] for Ser219's hydroxyl group, our model shows that Ser219 can use its hydroxyl to stabilize the location of the Tyr232 side chain (see Figure 1B). Nevertheless, it is also remarkable that the structure of the cofactor- and the substrate-binding sites of our 11 β -HSD2 homology model are well defined and agree with most of the proposed mechanisms that have been described in the literature [26,30,49-53].

Identification of the active pose that best discerns between active and non-active molecules

The 20 and 140 molecules that were assigned to the *active* and *non-active* sets (see supplementary tables S1 and S2) according to their capacity to inhibit 11 β -HSD2 were docked on the ligand-binding site of the 11 β -HSD2 homology model with the SP and XP running modes of Glide v5.5 [43,54]. During these calculations, Glide used a hierarchical series of filters to search for possible locations of the ligands in the active-site region of 11 β -HSD2. During that search, the shape and properties of 11 β -HSD2 were represented on a grid by several different sets of fields that provide progressively more accurate scoring of the ligand poses. Consequently, plausible bioactive poses were obtained for each molecule in the *active* and *non-active* sets by the two Glide modes. Then,

all poses for each active molecule were used as a query in a systematic shape and electrostatic-potential map comparison that was performed with EON v2.0.1 [44]. Thus, during such comparisons, each pose from the *active* set was used as a *query* and compared with all of the poses from: (a) the rest of the *active* set members (*i.e.*, the remaining poses of the query molecule were not included in this calculation) and (b) the molecules in the *non-active* set. At this point, it is worth mentioning that EON results only make sense when pre-aligned molecules are compared, and in that sense, poses derived from docking active and non-active molecules on the 11 β -HSD2 homology model fulfill such a requirement. Although the similarity between each pair of poses was evaluated with three different Tanimoto similarity scores (*i.e.*, *ET_pb*, *ST* and *ET_combo*), a further analysis of the characteristics of active molecules (see supplementary table S1) suggested to use *ET_combo* (the sum of *ST* and *ET_pb*) to measure the similarity between poses rather than using only *ET_pb* or *ST*.

The key requirement for success for any virtual screening (VS) filter is that it must be able to discern between active and non-active molecules. From a practical point of view, this means that when active and non-active molecules are merged in the same set and the filter is applied, active molecules must be ranked very early. In the present work, we searched for the combination of query pose and Glide's running condition that, with the *ET_combo* results from EON as the similarity score between poses, was able to remove from a list of predicted 11 β -HSD1 inhibitors those molecules that are also able to inhibit 11 β -HSD2. The analysis of the different enrichment metrics that were obtained from the EON results revealed that the percent of active molecules at the top 20% of the list was useful for that aim. Table 1 shows which combinations of query pose and Glide's running conditions were able to concentrate at the top 20% at least 90% of all of the actives. Moreover, the data in Table 1 also show the corresponding ROC values for these five combinations. Thus, considering that (a) the ROC gives the probability that a randomly chosen known active will rank higher than a randomly chosen non-active and (b) a ROC value of 1 corresponds to an ideal screen performance while 0.5 reflects random behavior, it can be concluded that the five combinations shown in Table 1 reflect a high VS yield. At this point, it is worth mentioning that, although the proportion of non-active molecules per active compound is not very high (*i.e.*, 6.7 instead of 36 for DUD [55]), which could bring into question the validity of the enrichment metric values obtained with our results, it must be considered that, as opposed to DUD, our non-active molecules have been selected according to experimental data. Therefore, this supports the correctness of our query choice.

It has recently been suggested that 11 β -HSD2 inhibitors bind to the 11 β -HSD2-binding site in a flipped way relative to their preferred binding to 11 β -HSD1, and thus, their 11-keto-oxygen points

away from the catalytic triad in 11 β -HSD2 [48]. Accordingly, when the five poses shown in Table 1 are analyzed in the context of the 11 β -HSD2-binding site, all of them are in this flipped orientation (see Figure 3). A further comparison of the spatial location of these five poses showed that the one that is more coherent with the position that is expected according to the results from Kratschmar et al. [48] is the fourth pose from **STX351** (**STX351/4** hereafter; see Figure 3B). Therefore, the combinations of query pose and Glide's running conditions that were selected for building a VS filter that is able to remove from a set of 11 β -HSD1 inhibitors those that are predicted to have a good chance of inhibiting 11 β -HSD2 are the ones formed by **STX351/4** and Glide SP (see Table 1 and Figure 3B). Moreover, further analysis of the characteristics of the query pose (see section below) confirmed the validity of this selection.

Description of Glide SP's STX351/4

STX351 is a non-selective 11 β -HSD inhibitor that, at a concentration of 10 μ M, produces percentages of inhibition in 11 β -HSD1 and 11 β -HSD2 of 70% and 94.4%, respectively [41]. Its scaffold was derived from 18 β -glycyrrhetic acid (GA), a natural molecule that is well known to be the principal active ingredient of glycyrrhizin, which is contained in the roots and rhizomes of licorice (*Glycyrrhiza spp.*) and which is a non-selective inhibitor of 11 β -HSD activity. Different studies have suggested that GA and sterols may serve as the lead compound for synthesizing new and selective 11 β -HSD1 blockers [23,34,48,56-59].

When **STX351/4** is analyzed, it can be seen that it is able to adopt the correct orientation on the 11 β -HSD2 substrate binding site [48], collapsing the cortisol entrance and making hydrophobic contacts on the 11 β -HSD2 steroid binding site (see Figure 4). Figures 4B and 4C show that this pose does not form hydrogen bonds with any binding site residue (a situation that is common for most of the active poses that were obtained by Glide; results not shown).

Figure 5 shows the results of the best EON comparison between the Glide SP's **STX351/4** (Figure 5A) and three different non-active molecules that were docked in the 11 β -HSD2 ligand-binding site in the same conditions as the query molecule. These three molecules were selected among the non-actives that gave the highest *ET_combo* values when compared with **STX351/4** (while also considering the greatest chemical structure variability possible). Thus, Figures 5B, 5C and 5D show the pose with the highest *ET_combo* value for comparison of the query pose with **Cpd04** [34], **10j2** [40] and **1** [38], respectively. Remarkably, these poses (all from molecules belonging to different chemical families) show different electrostatic potential maps but quite similar shapes relative to **STX351/4**, which translates into low *ET_combo* values (see Figure 5) and that confirms

the validity of using this parameter to measure the similarity between pose pairs. A further analysis of Figure 5 gives the following results:

- **Cpd04** is a known selective 11 β -HSD1 inhibitor [34]. According to its chemical structure, it is a pentacyclic triterpene of the ursan type (a family of molecules that are widespread in the plant kingdom). Remarkably, bioactivity has been described for many molecules from this family [34,60,61]. Thus, different studies have reported clinical interest in triterpenes, either from the ursan (*e.g.*, **Cpd04**) or from the oleanan (*e.g.*, **STX351**) types, because they show anti-inflammatory, hypoglycemic, anti-viral and apoptosis-inducing properties [8,34,62]. As mentioned before, our study shows that **Cpd04/8** and **STX351/4** have very different electrostatic potential maps (see Figures 5A and 5B), and this could be the reason for their different inhibition patterns (*i.e.*, **Cpd04** is a selective 11 β -HSD1 inhibitor whereas **STX351** is a non-selective 11 β -HSD inhibitor with a high percentage of inhibition for 11 β -HSD1 and 11 β -HSD2). One possible explanation for the different inhibition patterns for **Cpd04** and **STX351** could be the lack of a substituent on position 20 of the triterpene **Cpd04**'s scaffold (which is associated with 11 β -HSD2 inhibitors [48]). In parallel, the shape component of this *ET_combo* is quite high, as would be expected from the similarity between their triterpene-derived scaffolds. Summarizing, **Cpd04** and **STX351** present a clear example of two molecules with a common scaffold but different bioactivity.
- **10j2** is a potent and selective *in vitro* inhibitor of 11 β -HSD1 that was synthesized after a rationale-guided molecular model that took advantage of the two large hydrophobic pockets that are present in the ligand-binding site of 11 β -HSD1 [40]. The *ET_combo* for the comparison between **10j2/20** and **STX351/4** has a negative value (*i.e.*, -0.021; see Figure 5C) for the electrostatic potential contribution, which could be attributed to the large negative electrostatic potentials that the carboxyl and ureylene groups give to **10j2** (where the negative electrostatic potential for the latter group has no counterpart in **STX351/4**; see Figures 5A and 5C) and could explain the differences in the inhibition patterns for **Cpd04** and **STX351** (*i.e.*, potent and selective 11 β -HSD1 inhibitor *versus* non-selective 11 β -HSD inhibitor). In contrast, the shape Tanimoto contribution to *ET_combo* is significantly high (even slightly higher than the one obtained for the comparison between **Cpd04** and **STX351**, which share the triterpene scaffold; see Figure 5B).
- **1** is an adamantyl amide derivate that is a potent and selective 11 β -HSD1 inhibitor [38] that belongs to a class of well-known inhibitors of this enzyme [18,29,38,58,63]. The

ET_combo for the comparison between **1/11** and **STX351/4** has a modest value for the electrostatic potential contribution (*i.e.*, 0.125; see Figure 5D) that, nevertheless, is higher than those obtained for **Cpd04/8** and **10j2/20** (*i.e.*, 0.005 and -0.021, respectively; see Figures 5B and 5C). In contrast, the shape Tanimoto contribution to *ET_combo* is lower (*i.e.*, 0.435; see Figure 5D) than that obtained for **Cpd04/8** and **10j2/20** (*i.e.*, 0.557 and 0.582, respectively; see Figures 5B and 5C), which can be attributed to the scaffold differences and/or to the higher flexibility of the **1** structure when compared with the more rigid structures of **STX351**, **Cpd04** and **10j2**.

Conclusions

We have developed a robust, easy-to-use and accurate *in silico* approach that is able to discriminate between molecules that inhibit 11 β -HSD2 and those that do not inhibit that target in a molecular database. Thus, this VS filter can be of use either to significantly reduce the expenses needed to find potent and selective 11 β -HSD1 inhibitors (because it could decrease the number of molecules that are experimentally tested and increase the chance of finding molecules that inhibit 11 β -HSD1 but not 11 β -HSD2) or to find 11 β -HSD2 inhibitors (of current interest as anti-cancer drugs [19,48,64], to promote potassium excretion and hyperkalemia prevention in chronic hemodialysis patients [65] or for the treatment of chronic inflammation of the colon [66]). Thus, our data suggest that if: (1) Glide SP and our 11 β -HSD2 homology model are applied to a molecular database to generate probable poses at the steroid binding site, (2) the resulting poses are used by EON in a shape and electrostatic potential map comparison with **STX351/4**, (3) one pose per molecule is selected (*i.e.*, the one with the highest *ET_combo* value for the query), (4) molecules are sorted according to descending *ET_combo* values, and (5) the *ET_combo* values for the top 20% molecules are in the range [1.463, 0.562], then the chance that a molecule in the top 20% is a 11 β -HSD2 inhibitor is 94.4% (and 5.6% if the molecule is selected from the remaining 80%).

Obviously, this is a knowledge-based approach that would greatly benefit from the following: (a) delivery to the PDB of the experimental model for 11 β -HSD2 and (b) publication of activity data and structures for new families of specific and non-specific 11 β -HSD inhibitors. Nevertheless, although we assume that such new data could help to find an even better query pose, our results demonstrate that a very high predictive power can be achieved with the current data and workflow. At this point, we would like to emphasize the importance of using EON during the calculation of the similarity between two poses because this program has been shown to be very effective in

using known bioactive molecules for finding new lead-hopping candidates in molecular databases by smoothing differences in chemical structures and translating them into criteria that are important for their intermolecular interactions with the ligand-binding site [67].

Finally, it is worth to mention that a VS study for predicting 11 β -HSD1 inhibitors that do not inhibit 11 β -HSD2 (and that uses the *in silico* approach developed in the current work) is under way and its results will be published elsewhere.

Acknowledgements

This manuscript was edited for English language fluency by American Journal Experts. We thank Dr. Justyna Iwaskiewicz (Molecular Modeling Group at the Swiss Institute of Bioinformatics) and Dr. Daniel D. Robinson (Schrödinger, LLC) for their useful advice during the homology modeling of 11 β -HSD2. This study was supported by Grant Number AGL2008-00387/ALI from the Ministerio de Educación y Ciencia of the Spanish Government and the ACCIÓ (TECCT10-1-0008) program (Generalitat de Catalunya). The authors wish to thank the Servei de Disseny de Fàrmacs (Drug Design Service) of the Catalonia Supercomputer Center (CESCA) for providing access to Schrödinger software.

Electronic Supplementary Material

The supplementary material shows the 2D structures of all of the molecules in the *active* and the *non-active* sets together with their corresponding inhibitory activity relative to 11 β -HSD1 and 11 β -HSD2.

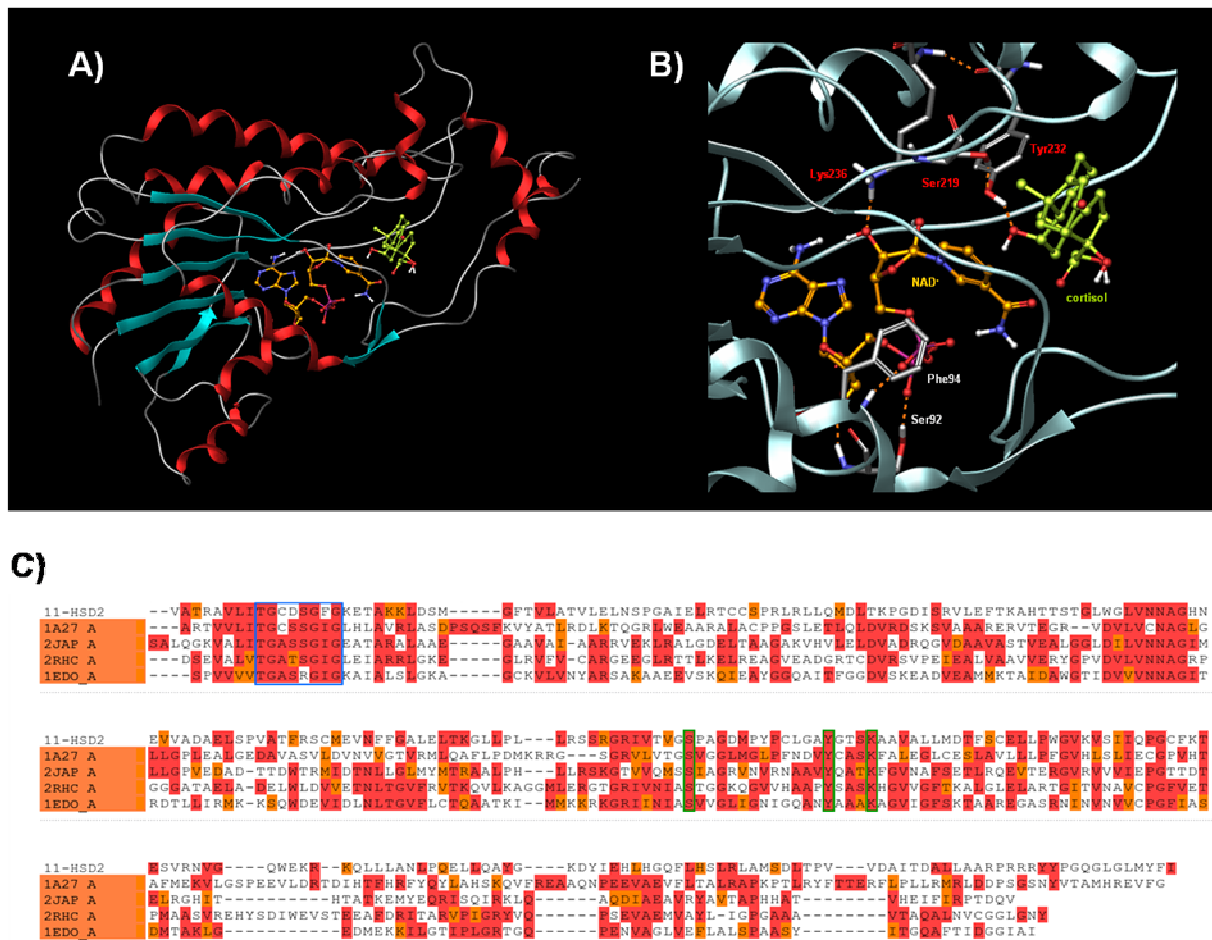


Figure 1. The homology model for human 11 β -HSD2. Figure 1A shows the folding predicted for 11 β -HSD2 with a bound substrate (*i.e.*, cortisol) and cofactor (*i.e.*, NAD⁺). In this panel, the protein is shown as a ribbon model (with α -helices and β -strands colored in red and cyan, respectively) and cortisol (predominantly in green) and NAD⁺ (predominantly in brown) as a ball and stick model. Figure 1B shows a close view of the ligand- and cofactor-binding sites. Here, hydrogen bonds are shown as dashed lines, and the labels that identify the 11 β -HSD2 catalytic triad (*i.e.*, Ser219, Tyr232 and Lys236) are shown in red. Figure 1C shows the multiple sequence alignment that was used to guide the modeling. In the multi-alignment, most significant SDR motifs are highlighted by means of boxes (*i.e.*, the TGxxx[AG]xG that corresponds to the cofactor-binding pattern is highlighted by a blue box, and the residues that form the catalytic triad are highlighted by a green box).

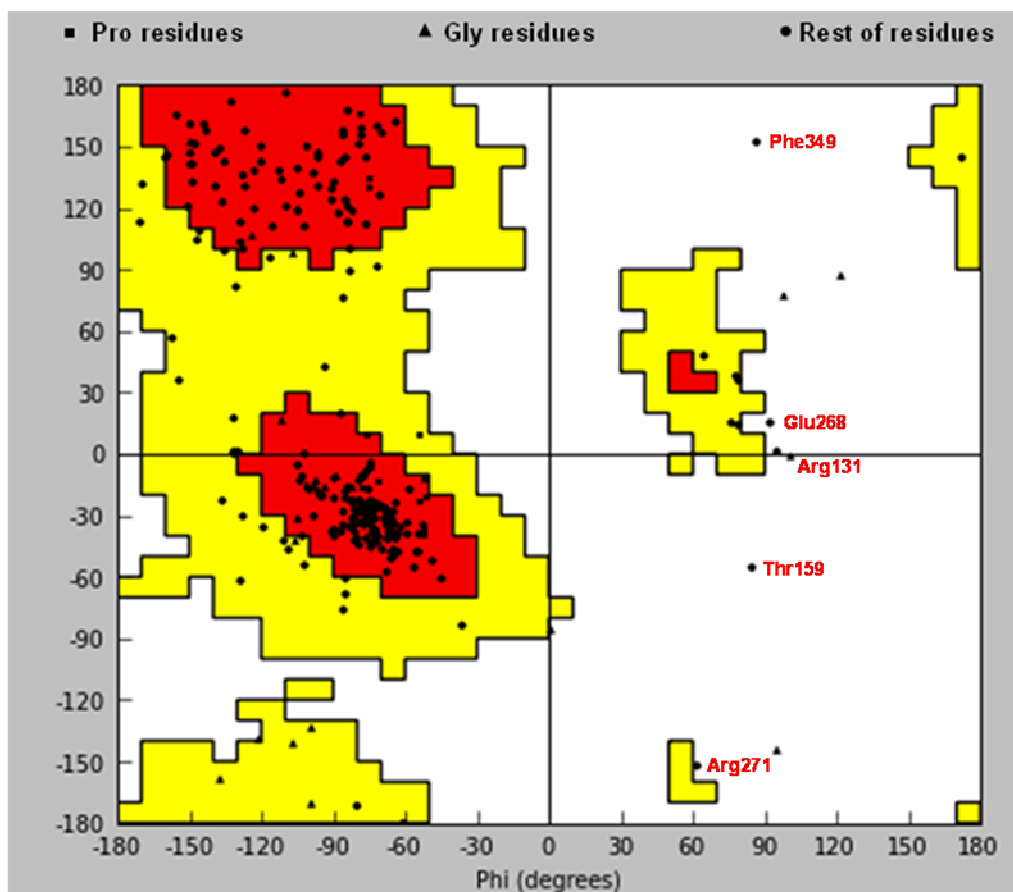


Figure 2. The Ramachandran plot of our 11β-HSD2 homology model. Residues different from Gly in disallowed regions of the Ramachandran plot are shown.

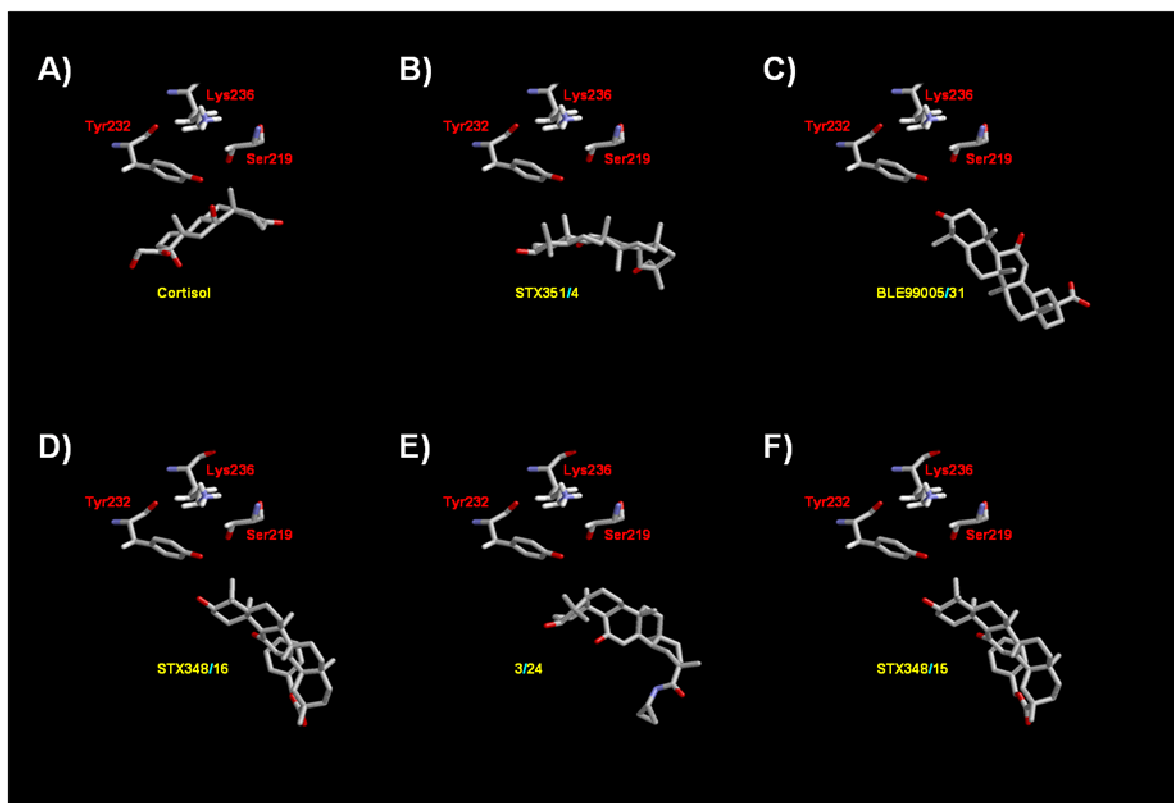


Figure 3. Relative location of cortisol and the best query poses in the context of the 11 β -HSD2 binding site. All panels have the same orientation to allow for an easier comparison between the different pose locations. The location of the 11 β -HSD2 catalytic triad (*i.e.*, Ser219, Tyr232 and Lys236) is shown as a reference. All panels in this figure were obtained with RasMol (<http://rasmol.org/>).

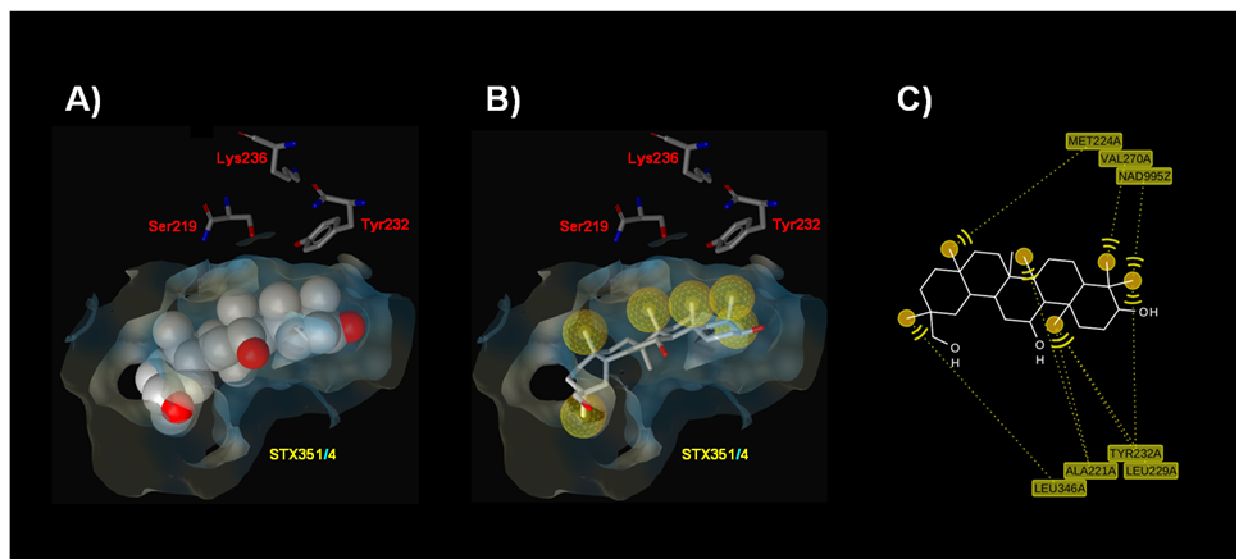


Figure 4. Intermolecular interactions between Glide SP's **STX351/4** and 11β-HSD2. Figure 4A shows how **STX351/4** (in a spacefill model) fits at the 11β-HSD2 binding site (schematically shown by its surface). Figure 4B shows the location of the hydrophobic sites that characterize the structure-derived pharmacophore resulting from the complex between 11β-HSD2 and **STX351/4**. Figure 4C shows the intermolecular interactions for this complex. In panels 4A and 4B, the location of the 11β-HSD2 catalytic triad (*i.e.*, Ser219, Tyr232 and Lys236) is shown as a reference. All panels in this figure were obtained with LigandScout v3.0 (<http://www.inteligand.com/ligandscout3/>).

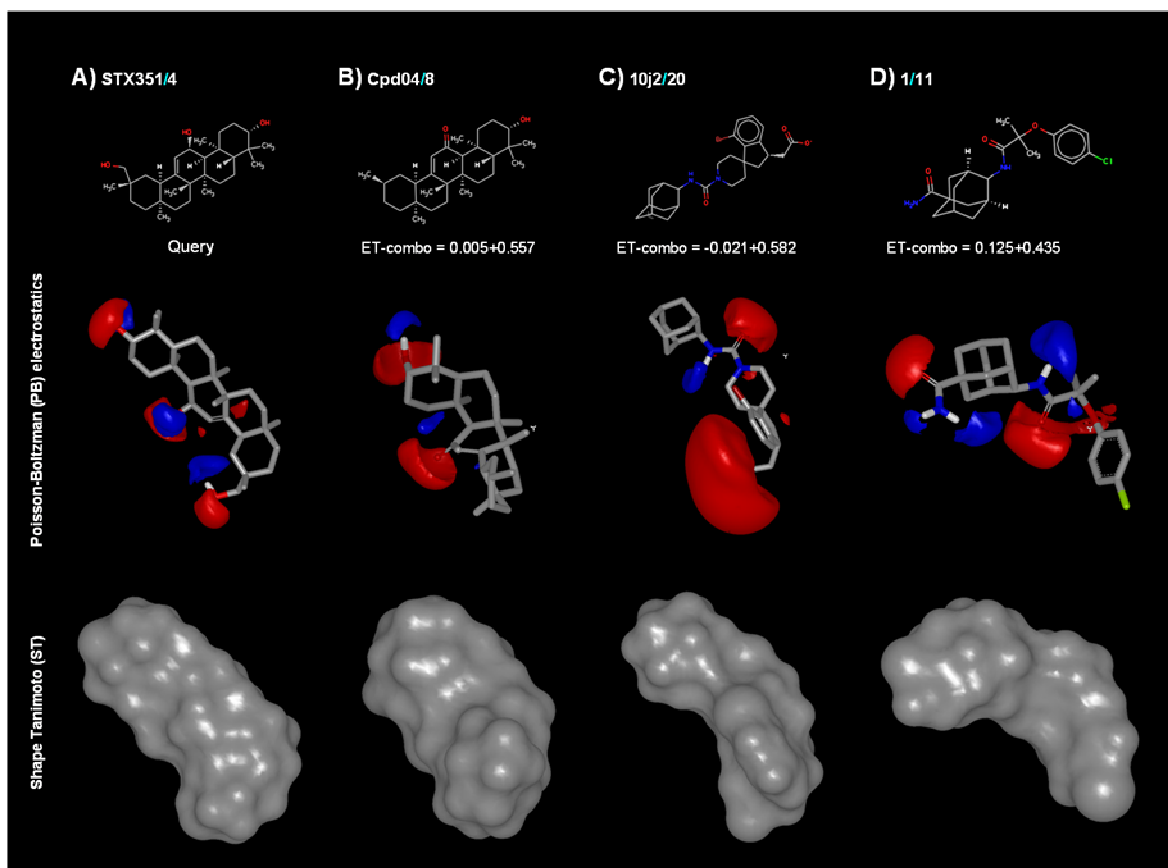


Figure 5. Results of the best EON comparison between Glide SP's **STX351/4** and three different non-active molecules. All molecules shown in this figure were docked into the 11 β -HSD2 ligand-binding site under the same conditions. Non-active molecules were selected among the ones that gave the highest *ET_combo* value when compared with **STX351/4** (while also considering the greatest chemical structure variability possible). For clarity, the electrostatic potential fields and molecular shapes are displayed separately. All panels have the same orientation to allow for an easier comparison and were obtained with VIDA v4.0.3 (<http://www.eyesopen.com/vida>).

Table 1. Top five predictive query poses for active molecules during EON comparisons.

| Query molecule/pose number | Reference | Glide's running mode | % 11 β -HSD inhibition | | % Actives 20% | % Non- actives 20% | ROC |
|----------------------------------|-----------|----------------------------|---------------------------------|--------|---------------------|--------------------------|------|
| | | | Type 1 | Type 2 | | | |
| STX351/4 | [41] | Glide-SP | 70.02 | 94.41 | 94.4 | 8.8 | 0.95 |
| BLE99005/31 | [41] | Glide-SP | 86.90 | 100.0 | 90.0 | 9.5 | 0.93 |
| STX348/16 | [41] | Glide-SP | 63.39 | 94.77 | 94.7 | 8.8 | 0.95 |
| 3/24 | [36] | Glide-SP | 90.00 | 103.0 | 94.4 | 8.8 | 0.94 |
| STX348/15 | [41] | Glide-SP | 63.39 | 94.77 | 94,7 | 8.8 | 0.95 |

References

1. Wamil M, Seckl JR (2007) *Drug Discov Today* 12(13-14):504
2. Galon J, Franchimont D, Hiroi N, Frey G, Boettner A, Ehrhart-Bornstein M, O'Shea JJ, Chrousos GP, Bornstein SR (2002) *FASEB J* 16(1):61
3. Inagaki K, Otsuka F, Miyoshi T, Watanabe N, Suzuki J, Ogura T, Makino H (2004) *Endocr J* 51(2):201
4. Seckl JR WB (2001) *Endocrinology* 142(4):1371
5. Staab CA, Maser E (2010) *J Steroid Biochem Mol Biol* 119(1-2):56
6. Morton NM (2010) *Mol Cell Endocrinol* 316(2):154
7. Zhu Y, Olson SH, Hermanowski-Vosatka A, Mundt S, Shah K, Springer M, Thieringer R, Wright S, Xiao J, Zokian H, Balkovec JM (2008) *Bioorg Med Chem Lett* 18(11):3405
8. Rosenstock J, Banarar S, Fonseca VA, Inzucchi SE, Sun W, Yao W, Hollis G, Flores R, Levy R, Williams WV, Seckl JR, Huber R, Investigators I-P (2010) *Diabetes Care* 33(7):1516
9. Tiwari A (2010) *IDrugs* 13(4):266
10. Jean DJS, Yuan C, Bercot EA, Cupples R, Chen M, Fretland J, Hale C, Hungate RW, Komorowski R, Veniant M, Wang MH, Zhang XP, Fotsch C (2007) *J Med Chem* 50(3):429
11. Schnackenberg CG (2008) *Curr Opin Investig Drugs* 9(3):295
12. St Jean DJ, Wang M, Fotsch C (2008) *Curr Top Med Chem* 8(17):1508
13. Ge R, Huang Y, Liang G, Li X (2010) *Curr Med Chem* 17(5):412
14. Boyle CD, Kowalski TJ (2009) *Expert Opin Ther Pat* 19(6):801
15. Rollinger JM, Kratschmar DV, Schuster D, Pfisterer PH, Gumy C, Aubry EM, Brandstötter S, Stuppner H, Wolber G, Odermatt A (2010) *Bioorg Med Chem* 18(4):1507

16. Gummy C, Thurnbichler C, Aubry EM, Balazs Z, Pfisterer P, Baumgartner L, Stuppner H, Odermatt A, Rollinger JM (2009) *Fitoterapia* 80(3):200
17. Wang H, Ruan Z, Li JJ, Simpkins LM, Smirk RA, Wu SC, Hutchins RD, Nirschl DS, Van Kirk K, Cooper CB, Sutton JC, Ma Z, Golla R, Seethala R, Salyan ME, Nayeem A, Krystek SR, Sheriff S, Camac DM, Morin PE, Carpenter B, Robl JA, Zahler R, Gordon DA, Hamann LG (2008) *Bioorg Med Chem Lett* 18(11):3168
18. Rohde JJ, Pliushchev MA, Sorensen BK, Wodka D, Shuai Q, Wang J, Fung S, Monzon KM, Chiou WJ, Pan L, Deng X, Chovan LE, Ramaiya A, Mullally M, Henry RF, Stolarik DF, Imade HM, Marsh KC, Beno DW, Fey TA, Droz BA, Brune ME, Camp HS, Sham HL, Frevert EU, Jacobson PB, Link JT (2007) *J Med Chem* 50(1):149
19. Walker EA, Stewart PM (2003) *Trends Endocrinol Metab* 14(7):334
20. Ferrari P (2010) *Biochim Biophys Acta* 1802(12):1178
21. Wan ZK, Chenail E, Xiang J, Li HQ, Ipek M, Bard J, Svenson K, Mansour TS, Xu X, Tian X, Suri V, Hahm S, Xing Y, Johnson CE, Li X, Qadri A, Panza D, Perreault M, Tobin JF, Saiah E (2009) *J Med Chem* 52(17):5449
22. Julian LD, Wang Z, Bostick T, Caille S, Choi R, DeGraffenreid M, Di Y, He X, Hungate RW, Jaen JC, Liu J, Monshouwer M, McMinn D, Rew Y, Sudom A, Sun D, Tu H, Ursu S, Walker N, Yan X, Ye Q, Powers JP (2008) *J Med Chem* 51(13):3953
23. Beseda I, Czollner L, Shah PS, Khunt R, Gaware R, Kosma P, Stanetty C, Del Ruiz-Ruiz MC, Amer H, Mereiter K, Da Cunha T, Odermatt A, Classen-Houben D, Jordis U (2010) *Bioorg Med Chem* 18(1):433
24. Yang H, Shen Y, Chen J, Jiang Q, Leng Y, Shen J (2009) *Eur J Med Chem* 44(3):1167
25. Classen-Houben D, Schuster D, Da Cunha T, Odermatt A, Wolber G, Jordis U, Kueenburg B (2009) *J Steroid Biochem Mol Biol* 113(3-5):248
26. Kavanagh KL, Jornvall H, Persson B, Oppermann U (2008) *Cell Mol Life Sci* 65(24):3895
27. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H,

- Westbrook JD, Zardecki C (2002) *Acta Crystallographica Section D-Biological Crystallography* 58:899
28. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R (2006) *J Med Chem* 49(2):534
 29. Patel JR, Shuai Q, Dinges J, Winn M, Pliushchev M, Fung S, Monzon K, Chiou W, Wang J, Pan L, Wagaw S, Engstrom K, Kerdesky FA, Longenecker K, Judge R, Qin W, Imade HM, Stolarik D, Beno DWA, Brune M, Chovan LE, Sham HL, Jacobson P, Link JT (2007) *Bioorg Med Chem Lett* 17(3):750
 30. Arnold P, Tam S, Yan L, Baker ME, Frey FJ, Odermatt A (2003) *Mol Cell Endocrinol* 201(1-2):177
 31. White PC, Mune T, Agarwal AK (1997) *Endocr Rev* 18(1):135
 32. Penning TM (1997) *Endocr Rev* 18(3):281
 33. Kim KW, Wang Z, Busby J, Tsuruda T, Chen M, Hale C, Castro VM, Svensson S, Nybo R, Xiong F, Wang M (2006) *Biochim Biophys Acta* 1764(4):824
 34. Blum A, Favia AD, Maser E (2009) *Mol Cell Endocrinol* 301(1-2):132
 35. Lee JH, Kang NS, Yoo S-E (2008) *Bioorg Med Chem Lett* 18(7):2479
 36. Vicker N, Su X, Lawrence H, Cruttenden A, Purohit A, Reed MJ, Potter BVL (2004) *Bioorg Med Chem Lett* 14(12):3263
 37. Xiang J, Ipek M, Suri V, Tam M, Xing Y, Huang N, Zhang Y, Tobin J, Mansour TS, McKew J (2007) *Bioorg Med Chem* 15(13):4396
 38. Yeh VSC, Patel JR, Yong H, Kurukulasuriya R, Fung S, Monzon K, Chiou W, Wang J, Stolarik D, Imade H, Beno D, Brune M, Jacobson P, Sham H, Link JT (2006) *Bioorg Med Chem Lett* 16(20):5414
 39. Xiang J, Ipek M, Suri V, Masefski W, Pan N, Ge Y, Tam M, Xing Y, Tobin JF, Xu X, Tam S (2005) *Bioorg Med Chem Lett* 15(11):2865
 40. Tice CM, Zhao W, Xu Z, Cacatian ST, Simpson RD, Ye Y-J, Singh SB, McKeever BM, Lindblom P, Guo J, Krosky PM, Kruk BA, Berbaum J, Harrison RK, Johnson JJ,

- Bukhtiyarov Y, Panemangalore R, Scott BB, Zhao Y, Bruno JG, Zhuang L, McGeehan GM, He W, Claremon DA (2010) *Bioorg Med Chem Lett* 20(3):881
41. Potter BVL, Purohit A, Reed MJ, Vicker N (2002) Patent WO072084
42. Vicker N, Su X, Ganeshapillai D, Purohit A, Reed MJ, Potter BVL (2004) Patent WO037251
43. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) *J Med Chem* 47(7):1739
44. Muchmore SW, Souers AJ, Akritopoulou-Zanze I (2006) *Chem Biol Drug Des* 67(2):174
45. Naylor E, Arredouani A, Vasudevan SR, Lewis AM, Parkesh R, Mizote A, Rosen D, Thomas JM, Izumi M, Ganesan A, Galione A, Churchill GC (2009) *Nat Chem Biol* 5(4):220
46. Rush TS, 3rd, Grant JA, Mosyak L, Nicholls A (2005) *J Med Chem* 48(5):1489
47. Truchon JF, Bayly CI (2007) *J Chem Inf Model* 47(2):488
48. Kratschmar DV, Vuorinen A, Da Cunha T, Wolber G, Classen-Houben D, Doblhoff O, Schuster D, Odermatt A (2011) *J Steroid Biochem Mol Biol*
49. Tanaka N, Nonaka T, Tanabe T, Yoshimoto T, Tsuru D, Mitsui Y (1996) *Biochemistry* 35(24):7715
50. Oppermann UC, Filling C, Berndt KD, Persson B, Benach J, Ladenstein R, Jornvall H (1997) *Biochemistry* 36(1):34
51. Carvajal CA, Gonzalez AA, Romero DG, Gonzalez A, Mosso LM, Lagos ET, Hevia MdP, Rosati MP, Perez-Acle TO, Gomez-Sanchez CE, Montero JA, Fardella CE (2003) *J Clin Endocrinol Metab* 88(6):2501
52. Jornvall H, Persson B, Krook M, Atrian S, Gonzalez-Duarte R, Jeffery J, Ghosh D (1995) *Biochemistry* 34(18):6003
53. Oppermann U, Filling C, Hult M, Shafqat N, Wu X, Lindh M, Shafqat J, Nordling E, Kallberg Y, Persson B, Jörnvall H (2003) *Chem Biol Interact* 143-144:247

54. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) *J Med Chem* 47(7):1750
55. Huang N, Shoichet BK, Irwin JJ (2006) *J Med Chem* 49(23):6789
56. Schuster D, Maurer E, Laggner C, Nashev L, Wilckens T, Langer T, Odermatt A (2006) *J Med Chem* 49(12):3454
57. Vicker N, Su X, Lawrence H, Cruttenden A, Purohit A, Reed MJ, Potter BV (2004) *Bioorg Med Chem Lett* 14(12):3263
58. Su X, Pradaux-Caggiano F, Thomas MP, Szeto MWY, Halem HA, Culler MD, Vicker N, Potter BVL (2010) *ChemMedChem* 5(7):1026
59. Yang H, Dou W, Lou J, Leng Y, Shen J (2008) *Bioorg Med Chem Lett* 18(4):1340
60. Yoshikawa M, Murakami T, Matsuda H (1997) *Chem Pharm Bull (Tokyo)* 45(12):2034
61. Kim DH, Yu KW, Bae EA, Park HJ, Choi JW (1998) *Biol Pharm Bull* 21(4):360
62. Hollis G, Huber R (2011) *Diabetes Obes Metab* 13(1):1
63. Patel JR, Shuai Q, Link JT, Rohde JJ, Dinges J, Sorensen BK, Winn M, Yong H, Yeh VS (2006) Patent WO074244
64. Stanetty C, Czollner L, Koller I, Shah P, Gaware R, Cunha TD, Odermatt A, Jordis U, Kosma P, Classen-Houben D (2010) *Bioorg Med Chem* 18(21):7522
65. Farese S, Kruse A, Pasch A, Dick B, Frey BM, Uehlinger DE, Frey FJ (2009) *Kidney Int* 76(8):877
66. Zhang MZ, Xu J, Yao B, Yin H, Cai Q, Shrubsole MJ, Chen X, Kon V, Zheng W, Pozzi A, Harris RC (2009) *J Clin Invest* 119(4):876
67. Sala E, Guasch L, Iwaszkiewicz J, Mulero M, Salvadó MJ, Pinent M, Zoete V, Grosdidier A, Garcia-Vallvé S, Michielin O, Pujadas G (2011) *PLoS ONE* 6(2): e16903. doi:10.1371/journal.pone.0016903

SUPPORTING INFORMATION

Manuscript 4

A Structure-Based *In Silico* Methodology for Predicting 11beta-HSD2 Inhibitors

Esther Sala, Laura Guasch, Miquel Mulero, Cristina Valls, Maria-Josepa Salvadó,
Anna Arola-Arnal, Santiago Garcia-Vallvé, Gerard Pujadas.

Submitted needing revision to Journal of Computer-Aided Molecular Design

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

SUPPORTING INFORMATION

A Structure-Based *In Silico* Methodology for Predicting 11beta-HSD2 Inhibitors.

Esther Sala[§], Laura Guasch[§], Miquel Mulero[§], Cristina Valls[§], Maria-Josepa Salvadó[§], Anna Arola-Arnal[§], Santiago Garcia-Vallvé^{§,†}, Gerard Pujadas^{§,†*}.

[§]Grup de Recerca en Nutrigenòmica, Departament de Bioquímica i Biotecnologia, Universitat Rovira i Virgili, Campus de Sescelades, Tarragona, Catalonia, Spain

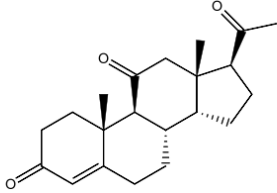
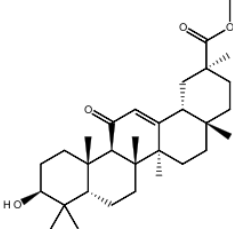
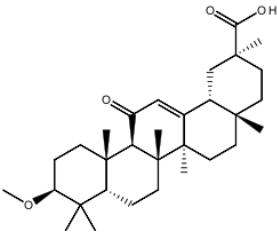
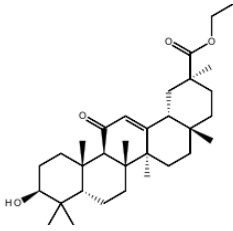
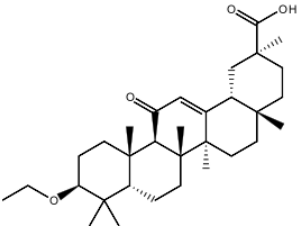
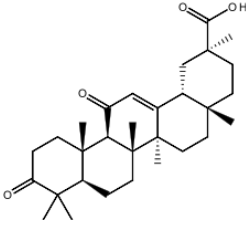
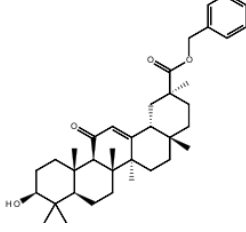
[†]Centre Tecnològic de Nutrició i Salut, Reus, Catalonia, Spain

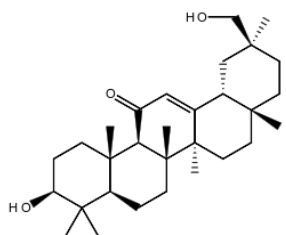
*e-mail: gerard.pujadas@urv.cat; Tel: +34977 559565; Fax: +34977558232

CONTENTS

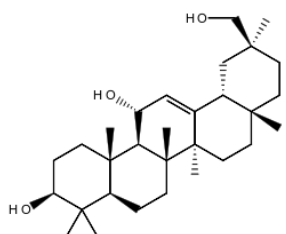
- Molecules in the *active* set. **S1**
- Molecules in the *non-active* set. **S2**
- Bibliography cited in Tables **S1** and **S2**

Table S1. Molecules in the *active set*.

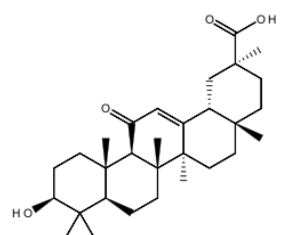
| 2D structure | Name | 11 β -HSD inhibition * | |
|---|-------------|------------------------------|--------|
| | | Type 1 | Type 2 |
|  | STX124 [1] | 100 | 92.6 |
|  | STX195a [1] | 53 | 90.3 |
|  | STX195 [1] | 85.4 | 105.2 |
|  | STX196a [1] | 55 | 93.9 |
|  | STX196 [1] | 80.6 | 97.5 |
|  | STX347 [1] | 89.3 | 102.2 |
|  | STX348 [1] | 63.3 | 94.7 |



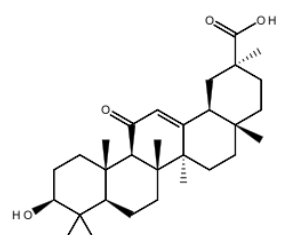
STX349 [1] 89.6 100



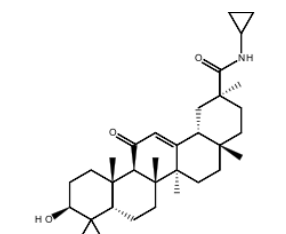
STX351 [1] 70 94.4



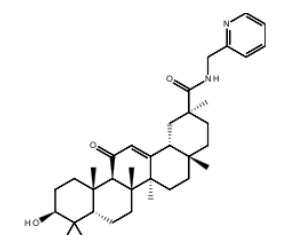
STX352 [1] 85.1 101



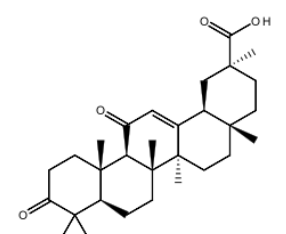
STX353 [1] 89 100.4



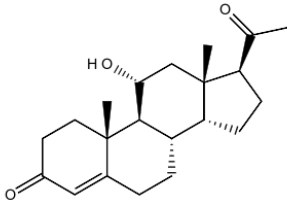
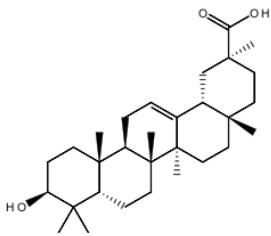
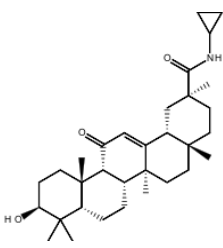
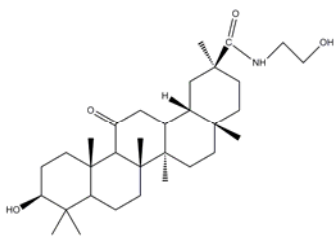
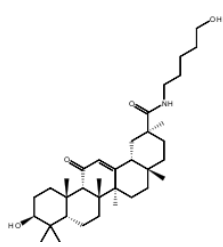
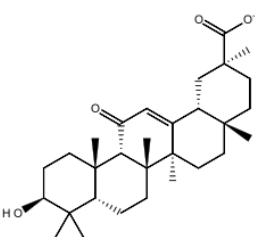
STX370 [1] 89.6 102.5



STX371 [1] 57.1 92.9

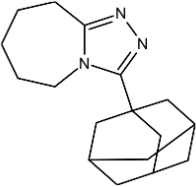
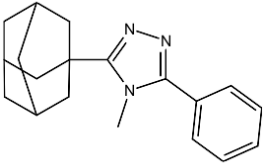
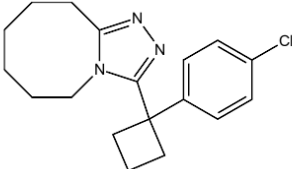
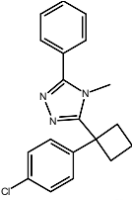
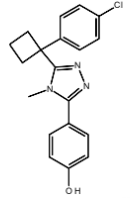
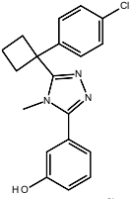
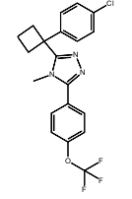
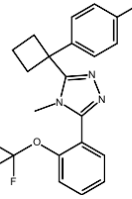


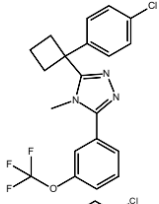
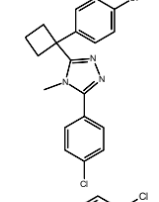
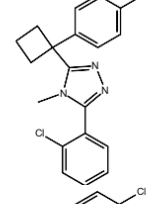
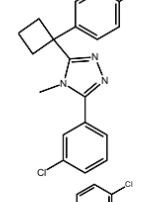
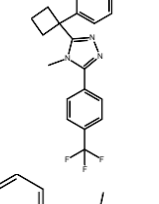
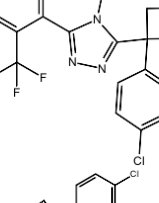
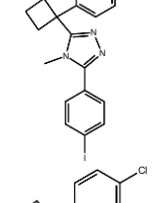
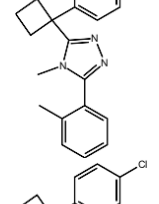
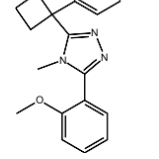
BLE99005 [1] 86.9 100

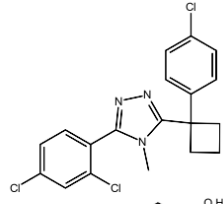
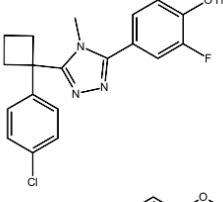
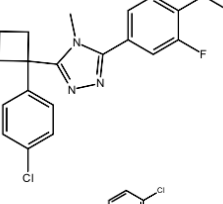
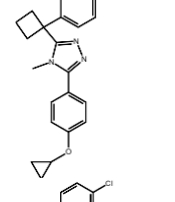
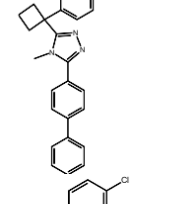
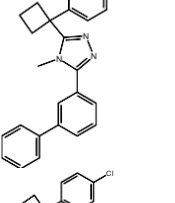
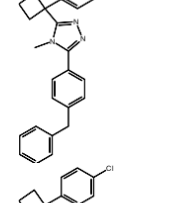
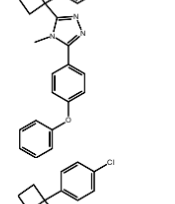
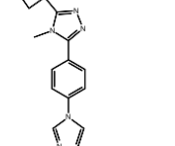
| | | | |
|---|-------------------|------|------|
|  | DG322B [1] | 32.6 | 83.8 |
|  | DG381A [1] | 92.6 | 109 |
|  | 3 [2] | 90 | 103 |
|  | 5 [2] | 36 | 92 |
|  | 6 [2] | 88 | 98 |
|  | Cpd01 [3] | 12 | 8 |

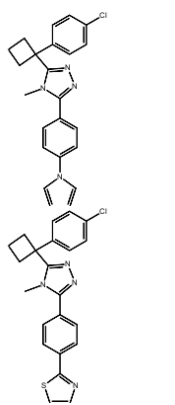
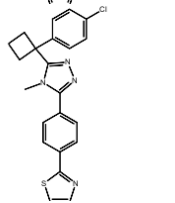
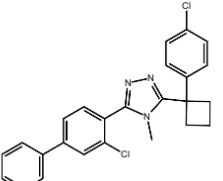
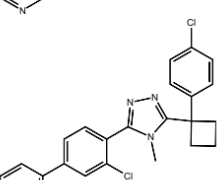
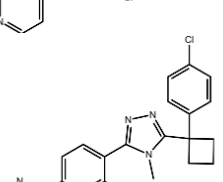
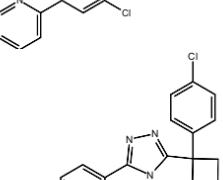
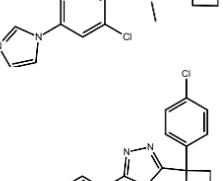
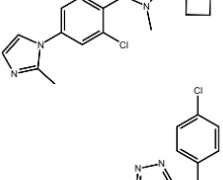
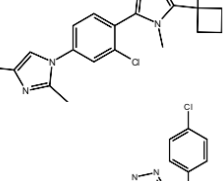
* The inhibitory activity is expressed either as a percent of inhibition [1,2] (by using a ligand concentration of 10 μ M) or IC_{50} (in nM) [3].

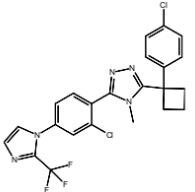
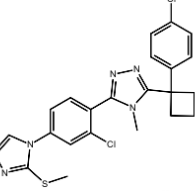
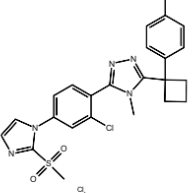
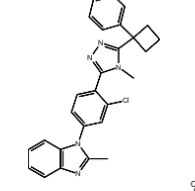
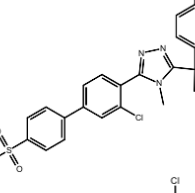
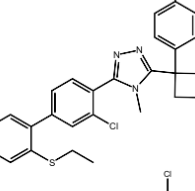
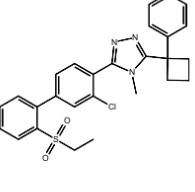
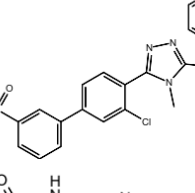
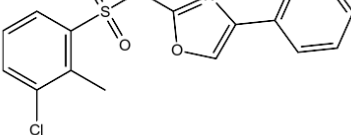
Table S2. Molecules in the *non-active* set.

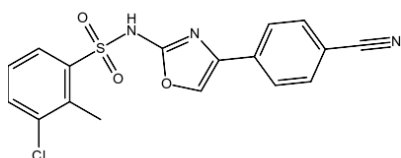
| 2D structure | Name | 11 β -HSD inhibition* | |
|---|--------------|-----------------------------|--------------------|
| | | Type 1 | Type 2 |
|  | 1 [4] | 7.8 | >3*10 ³ |
|  | 2 [4] | 37 | >4*10 ³ |
|  | 3 [4] | 1.3 | >4*10 ³ |
|  | 4 [4] | 1.7 | >4*10 ³ |
|  | 5 [4] | 0.34 | >4*10 ³ |
|  | 7 [4] | 1 | >4*10 ³ |
|  | 8 [4] | 13 | >4*10 ³ |
|  | 9 [4] | 2.3 | >4*10 ³ |

| | | | |
|---|---------------|------|-----------------|
|  | 10 [4] | 2.2 | $>4 \cdot 10^3$ |
|  | 11 [4] | 2.7 | $>4 \cdot 10^3$ |
|  | 12 [4] | 0.28 | $>4 \cdot 10^3$ |
|  | 13 [4] | 1 | $>4 \cdot 10^3$ |
|  | 17 [4] | 9.3 | $>4 \cdot 10^3$ |
|  | 18 [4] | 2 | $>4 \cdot 10^3$ |
|  | 19 [4] | 3.8 | $>4 \cdot 10^3$ |
|  | 20 [4] | 5 | $>4 \cdot 10^3$ |
|  | 21 [4] | 2.7 | $>4 \cdot 10^3$ |

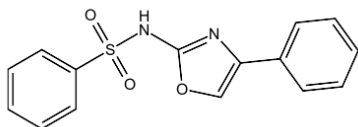
| | | | |
|---|---------------|-----|-----------------|
|  | 22 [4] | 1 | $>4 \cdot 10^3$ |
|  | 23 [4] | 2 | $>4 \cdot 10^3$ |
|  | 24 [4] | 3 | $>4 \cdot 10^3$ |
|  | 25 [4] | 10 | $>4 \cdot 10^3$ |
|  | 26 [4] | 11 | $>4 \cdot 10^3$ |
|  | 27 [4] | 4.3 | $>4 \cdot 10^3$ |
|  | 29 [4] | 68 | $>4 \cdot 10^3$ |
|  | 31 [4] | 26 | $>4 \cdot 10^3$ |
|  | 32 [4] | 3.5 | $>4 \cdot 10^3$ |

| | | | |
|---|---------------|-----|-----------------|
|  | 33 [4] | 3 | $>4 \cdot 10^3$ |
|  | 34 [4] | 44 | $>4 \cdot 10^3$ |
|  | 36 [4] | 1 | $>4 \cdot 10^3$ |
|  | 37 [4] | 1 | $>4 \cdot 10^3$ |
|  | 38 [4] | 1 | $>4 \cdot 10^3$ |
|  | 39 [4] | 1 | $>4 \cdot 10^3$ |
|  | 40 [4] | 2.6 | $>4 \cdot 10^3$ |
|  | 41 [4] | 6.9 | $>4 \cdot 10^3$ |
|  | 42 [4] | 9.8 | $>4 \cdot 10^3$ |

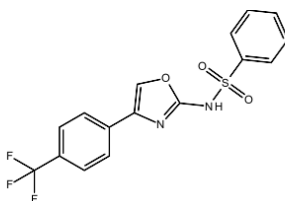
| | | | |
|---|----------------|-------------------|-----------------|
|  | 43 [4] | 0.98 | $>4 \cdot 10^3$ |
|  | 44 [4] | 3.9 | $>4 \cdot 10^3$ |
|  | 45 [4] | 14 | $>4 \cdot 10^3$ |
|  | 46 [4] | 12 | $>4 \cdot 10^3$ |
|  | 47 [4] | 2.1 | $>4 \cdot 10^3$ |
|  | 48 [4] | 11 | $>4 \cdot 10^3$ |
|  | 49 [4] | 0.98 | $>4 \cdot 10^3$ |
|  | 50 [4] | 0.98 | $>4 \cdot 10^3$ |
|  | 11D [5] | $17.3 \cdot 10^2$ | $>2 \cdot 10^5$ |



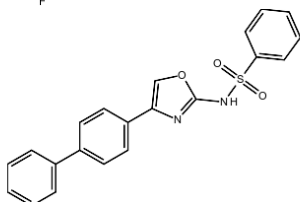
11E [5] $48 \cdot 10^3$ $>2 \cdot 10^5$



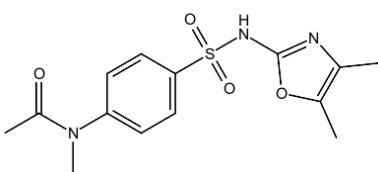
11F [5] $23 \cdot 10^3$ $>2 \cdot 10^5$



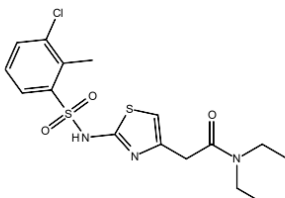
11G [5] $9 \cdot 10^3$ $>2 \cdot 10^5$



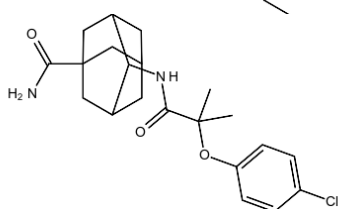
11H [5] $2.3 \cdot 10^2$ $>2 \cdot 10^5$



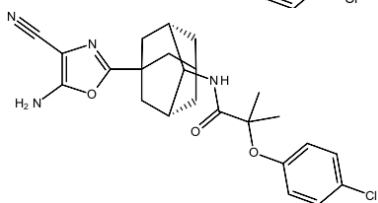
12 [5] $1 \cdot 10^5$ $>2 \cdot 10^5$



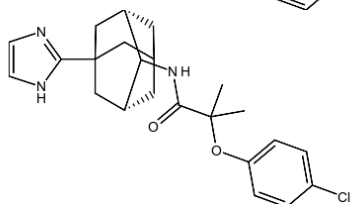
BVT-14225 [5] 1500 $>2 \cdot 10^5$



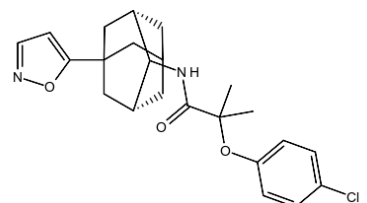
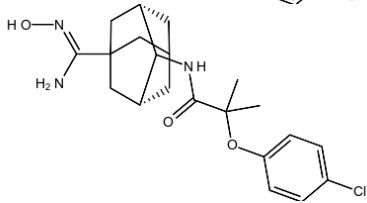
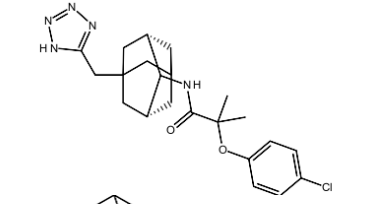
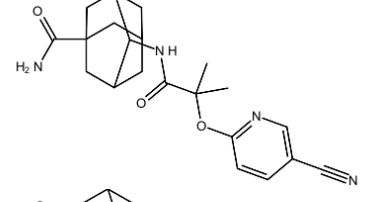
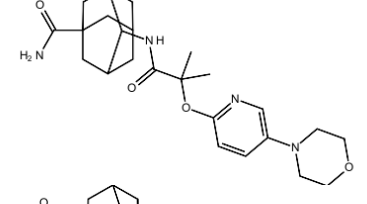
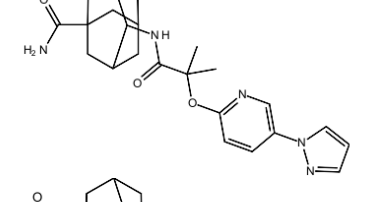
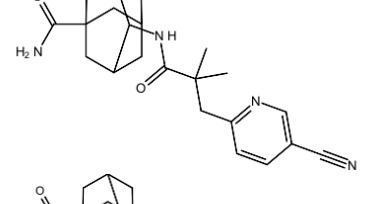
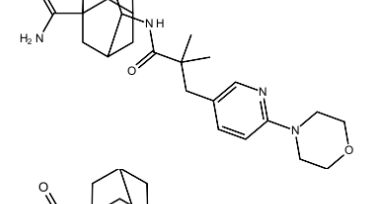
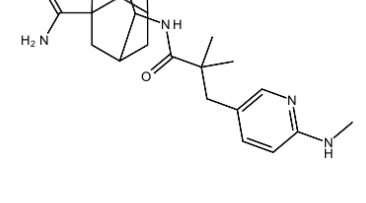
1 [6] 32 $>1 \cdot 10^5$

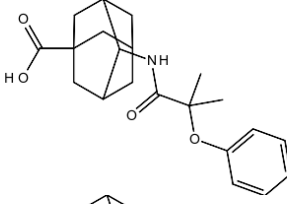
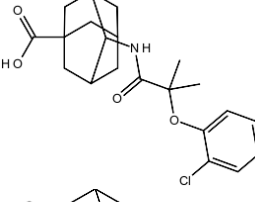
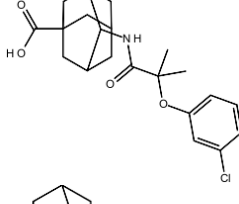
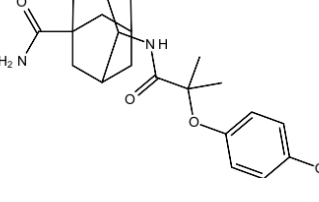
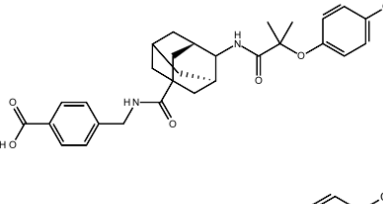
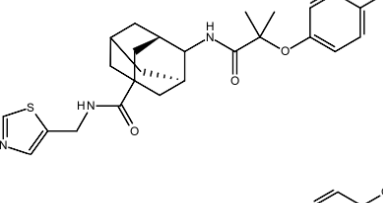
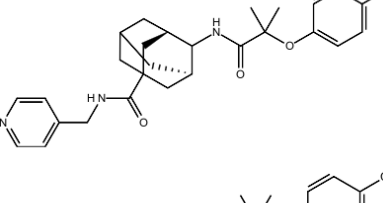
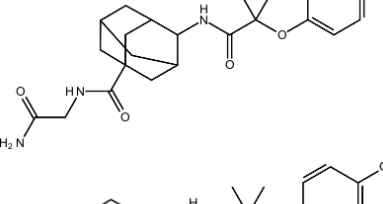
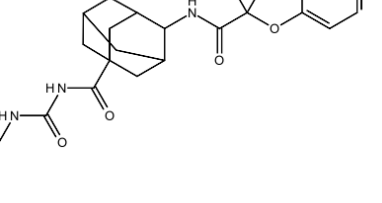


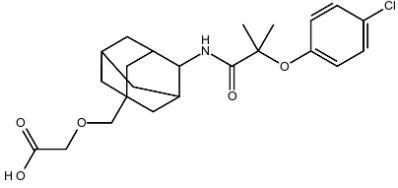
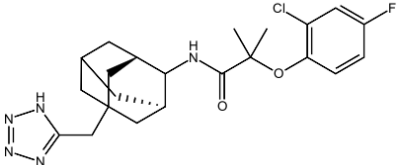
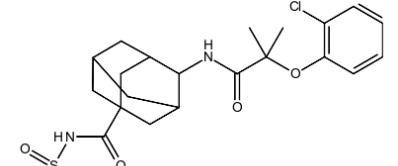
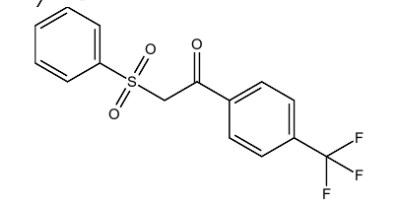
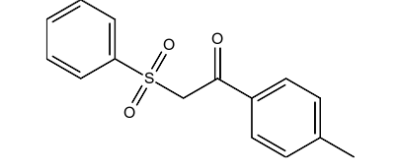
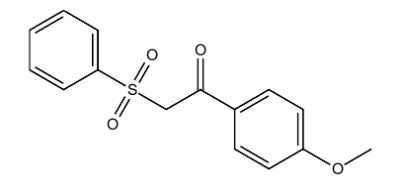
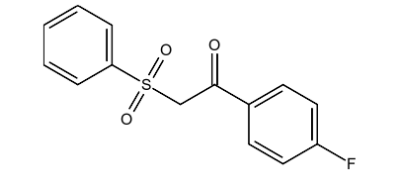
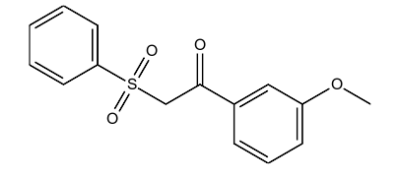
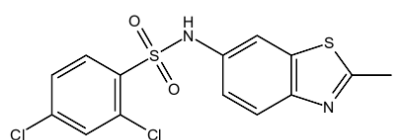
7 [6] 146 $>1 \cdot 10^5$

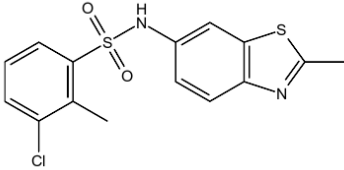
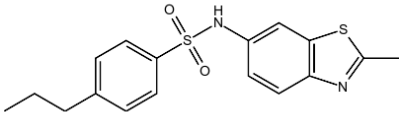
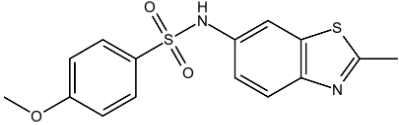
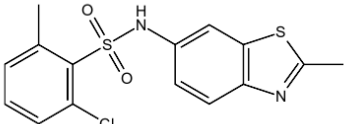
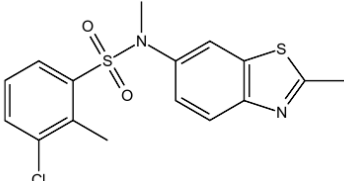
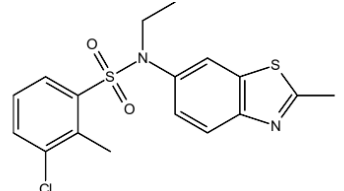
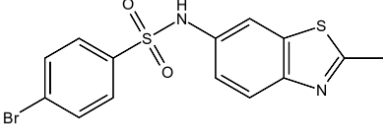
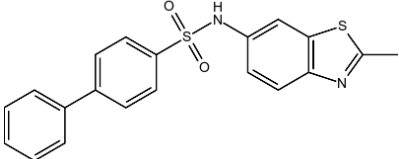
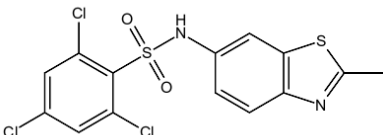


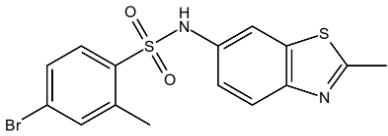
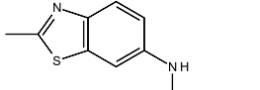
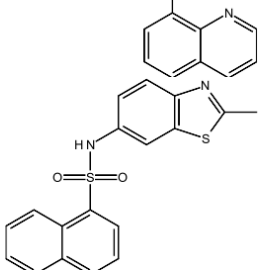
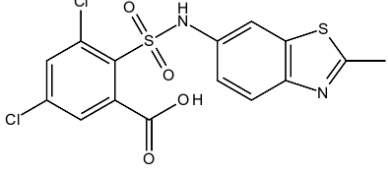
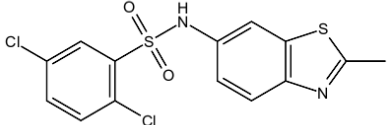
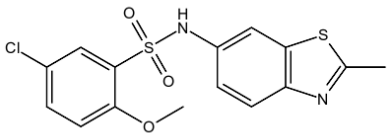
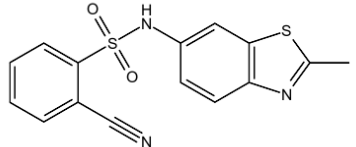
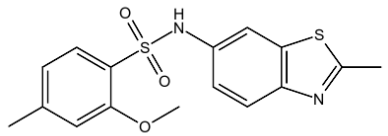
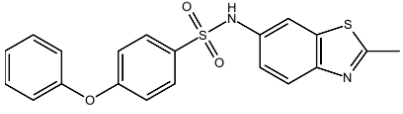
9 [6] 10 $>11 \cdot 10^3$

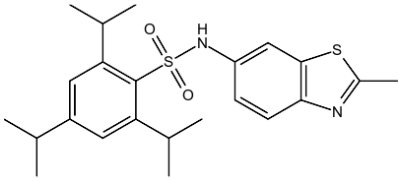
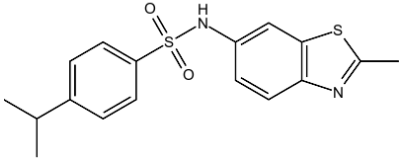
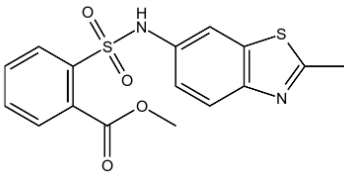
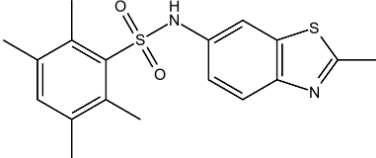
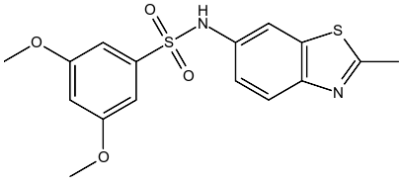
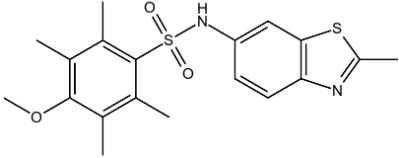
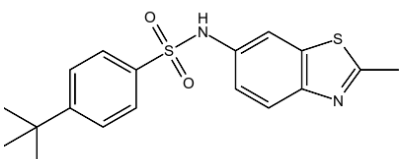
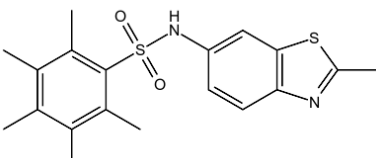
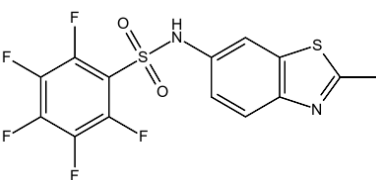
| | | | |
|---|---------------|-----|------------------|
|  | 10 [6] | 105 | $>15 \cdot 10^3$ |
|  | 11 [6] | 39 | $>1 \cdot 10^5$ |
|  | 14 [6] | 41 | $>1 \cdot 10^5$ |
|  | 23 [6] | 28 | $>1 \cdot 10^5$ |
|  | 24 [6] | 44 | $>1 \cdot 10^5$ |
|  | 26 [6] | 5 | $>1 \cdot 10^5$ |
|  | 27 [6] | 14 | $>1 \cdot 10^5$ |
|  | 28 [6] | 47 | $>1 \cdot 10^5$ |
|  | 29 [6] | 29 | $>1 \cdot 10^5$ |

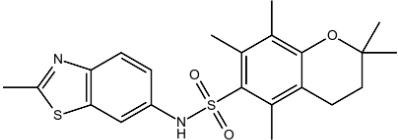
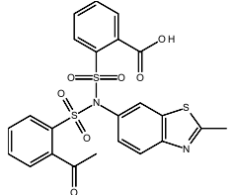
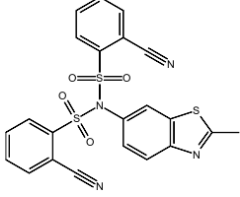
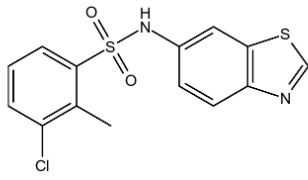
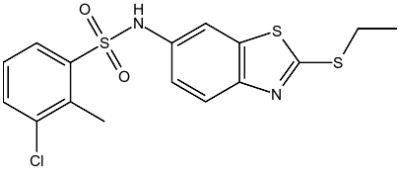
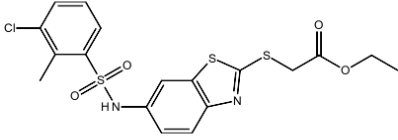
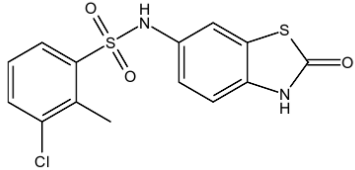
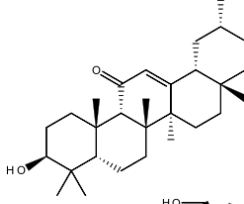
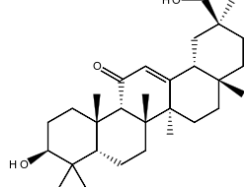
| | | | |
|---|---------------|----|-----------------|
|  | 5 [7] | 39 | $17 \cdot 10^3$ |
|  | 7 [7] | 8 | 1800 |
|  | 8 [7] | 15 | 3500 |
|  | 10 [7] | 6 | $12 \cdot 10^3$ |
|  | 16 [7] | 15 | $>2 \cdot 10^4$ |
|  | 20 [7] | 27 | $13 \cdot 10^3$ |
|  | 21 [7] | 25 | 5300 |
|  | 22 [7] | 12 | $2 \cdot 10^4$ |
|  | 23 [7] | 13 | $12 \cdot 10^3$ |

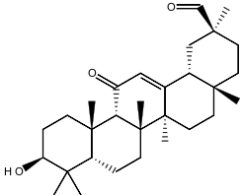
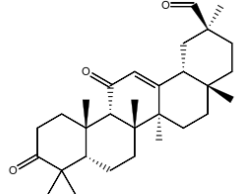
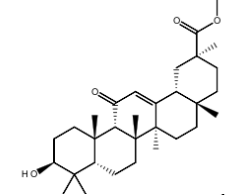
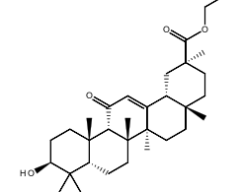
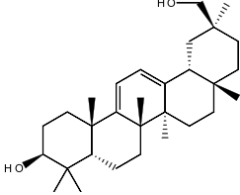
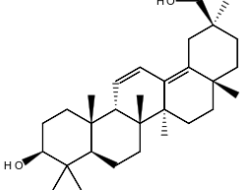
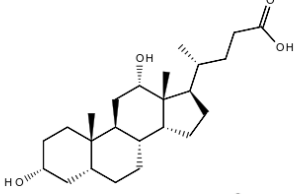
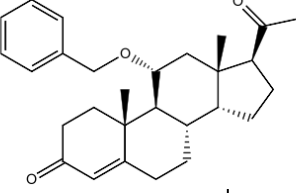
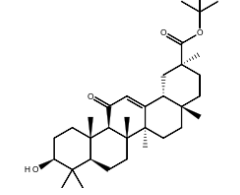
| | | | |
|---|----------------|-----|-----------------|
|  | 25 [7] | 110 | $>1 \cdot 10^5$ |
|  | 26 [7] | 11 | 550 |
|  | 27 [7] | 96 | 580 |
|  | 2a [8] | 187 | $>2 \cdot 10^5$ |
|  | 2b [8] | 102 | $>2 \cdot 10^5$ |
|  | 2c [8] | 75 | $>2 \cdot 10^5$ |
|  | 2d [8] | 90 | $>2 \cdot 10^5$ |
|  | 2f [8] | 60 | $>2 \cdot 10^5$ |
|  | 412 [9] | 27 | 3 |

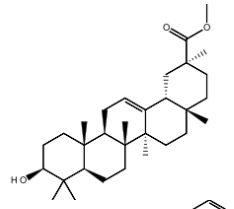
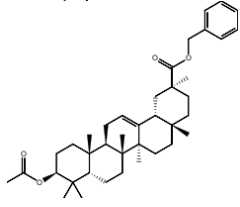
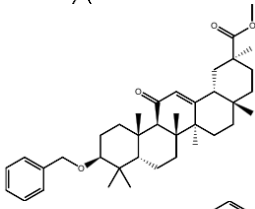
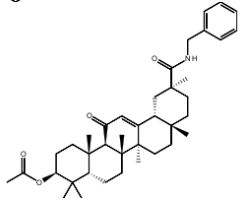
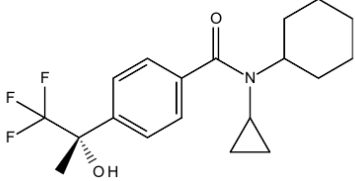
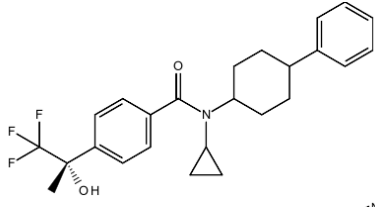
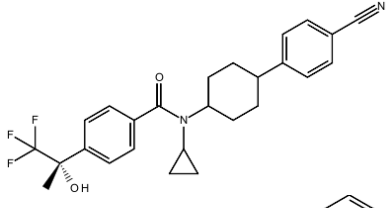
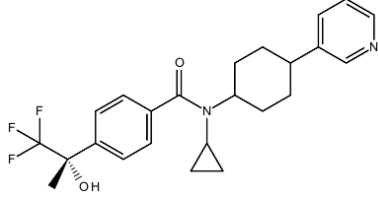
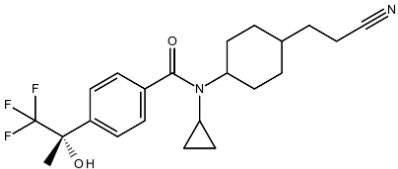
| | | | |
|---|----------------|-----|-----|
|  | 413 [9] | 53 | 0.2 |
|  | 421 [9] | 60 | 0.9 |
|  | 424 [9] | 24 | 0.7 |
|  | 425 [9] | 40 | 0 |
|  | 469 [9] | 63 | 29 |
|  | 470 [9] | 39 | 30 |
|  | 519 [9] | 48 | 8 |
|  | 521 [9] | 0.5 | 5 |
|  | 522 [9] | 37 | 6 |

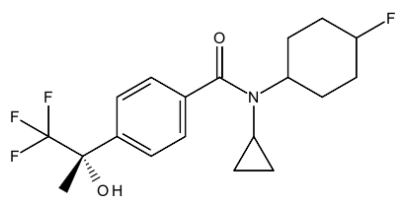
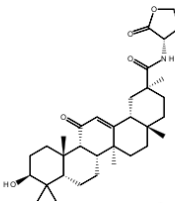
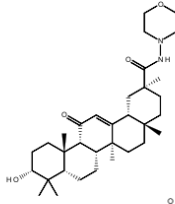
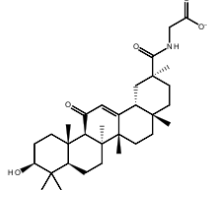
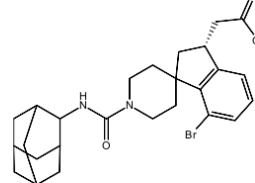
| | | | |
|---|----------------|-----|-----|
|  | 523 [9] | 21 | 8 |
|  | 552 [9] | 18 | 24 |
|  | 553 [9] | 0.7 | 18 |
|  | 575 [9] | 62 | 1.6 |
|  | 580 [9] | 75 | 1.4 |
|  | 582 [9] | 40 | 0.7 |
|  | 583 [9] | 29 | 0.4 |
|  | 584 [9] | 48 | 10 |
|  | 585 [9] | 48 | 1.6 |

| | | | |
|---|----------------|----|----|
|  | 703 [9] | 35 | 4 |
|  | 704 [9] | 38 | 4 |
|  | 705 [9] | 6 | 6 |
|  | 706 [9] | 29 | 7 |
|  | 707 [9] | 21 | 11 |
|  | 708 [9] | 39 | 11 |
|  | 709 [9] | 10 | 13 |
|  | 710 [9] | 55 | 10 |
|  | 712 [9] | 24 | 3 |

| | | | |
|---|------------------|-----|--------------------|
|  | 713 [9] | 26 | 3 |
|  | 730 [9] | 32 | 9 |
|  | 731 [9] | 45 | 12 |
|  | 750 [9] | 4 | 10 |
|  | 751 [9] | 10 | 5 |
|  | 752 [9] | 5 | 1 |
|  | 753 [9] | 8 | 2 |
|  | Cpd04 [3] | 244 | 12*10 ³ |
|  | Cpd05 [3] | 149 | 16*10 ³ |

| | | | |
|---|-----------------------------|------|-----------------|
|  | Cpd06 [3] | 122 | $25 \cdot 10^3$ |
|  | Cpd07 [3] | 124 | $16 \cdot 10^3$ |
|  | Cpd08 [3] | 119 | $>2 \cdot 10^5$ |
|  | Cpd09 [3] | 141 | $>2 \cdot 10^5$ |
|  | Cpd10 [3] | 39 | $>2 \cdot 10^5$ |
|  | Cpd11 [3] | 37 | $>2 \cdot 10^5$ |
|  | Deoxycholic acid [1] | 27.6 | -1.1 |
|  | STX185 [1] | 16.1 | 18.9 |
|  | STX298 [1] | 50.5 | 22.5 |

| | | | |
|---|-------------------|------|-----------------|
|  | STX299 [1] | 20.2 | 19.5 |
|  | STX354 [1] | -1.1 | 13.6 |
|  | STX359 [1] | 17.8 | 17.0 |
|  | STX367 [1] | 36.1 | 27.2 |
|  | 9 [10] | 0.8 | $>1 \cdot 10^4$ |
|  | 10 [10] | 1.3 | $>1 \cdot 10^4$ |
|  | 12 [10] | 0.7 | $>1 \cdot 10^4$ |
|  | 13 [10] | 1.4 | $>1 \cdot 10^4$ |
|  | 14 [10] | 0.7 | $>1 \cdot 10^4$ |

| | | | |
|--|------------------|-----|-----------------|
|  | 15 [10] | 6.6 | $>1 \cdot 10^4$ |
|  | 7 [2] | 71 | 25 |
|  | 9 [2] | 34 | 20 |
|  | 26i [11] | 2.7 | 21.4 |
|  | 10j2 [12] | 1.1 | $>1 \cdot 10^4$ |

* The inhibitory activity is expressed either as IC_{50} [3-8,10,12] (in nM) or a percent of inhibition [1,2,9,11] [by using a ligand concentration of 10 μ M (except for **26i** [11], for which 1 μ M was used)].

References

1. Potter BVL, Purohit A, Reed MJ, Vicker N (2002) Patent WO072084
2. Vicker N, Su X, Lawrence H, Cruttenden A, Purohit A, Reed MJ, Potter BV (2004) *Bioorg Med Chem Lett* 14(12):3263
3. Blum A, Favia AD, Maser E (2009) *Mol Cell Endocrinol* 301(1-2):132
4. Zhu Y, Olson SH, Hermanowski-Vosatka A, Mundt S, Shah K, Springer M, Thieringer R, Wright S, Xiao J, Zokian H, Balkovec JM (2008) *Bioorg Med Chem Lett* 18(11):3405
5. Xiang J, Ipek M, Suri V, Masefski W, Pan N, Ge Y, Tam M, Xing Y, Tobin JF, Xu X, Tam S (2005) *Bioorg Med Chem Lett* 15(11):2865
6. Yeh VSC, Patel JR, Yong H, Kurukulasuriya R, Fung S, Monzon K, Chiou W, Wang J, Stolarik D, Imade H, Beno D, Brune M, Jacobson P, Sham H, Link JT (2006) *Bioorg Med Chem Lett* 16(20):5414
7. Patel JR, Shuai Q, Dinges J, Winn M, Pliushchev M, Fung S, Monzon K, Chiou W, Wang J, Pan L, Wagaw S, Engstrom K, Kerdesky FA, Longenecker K, Judge R, Qin W, Imade HM, Stolarik D, Beno DWA, Brune M, Chovan LE, Sham HL, Jacobson P, Link JT (2007) *Bioorg Med Chem Lett* 17(3):750
8. Xiang J, Ipek M, Suri V, Tam M, Xing Y, Huang N, Zhang Y, Tobin J, Mansour TS, McKew J (2007) *Bioorg Med Chem* 15(13):4396
9. Vicker N, Su X, Ganeshpillai D, Purohit A, Reed MJ, Potter BVL (2004) Patent WO037251
10. Julian LD, Wang Z, Bostick T, Caille S, Choi R, DeGraffenreid M, Di Y, He X, Hungate RW, Jaen JC, Liu J, Monshouwer M, McMinn D, Rew Y, Sudom A, Sun D, Tu H, Ursu S, Walker N, Yan X, Ye Q, Powers JP (2008) *J Med Chem* 51(13):3953
11. Beseda I, Czollner L, Shah PS, Khunt R, Gaware R, Kosma P, Stanetty C, Del Ruiz-Ruiz MC, Amer H, Mereiter K, Da Cunha T, Odermatt A, Classen-Houben D, Jordis U (2010) *Bioorg Med Chem* 18(1):433
12. Tice CM, Zhao W, Xu Z, Cacatian ST, Simpson RD, Ye Y-J, Singh SB, McKeever BM, Lindblom P, Guo J, Krosky PM, Kruk BA, Berbaum J, Harrison RK, Johnson JJ, Bukhtiyarov Y, Panemangalore R, Scott BB, Zhao Y, Bruno JG, Zhuang L, McGeehan GM, He W, Claremon DA (2010) *Bioorg Med Chem Lett* 20(3):881

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

Manuscript 5

In silico prediction of selective 11 β -HSD1 inhibitors
of Natural Origin

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

***In silico* prediction of selective 11beta-HSD1 inhibitors of natural origin**

Esther Sala[§], Laura Guasch[§], Eloy Rodríguez-Freire[§], Miquel Mulero[§], Cristina Valls[§], Anna Arola-Arnal[§], Lluís Arola^{§,†}, Santiago Garcia-Vallvé^{§,†}, Gerard Pujadas^{§,†*}.

[§]Grup de Recerca en Nutrigenòmica, Departament de Bioquímica i Biotecnologia, Universitat Rovira i Virgili, Campus de Sescelades, Tarragona, Catalonia, Spain

[†]Centre Tecnològic de Nutrició i Salut, Reus, Catalonia, Spain

*e-mail: gerard.pujadas@urv.cat; Tel: +34977 559565; Fax: +34977 558232

Abstract

Glucocorticoids are ubiquitous hormones that play a key role in modulating immune and inflammatory responses and in regulating energy metabolism, cardiovascular homeostasis and the body's responses to stress. The principal glucocorticoid is cortisol, which has a concentration in target tissues that is modulated by 11 β -HSD1 (which reduces cortisone to cortisol) and by 11 β -HSD2 (which converts cortisol to cortisone). A growing body of evidence suggests that increased 11 β -HSD1 activity within target tissues may promote insulin resistance, obesity, hypertension and dyslipidemia. Therefore, the search for 11 β -HSD1 inhibitors has been an active area of study in recent years. Unfortunately, there are many examples of 11 β -HSD1 inhibitors that cause side effects, such as sodium retention and hypertension, because they also inhibit 11 β -HSD2. Consequently, the increasing interest in detecting selective 11 β -HSD1 inhibitors for the therapeutic treatment of glucocorticoid-dependent diseases has encouraged us to develop a virtual screening protocol for finding such inhibitors in natural products databases. Our study was able to predict that 11 of the 89,165 compounds in the starting database have a good chance of being selective 11 β -HSD1 inhibitors. Interestingly, these compounds correspond to 5 different chemical scaffolds, and only one of them, which is shared by 7 out of 11 virtual screening hits, has been reported in the development of selective 11 β -HSD1 inhibitors by structure-activity relationship studies. Therefore, the remaining 4 hits (each from a different scaffold) can provide new lead compounds for deriving new 11 β -HSD1 inhibitors through structure-activity studies.

Keywords

Metabolic syndrome, structure-based pharmacophore search, protein-ligand docking, 11beta-hydroxysteroid dehydrogenase type 1, 11beta-hydroxysteroid dehydrogenase type 2.

Abbreviations

11 β -HSD1: 11 β -hydroxysteroid dehydrogenase type 1; 11 β -HSD2: 11 β -hydroxysteroid dehydrogenase type 2; GCs: glucocorticoids; NPs: Natural products; PDB: Protein Data Bank; SDR: short-chain dehydrogenase/reductase; VS: virtual screening.

Introduction

Natural products (NPs) are ideal starting points for drug design and development because of their large scaffold diversity and properties, such as structural complexity and drug similarity claims [1-4]. In fact, over 60% of the current anticancer drugs are natural product related molecules (*i.e.*, either NPs or derivatives/analogues that have been inspired by a natural compound) [1]. Despite these clear advantages, it is estimated that only 5 to 15% of the approximately 250,000 described higher plant species have been tested for some type of biological activity [5], and the use of other organisms, such as micro-organisms, insects, fungi or marine species, for this goal is just beginning [6,7]. It has recently been suggested that bioinformatics/chemoinformatics tools could be used to screen large NP databases and to identify new bioactive molecules for specific targets [3,5,8].

One target of interest for NPs is the 11 β -hydroxysteroid dehydrogenase type 1 (11 β -HSD1), a NADPH-dependent enzyme that is predominantly expressed in liver, adipose tissue and skeletal muscles, where it increases the intracellular glucocorticoid (GC) action by catalyzing the reduction of inactive 11-ketoglucocorticoids (*i.e.*, cortisone in humans) to active 11 β -hydroxyglucocorticoids (*i.e.*, cortisol in humans) [9-12]. The opposite reaction is catalyzed by 11 β -hydroxysteroid dehydrogenase type 2 (11 β -HSD2), which is a NAD⁺-dependent enzyme that is highly expressed in classical aldosterone selective target tissues, like distal nephron, colon, sweat glands and the placenta, where it protects the mineralocorticoid receptor from active 11 β -hydroxyglucocorticoids excess [13,14]. GCs modulate the expression of up to 20% of the genes in the mammalian genome [15], and their concentration in target tissues is tightly regulated by 11 β -HSD1, 11 β -HSD2 and GC receptors [16]. The normalization of GC levels can reverse metabolic syndrome by improving obesity, insulin resistance, hypertension and the lipid and lipoprotein profiles [17,18].

A growing body of evidence suggests that increased 11 β -HSD1 activity within target tissues may promote insulin resistance, obesity, hypertension and dyslipidemia. Therefore, looking for potent and selective 11 β -HSD1 inhibitors has been an active area of study in recent years in both academia and in the pharmaceutical industry [9,10,18-42]. Unfortunately, there are many of examples of 11 β -HSD1 inhibitors that cause side effects like sodium retention and hypertension because these inhibitors also inhibit 11 β -HSD2 [19,38,39,43,44]. Thus, one important task when finding new 11 β -HSD1 inhibitors is the evaluation of their activity on 11 β -HSD2 [13,23,29,40-42,45].

The human 11 β -HSD1 is a protein that belongs to the classical short dehydrogenase/reductase family (SDR) [46-50] and possesses a sequence length of 291 residues. It is an endoplasmic reticulum-membrane protein with luminal orientation that contains an N-terminal membrane anchor that is part of a single transmembrane helix (residues 8 to 24). The rest of the C-terminal segment (residues 25 to 291) corresponds to the catalytic domain, containing the relevant sequence signature and active site motifs of the SDR superfamily and protruding into the endoplasmic reticulum lumen [51,52]. At present, there are 16 different entries at the Protein Data Bank (PDB; <http://www.pdb.org>)[53] that correspond to complexes between the 11 β -HSD1 catalytic domain and synthetic inhibitors (see Table 1).

The increasing interest in detecting specific 11 β -HSD1 inhibitors for the therapeutic treatment of GC-dependent diseases has encouraged us to develop a virtual screening (VS) protocol for finding such inhibitors in a NP database. To achieve this goal, we have (a) used the corresponding electron density maps to assess the reliability of ligand and binding-site coordinates in the PDB complexes between 11 β -HSD1 and 11 β -HSD1 inhibitors; (b) used the reliable complexes to derive structure-based common pharmacophores that contain the key intermolecular interactions between 11 β -HSD1 and different inhibitor scaffolds; (c) added to each pharmacophore a shell of excluded volumes that schematically represents the locations of the residues forming the ligand-binding site; (d) used these structure-based common pharmacophores as part of a VS protocol that has been applied to a NP database and that also includes an ADME/Tox filter, ligand-pose generation through protein-ligand docking, a shape and electrostatic comparison with experimental poses of 11 β -HSD1 inhibitors and a filter to discard 11 β -HSD2 inhibitors. The results of our study show that the VS protocol is able to predict new selective 11 β -HSD1 inhibitors that belong to chemical families that are different from the known 11 β -HSD1 inhibitors (including those used to build the structure-based pharmacophores). These VS hits suggest new lead compounds for deriving more potent and selective 11 β -HSD1 inhibitors through structure-activity studies.

Materials and computational methods

Validation of the reliability of the intermolecular interactions between 11 β -HSD1 and co-crystallized inhibitors

The reliability of the intermolecular interactions between 11 β -HSD1 and co-crystallized inhibitors in PDB files was assessed by a visual inspection of the fit of the inhibitors and their surrounding residues (only residues with at least one atom at a distance ≤ 4.0 Å were considered) with the corresponding section of the electron density map at a contour level of 1.00 σ . This comparison

was performed with the help of the resources at the Electron Density Server (<http://eds.bmc.uu.se/eds/>) [54]. Moreover, special care was taken to ensure that all of the residues that are reported to interact with these inhibitors at the LigPlot diagrams [55] of the PDBsum webserver (<http://www.ebi.ac.uk/pdbsum/>) [56] were included in this task.

Common structure-based pharmacophore generation

Those complexes with no ambiguities in the coordinates of the set formed by the inhibitor and its ligand-binding site (see Table 1) were superimposed using DeepView v4.0 (<http://spdbv.vital-it.ch/>) [57]. Then, each superimposed complex was used by LigandScout v2.00 [58] (Inte:Ligand Softwareentwicklungs- und Consulting GmbH, Vienna, Austria; <http://www.inteligand.com/ligandscout/>) to derive its structure-based pharmacophore. The resulting structure-based pharmacophores were further analyzed to find common patterns of intermolecular interactions that could show the minimum functional-group requirements that are needed to inhibit 11 β -HSD1. Moreover, this analysis provided the framework to (a) consider sites either as mandatory or optional (*i.e.*, a site was considered as mandatory when it was common to most of the ligands that contribute to the same common pharmacophore); and (b) set up the tolerance of the different sites' locations (see Table 2). The resulting common structure-based pharmacophores were completed with excluded volumes calculated with the *Receptor-Based Excluded Volumes* graphic front-end from Phase v3.1 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) and by setting the *Sphere filters* parameter values to **(a) ignore receptor atoms whose surfaces were within 0.25 Å of ligand surface**; and **(b) limit the excluded volume shell thickness to 10 Å**. The rest of parameter values used in that front-end were the default settings.

Virtual-screening workflow description

Initially, the 3D structures of the ligands used in this work were downloaded from the ZINC Natural Products Database (http://wiki.compbio.ucsf.edu/wiki/index.php/Natural_products_database) [59]. Then, all of these molecules were submitted to an ADME/Tox filter with the FAF-Drugs2 tool [60], which aimed to discard molecules that either had poor ADME properties or were potentially toxic. Thus, the drug-like properties of a compound were evaluated using the Lipinski rule [61], and only one violation of this rule was allowed. This rule is based on a set of property values (*i.e.*, the number of hydrogen-bond donors and acceptors, the molecular weight and the logP) that were derived from drugs with good ADME characteristics [61]. Therefore, molecules that pass the Lipinski rule are expected to be active in humans after oral admission. Moreover, molecules containing toxic groups were filtered using the 204 substructures

for “warhead” chelators, frequent hitters, promiscuous inhibitors and other undesirable functional groups available in the FAF-Drugs2 tool [60].

Molecules with appropriate ADME/Tox properties were then set up with LigPrep v2.3 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) and were improved by cleaning. The cleaning process was carried out with the following parameters: **(a)** the force field used was OPLS 2005; **(b)** all possible ionization states at pH 7.0 \pm 2.0 were generated with Ionizer; **(c)** the *desalt* option was activated; **(d)** tautomers were generated for all ionization states at pH 7.0 \pm 2.0; **(e)** chiralities, when present, were determined from the 3D structure; and **(f)** one low-energy ring conformation per ligand was generated. Conformations and sites for the resulting ligand structures were determined during the generation of the corresponding Phase databases with the *Generate Phase Database* graphic front-end. The parameter values used during this *in vacuo* conformer generation were the default settings, with the exception of the maximum number of conformers per structure, which was increased from 100 (the default value) to 200. The conformer sites were generated with definitions made by adding the ability to consider aromatic rings as hydrophobic groups to the default built-in Phase definitions. Next, this database was filtered with Phase v3.1 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) [62] through each common pharmacophore/environment combination by using the following running conditions: **(a)** search in the conformers database, **(b)** do not *score in place* the conformers into the pharmacophore (*i.e.*, allow reorientation of the conformers to determine if they match the pharmacophore or not), **(c)** match at least 3 out of the 4 sites of the corresponding common pharmacophore according to conditions specified in Table 2, **(d)** do not prefer partial matches involving more sites and **(e)** use the excluded volumes from the corresponding binding-site environment (see Table 2). The rest of options and parameter values used during that search were the default values.

Then, those ligands with at least one pharmacophore hit in the Phase search were docked to the binding site of the same PDB file that was used to derive the pharmacophore’s excluded volumes (see Table 2). This protein-ligand docking was performed with eHiTS v2009 [63,64] by considering the receptor as a rigid body and the ligands as flexible (*i.e.*, free rotation was allowed around ligands’ single bonds). Docking conditions were by default, except for the size of the sides of the cubic box that encompassed the steroid binding pocket, which was increased to 15 Å (default value is 10 Å). Then, the resulting ligand poses were again filtered with Phase through the same pharmacophore and using the identical filtering conditions as in the first Phase run, with the exception that no reorientation of the poses was allowed during the search (*i.e.*, the *score in place* option was used) to find docking poses that were compatible with the pharmacophore. Finally, the poses that were hits in this second pharmacophore screening were submitted to a shape and electrostatic-potential comparison with the experimental poses of the 11 β -HSD1 inhibitors that

belong to the PDB files from which the common pharmacophore was derived. This comparison was performed with EON v2.0.1 (OpenEye Scientific Software, Inc., Santa Fe, New Mexico, USA; <http://www.eyesopen.com>) using the three different scores provided by EON. Thus, only VS hits that simultaneously accomplished the *ET_pb* (the Poisson-Boltzmann Electrostatic Tanimoto that compares the electrostatic potential of two small molecules) [65], the *Shape Tanimoto* (which is a quantitative measure of three-dimensional overlap) [66] and the *ET_coul* (the Coulombic part of the Poisson-Boltzmann Electrostatic Tanimoto) values that were equal or higher than 0.3, 0.5 and 0.3, respectively, toward one of the experimental poses were predicted to inhibit 11 β -HSD1. All of these processes (from the first Phase search to the EON similarity analysis) were repeated for the Phase's hits of each pharmacophore.

Finally, the last step of the VS workflow consisted of applying a filter whose aim was to remove those molecules that have a large chance of also inhibiting 11 β -HSD2 from the sample [67]. In essence, the filter consisted of (1) using Glide v5.5 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>) [68] in Standard Precision (SP) mode to generate the probable poses for all VS hits identified by EON at the steroid binding site of our 11 β -HSD2 homology model [67], (2) using the resulting poses in a shape and electrostatic potential map comparison performed by EON that used the fourth Glide SP pose for **STX351** [69] at the steroid binding site of our 11 β -HSD2 homology model as the query molecule, (3) selecting one pose per molecule (*i.e.*, the one with the highest *ET_combo* value for the query), (4) sorting molecules according to descending *ET_combo* values, and (5) considering those molecules with an *ET_combo* lower than 0.562 as selective 11 β -HSD1 inhibitors.

Similarity analysis with known 11 β -HSD1 inhibitors

To determine whether the VS hits obtained with our VS workflow can be scaffold-hopping candidates for 11 β -HSD1 inhibition, (1) all hits were merged with all the known 11 β -HSD1 inhibitors available at the BindingDB database (*i.e.*, 1054 molecules; <http://www.bindingdb.org>) [70] and (2) the resulting set was clustered using Canvas v1.2 (Schrödinger LLC., Portland, USA; <http://www.schrodinger.com>). MOLPRINT2D fingerprints [71], using a fingerprint precision of 32 bits, were calculated for each molecule, and then a hierarchical clustering based on Tanimoto similarities was obtained. The number of clusters was defined using the Kelley criterion [72]. Clusters that did not contain any known 11 β -HSD1 inhibitors were defined as clusters with new scaffolds for 11 β -HSD1 inhibitors.

Results and discussion

Structure-based pharmacophore generation for 11 β -HSD1 inhibition

Currently, there are 18 human 11 β -HSD1 complexes at the PDB that correspond to its catalytic domain (see Table 1). Two of these (*i.e.*, 1XU7 [73] and 1XU9 [73]) correspond to complexes with the steroidal detergent CHAPS and, therefore, were not further considered in our study. Another complex (*i.e.*, 2BEL) corresponds with carbenoxolone—a very potent 11 β -HSD2 inhibitor—and, therefore, was not further considered. The remaining 15 structures [32-38,40,41,74-76] correspond to complexes with 11 β -HSD1 inhibitors, but 5 of them (*i.e.*, 3BZU [74], 3D3E [41], 3D4N [41,77], 3D5Q [77] and 3EY4) have no electron density map available at the Electron Density Server [54], and therefore, the reliability of the intermolecular interactions at the ligand-binding site could not be checked. Thus, the fit of the inhibitors and their surrounding residues with the electron density map was examined for 10 complexes, and this analysis showed that only 8 of these complexes (*i.e.*, 2ILT [32], 2IRW [33], 2RBE [34], 3BYZ [35], 3CH6 [38], 3CZR [37], 3FRJ [36] and 3HFG [40]) were suitable for deriving reliable structure-based pharmacophores with LigandScout (see Table 1).

According to the LigandScout results (see Figure 1), the intermolecular interactions at the 11 β -HSD1 ligand-binding site allowed us to group the 8 structure-based pharmacophores into two different common pharmacophores (see Figures 2A and 2B). Thus, the first one (*i.e.*, **ph4-1**) was built as a result of the consensus of the pharmacophores for **10** (from 2ILT; hereafter **10-2ILT**) [32], **15** (from 2IRW; hereafter **15-2IRW**) [33], **3** (from 2RBE; hereafter **3-2RBE**) [34], **30** (from 3CH6; hereafter **30-3CH6**) [38] and **15** (from 3FRJ; hereafter **15-3FRJ**) [36] and, as may be seen in Table 2 and Figure 2A, is formed by three hydrophobic sites (*i.e.*, **H1.1**, **H2.1** and **H3.1**) and one acceptor/donor site (*i.e.*, **A1/D1**). Interestingly, this pharmacophore is built from inhibitors that belong to very different chemical families (*i.e.*, **10-2ILT** and **15-2IRW** share the same scaffold but the rest of ligands have very different structures; see Table 1) and have very different sizes. However, the inhibitors share essentially the same binding pattern to 11 β -HSD1 (see Figure 1). At this point, it is worth mentioning that although **10-2ILT**, **15-2IRW**, **30-3CH6** and **15-3FRJ** use their oxygen carbonyl as an acceptor in a hydrogen bond with Ser170 and Tyr183 (see Figures 1 and 2A), an equivalent interaction with the hydroxyl groups of these two residues would be also possible if a hydrogen bond donor were present in a location equivalent to the mentioned carbonyl oxygen (as is the case for **3-2RBE**). Therefore, this possibility has led us to consider a dual acceptor/donor behavior for this pharmacophore site. The second common pharmacophore (*i.e.*, **ph4-2**) has been built as the result of the consensus of the pharmacophores for **6d** (from 3BYZ; hereafter **6d-3BYZ**) [35], **45** (from 3CZR; hereafter **45-3CZR**) [37] and **20** (from 3HFG; hereafter

20-3HFG) [40] and, as may be seen in Table 2 and Figure 2B, is also formed by three hydrophobic sites (*i.e.*, **H1.2**, **H2.2** and **H3.2**) and one acceptor site (*i.e.*, **A2**). Interestingly, although the number and characteristics of **ph4-1** and **ph4-2** are the same (with the exception that the duality acceptor/donor is not allowed for **ph4-2**), the spatial location of their sites is very different (see Figures 2A and 2B). For instance, the key hydrogen bonding interactions of these inhibitors are made with Ala172 and not with Ser170 and Tyr183 (see Figure 1). It is also noteworthy that while **45-3CZR** and **20-3HFG** share the same scaffold, their structure and size is very different from the third 11 β -HSD1 experimental pose that contributes to this second common pharmacophore (*i.e.*, **6d-3BYZ**; see Table 1).

Moreover, to increase the discriminating power of the two common pharmacophores, a shell of excluded volumes that schematically represents the locations of the residues forming the steroid-binding site was added. Thus, for **ph4-1** and **ph4-2**, where more than one complex contributes to the same common pharmacophore, the steroid-binding sites were compared, and the results were as follows:

- The five structures that contribute to **ph4-1** have three different protein environments at the steroid-binding site: (a) the first corresponding to 2ILT, 2IRW and 3CH6 (*i.e.*, **ev1.1**; see Table 2); (b) the second corresponding to 2RBE (*i.e.*, **ev1.2**; see Table 2); and (c) the third corresponding to 3FRJ (*i.e.*, **ev1.3**; see Table 2). The main differences between these three environments were the conformations of Thr124 and Tyr177 (see Figure 2C) that, upon ligand-binding, changed in order to adapt the binding-site conformation to the inhibitor's size. Thus, for instance, **3-2RBE** (in orange in Figure 2C) could produce steric hindrance with Tyr177 if this residue remained at its 3FRJ location (in green in Figure 2C). Therefore, this justifies completing **ph4-1** with three different excluded-volume environments. From a practical point of view, this means that three different pharmacophores were derived from **ph4-1** (*i.e.*, **ph4-1_ev1.1**, **ph4-1_ev1.2** and **ph4-1_ev1.3**).
- The three structures that contributed to **ph4-2** have three different protein environments at the steroid-binding site (*i.e.*, **ev2.1**, **ev2.2** and **ev2.3**; see Table 2) because there are two residues (*i.e.*, Val231 and Met233) that have a different conformation at this location in 3BYZ, 3CZR and 3HFG (see Figure 2D). Accordingly, three different excluded-volume environments were built for **ph4-2**, which results in three different pharmacophores (*i.e.*, **ph4-2_ev2.1** for 3BYZ, **ph4-2_ev2.2** for 3CZR and **ph4-2_ev2.3** for 3HFG).

As a summary, when the excluded volumes that schematize the ligand-binding site at 11 β -HSD1 were added to the **ph4-1** and **ph4-2**, this resulted in six different common pharmacophores (*i.e.*, **ph4-1_ev1.1**, **ph4-1_ev1.2** and **ph4-1_ev1.3** derived from **ph4-1**; and **ph4-2_ev2.1**, **ph4-2_ev2.2** and **ph4-2_ev2.3** derived from **ph4-2**). Therefore, all of these pharmacophores were subsequently used in the VS workflow (and all the different environments were considered during the VS step that involves protein-ligand docking).

Prediction of new NP scaffolds for the development of selective 11 β -HSD1 inhibitors.

The VS workflow designed for 11 β -HSD1 (see Figure 3) was applied to the NP subset of the ZINC database [59] that is formed by 89,165 compounds, and it was predicted that 11 molecules (see Table 3) are selective 11 β -HSD1 inhibitors (*i.e.*, they do not inhibit 11 β -HSD2). When these 11 VS hits were merged with the 1054 known 11 β -HSD1 inhibitors and when the resulting set was classified in clusters, the following results were obtained: (1) the 1065 molecules were classified in 106 different clusters; (2) the 11 VS hits were classified in 5 clusters (*i.e.*, one hit at clusters 11, 19, 78 and 84 and seven hits at cluster 85; see Tables 3 and 4); and (3) the 1054 known 11 β -HSD1 inhibitors downloaded from the BindingDB database [70] were classified in the remaining 101 clusters. Therefore, it can be concluded that our VS workflow is able to predict 11 β -HSD1 inhibitory activities for 5 scaffolds that were not yet reported at the BindingDB. Remarkably, a SciFinder search (Chemical Abstracts Service, Columbus, Ohio, USA; <http://www.cas.org/products/sfacad>) of the literature in which the 11 hits are cited revealed that three of these hits (*i.e.*, ZINC00020259, ZINC00074836 and ZINC00074840; see Table 4) have been recently used to derive 11 β -HSD1 inhibitors (although they have not been yet included at the BindingDB) [78]. Therefore, this finding strongly supports that our VS workflow is able to discern between molecules that inhibit 11 β -HSD1 and those that do not in a molecular database.

Interestingly, the results in Table 3 justify the independent use of both **ph4-1** and **ph4-2** during the VS workflow because their results can be seen as *complementary*. **Ph4-1** is able to identify the single hit at cluster 19 and 6 out of 7 hits at cluster 85, whereas **ph4-2** identifies the single hits from clusters 11, 78 and 84 and hit in cluster 85 that was not identified by **ph4-1**.

Figure 4 displays the EON results for the comparison between the poses for the VS hits that have the highest *ET_combo* values relative to **15-3FRJ**, which contributes to **ph4-1**, and to **6d-3BYZ**, which contributes to **ph4-2** (see Table 3). Thus, the comparison between **15-3FRJ** and ZINC03643826 (Figure 4A) shows that both poses have a similar surface distribution of positive charges (in blue) whereas the common negative contribution (in red) is located at the hydroxyl group of ZINC03643826 and at the amide oxygen of **15-3FRJ**, where it allows both ligands to

make hydrogen bonds interactions with Ser170 and Tyr183 (see Figure 5). The shape comparison between **15-3FRJ** and ZINC03643826 clearly reflects the volume difference between both poses (see Figure 4A). The comparison between **6d-3BYZ** and ZINC03851930 in Figure 4B shows that there is a good agreement between the electrostatic potential and the shape of both poses that translates in the highest *ET_combo* value for the 11 VS hits (*i.e.*, 1.201; see Table 3). Moreover, the common negative contribution allows **6d-3BYZ** and ZINC03851930 to accept the amide backbone hydrogen from Ala172 (see Figure 5). Therefore, the results in Figures 4 and 5 support the ability of the electrostatic-potential and shape comparison performed by EON to smooth differences in chemical structures and to translate them into criteria important for their intermolecular interactions with the ligand-binding site.

Conclusions

The challenge of any VS protocol consists of using *in silico* tools to predict the molecules in a database that have the required activity against a specific target. Thus, the results of the present study demonstrate that our VS protocol successfully identified selective 11 β -HSD1 inhibitors with no chemical structure similarities to known active compound, and therefore, it is suitable for scaffold hopping on this target. In that sense—and considering our previous experience with similar VS workflows on other targets [79]—the hits with the highest *ET_combo* value (*i.e.*, ZINC03851930, ZINC00099956 and ZINC01774351; see Table 3) as compared with known actives would be optimal starting points for the rational drug design of potent and selective 11 β -HSD1 inhibitors with new chemical scaffolds. Thus, the *ET_combo* values of these hits are higher than those from the three VS hits that have been reported to be used to derive 11 β -HSD1 inhibitors (*i.e.*, ZINC00020259, ZINC00074836 and ZINC00074840; see Table 3) [78].

Obviously, the current VS protocol could be used to predict other 11 β -HSD1 inhibitors from larger molecular databases irrespective of their natural or synthetic nature. In that sense, its application to the complete ZINC database [59] is developing, and the most significant results of this VS will be published elsewhere.

Finally, we would like to remark that all of the tools developed in the current study (*i.e.*, the common structure-based derived pharmacophores—with the corresponding excluded volumes—and the filter to remove VS hits that also inhibit 11 β -HSD2 [67]) are available to interested readers upon request.

Acknowledgements

This manuscript was edited for English language fluency by American Journal Experts. This study was supported by Grant Number AGL2008-00387/ALI from the Ministerio de Educación y Ciencia of the Spanish Government and the ACCIÓ (TECCT10-1-0008) program (Generalitat de Catalunya). The authors wish to thank the Servei de Disseny de Fàrmacs (Drug Design Service) of the Catalonia Supercomputer Center (CESCA) for providing access to Schrödinger and OpenEye software.

| | ph4-1 | | | | | ph4-2 | | |
|---------------|--------------|-------------|-------------|-------------|-------------|--------------|-------------|-------------|
| | 2ILT | 2IRW | 2RBE | 3CH6 | 3FRJ | 3BYZ | 3CZR | 3HFG |
| Ile121 | | | | o | o | o | | o |
| Thr124 | x | | o | o | o | | o | o |
| Leu126 | o | o | o | o | o | o | o | o |
| Ser170 | x | x | x | x | x | | | |
| Leu171 | o | | o | o | o | | o | o |
| Ala172 | o | o | o | o | o | x | x | x |
| Tyr177 | o | o | o | o | o | o | o | o |
| Pro178 | | | | | o | | | |
| Met179 | | | | o | o | o | | |
| Val180 | o | o | o | o | | o | o | o |
| Tyr183 | x | x | x | x | x | o | o | o |
| Leu217 | o | | | o | | | o | o |
| Thr222 | | | | o | o | o | | o |
| Ala223 | | | o | o | o | o | o | o |
| Ala226 | | | o | | o | | o | o |
| Val227 | | o | | | | | | o |
| Ile230 | | | | o | | | | |
| Val231 | o | o | o | o | o | o | | |
| Met233 | o | | o | o | o | | o | |
| NADPH | | | | | o | | | o |

o hydrophobic contact
 x hydrogen bond

Figure 1. Intermolecular interactions between 11 β -HSD1 and 11 β -HSD1's inhibitors at reliable complexes. The results for the PDB complexes that contribute to **ph4-1** and **ph4-2** are grouped.

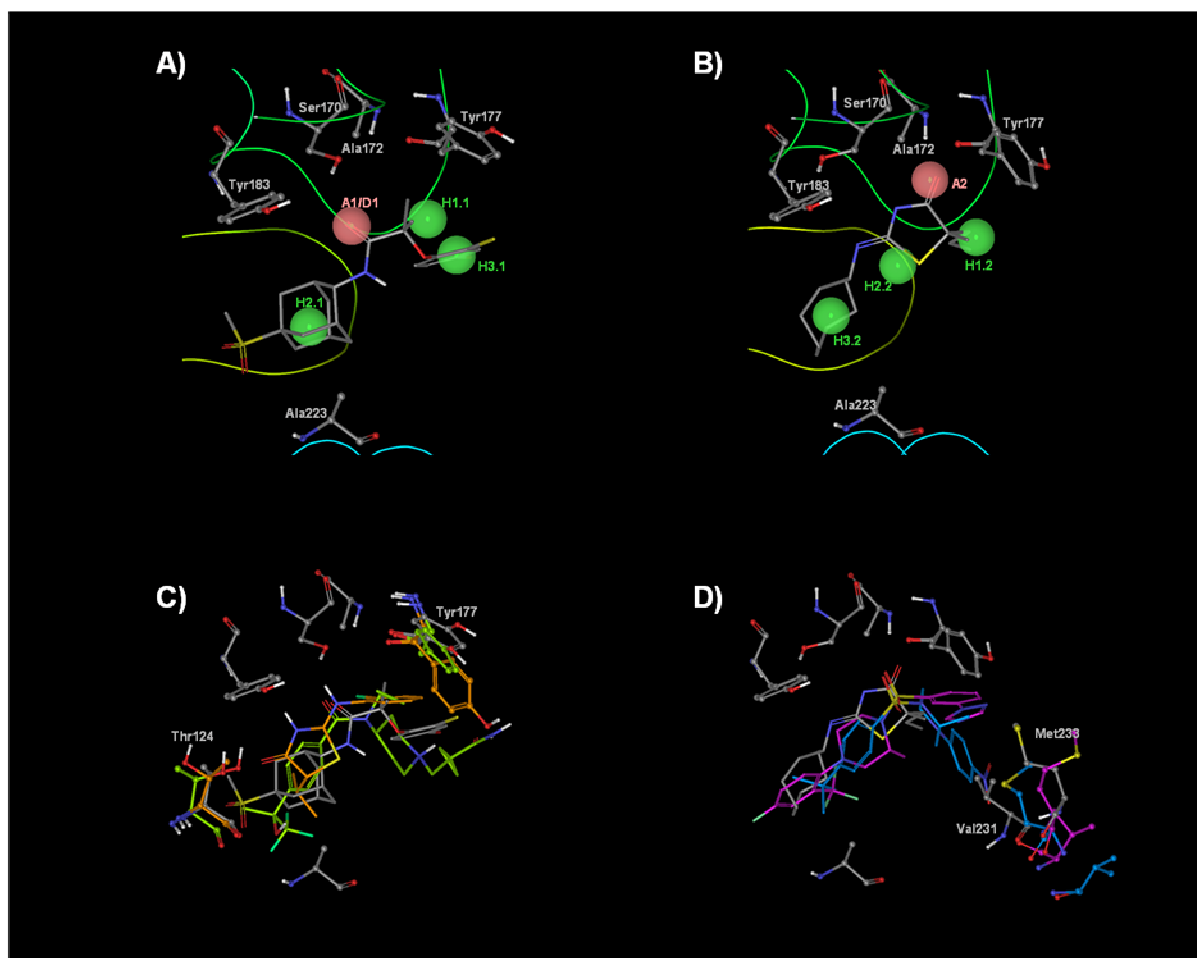


Figure 2. Structure-based pharmacophores derived from reliable 11 β -HSD1-inhibitor complexes. Figures 2A and 2B show the sites that characterize **ph4-1** and **ph4-2** (where hydrophobic and acceptor/donor sites are colored in green and pink, respectively) in the context of 2ILT and 3BYZ ligand-binding sites. Tolerances and mandatory/optional character for all of these sites during pharmacophore-based searches are shown in Table 2. Figures 2C and 2D show the main differences in the protein environments that surround each pharmacophore (that translate in three different excluded volume sets for **ph4-1** and for **ph4-2**). Residues with different conformations at the binding site of **ph4-1** are shown in grey, orange and green for 2ILT, 2RBE and 3FRJ, respectively. Residues with different conformations at the binding site of **ph4-2** are shown in grey, blue and pink for 3BYZ, 3CZR and 3HFG, respectively.

| Workflow steps | NPs subset at the ZINC database |
|---|--|
| Starting Database | 89,425 |
| ADME/Toxicity analysis FaF-Drug2 | 53,656 |
| Structure-based Pharmacophore Phase v3.1 Rigid Docking eHit's v2009 Structure-based Pharmacophore Phase v3.1 | 1,208 |
| Similarity & Electrostatic analysis EON v2.0.1 | 15 |
| 11β-HSD2 filter Glide-SP v5.5 and EON v2.0.1 | 11 |

Figure 3. The VS workflow used in the present work. The flow chart outlines the process and the progress of the VS used by this study. The data, aside from that for each VS step, show the number of molecules that *survived* it.

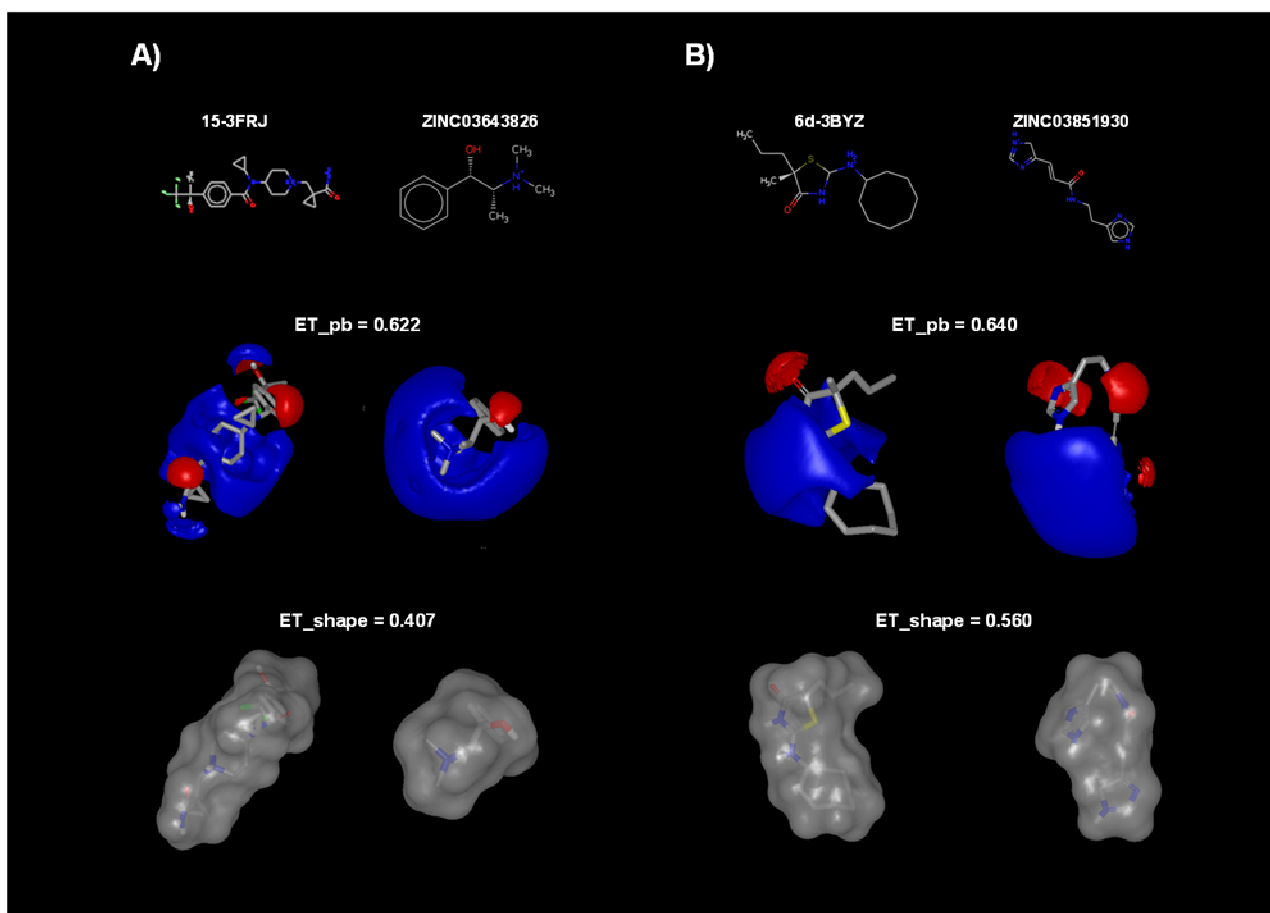


Figure 4. The EON results for the highest ET_{combo} values between poses for VS hits and **15-3FRJ (ph4-1)** and **6d-3BYZ (ph4-2)**. The chemical structure, the Poisson-Boltzmann electrostatic-potential and the shape for both query and VS hits are located side-by-side to allow for an easier comparison between them. Figures 4A and 4B compare **15-3FRJ** with ZINC03643826 and **6d-3BYZ** with ZINC03851930. This figure was drawn with VIDA v4.03 (OpenEye Scientific Software, Inc., Santa Fe, New Mexico, USA; <http://www.eyesopen.com>).

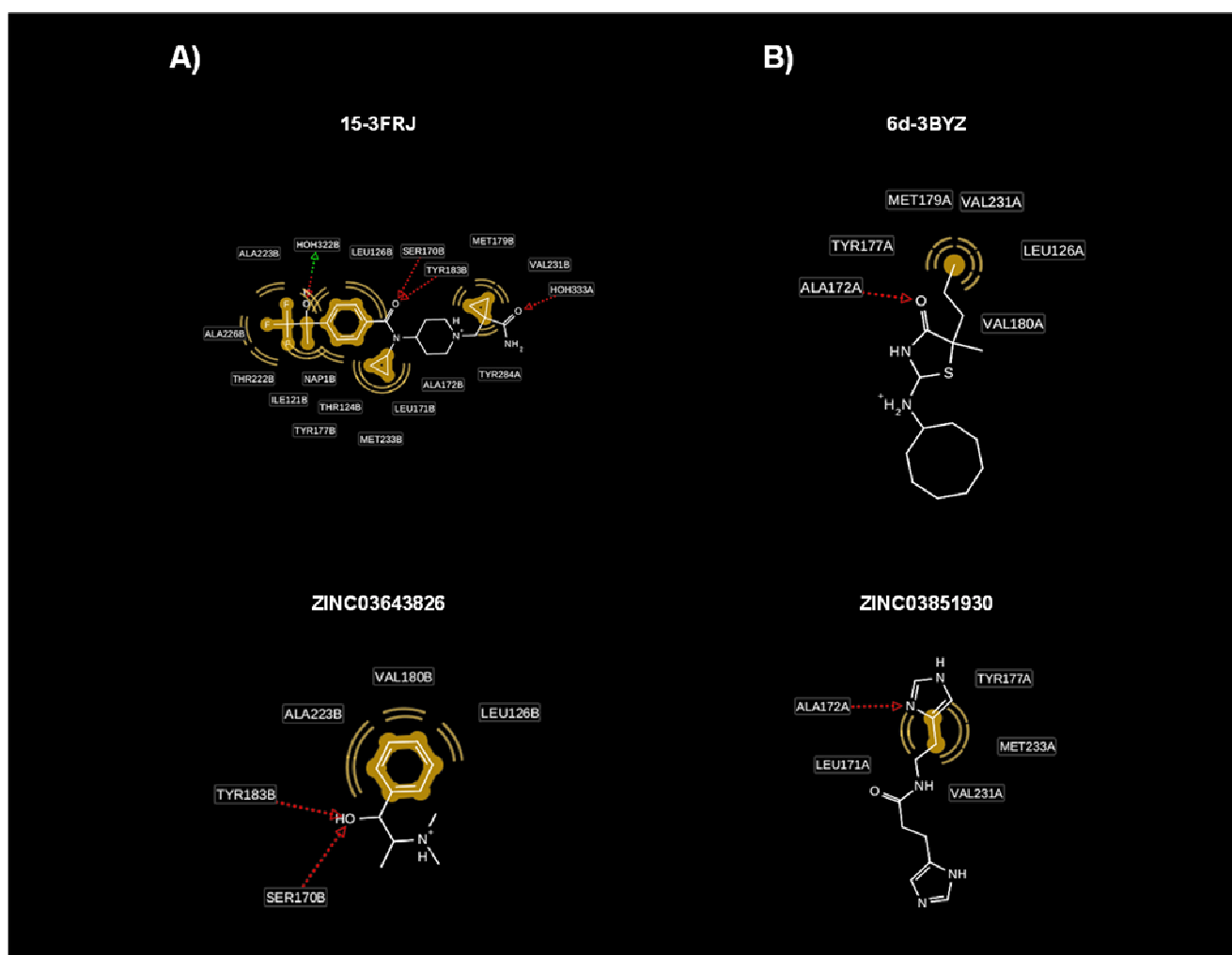
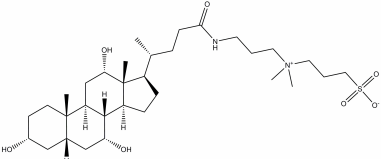
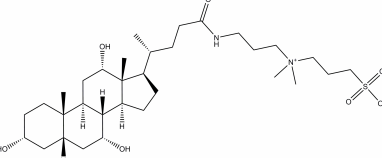
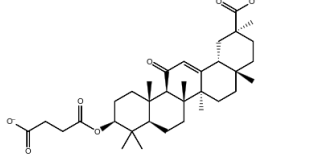
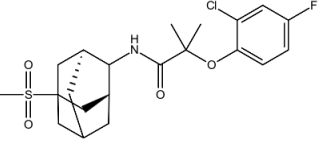
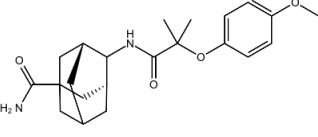
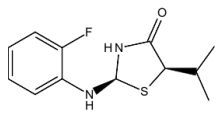
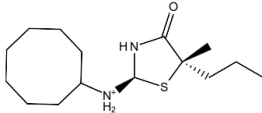
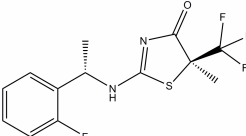
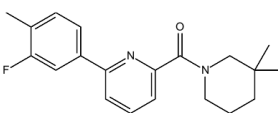
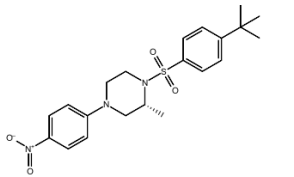
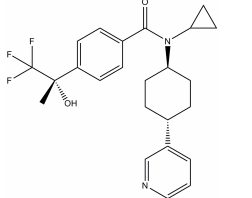
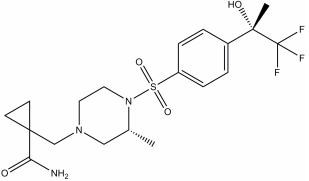
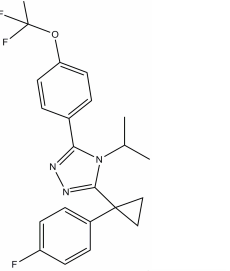
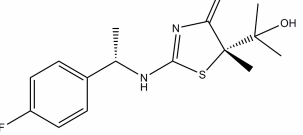
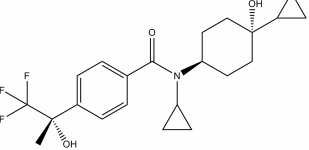
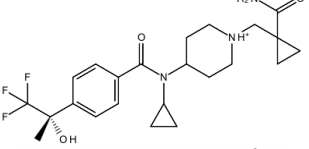
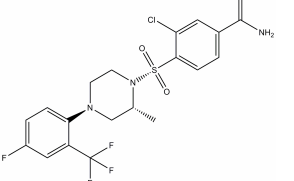
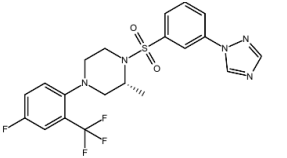


Figure 5. Intermolecular interactions for ZINC03643826/15-3FRJ and ZINC03851930/6d-3BYZ at the steroid-binding site of 3FRJ and 3BYZ. The intermolecular interactions of the poses that are compared at Figure 4 are shown here. Figures 5A and 5B show results for ZINC03643826/15-3FRJ and ZINC03851930/6d-3BYZ, respectively. This figure was derived with LigandScout v2.00 from the analysis of the corresponding docked or experimental poses.

Table 1. Current PDB structures for human 11 β -HSD1 catalytic domain

| PDB code | Ligand information | | | | | Bibliographic reference | Reliable coordinates |
|----------|---|---------------|------------------|-----------------------------|------------------------------------|-------------------------|----------------------|
| | Chemical structure | Name | Is an inhibitor? | 11 β -HSD1 inhibition | 11 β -HSD2 inhibition | | |
| 1XU7 |  | CHAPS | No | - | - | [73] | a |
| 1XU9 |  | CHAPS | No | - | - | [73] | a |
| 2BEL |  | carbenoxolone | Yes | 52.2% | 100% | - | Yes |
| 2ILT |  | 10 | Yes | 7 | 26,000 | [32] | Yes |
| 2IRW |  | 15 | Yes | 6 | 18,000 | [33] | Yes |
| 2RBE |  | 3 | Yes | 15 | n.a. | [34] | Yes |
| 3BYZ |  | 6d | Yes | 28 (Ki) | No significant inhibition detected | [35] | Yes |
| 3BZU |  | 2922 | Yes | 14 | No significant inhibition detected | [74] | b |
| 3CH6 |  | 30 | Yes | 0.1 | No significant inhibition detected | [38] | Yes |

| | | | | | | | |
|------|---|-----|-----|-----|------------------------------------|----------|-----|
| 3CZR |  | 45 | Yes | 3 | n.a. | [37] | Yes |
| 3D3E |  | 13 | Yes | 1.4 | >10,000 | [41] | b |
| 3D4N |  | 20 | Yes | 1.2 | No significant inhibition detected | [41, 77] | b |
| 3D5Q |  | 2 | Yes | 1.6 | No significant inhibition detected | [77] | b |
| 3EY4 |  | 352 | Yes | 34 | n.a. | - | b |
| 3FCO |  | 23 | Yes | 14 | n.a. | [76] | No |
| 3FRJ |  | 15 | Yes | 14 | n.a. | [36] | Yes |
| 3H6K |  | 1 | Yes | 10 | No significant inhibition detected | [40] | No |
| 3HFG |  | 20 | Yes | 26 | No significant inhibition detected | [40] | Yes |

^a The structural factors of these structures were not deposited at the PDB

^b The scripts at the Electron Density Server (<http://eds.bmc.uu.se>) were not able to generate the corresponding electron density maps from the structural factors deposited at the PDB

The inhibitory activity is expressed as IC₅₀ (in nM) for all of the molecules with the exception of **6d-3BYZ** (whose inhibitory activity on 11 β -HSD1 is expressed as K_i) and carbenoxolone [whose inhibitory activity on 11 β -HSD1 and 11 β -HSD2 is expressed as a percent of inhibition (by using a ligand concentration of 10 μ M; Patent WO 02/072084)].

Table 2. Relevant features of 11 β -HSD1 structure-based pharmacophores

| Pharmacophore label | Site type | Site tolerance (in Å) | Site condition during pharmacophore-based searches | PDB code (label of the derived environment) ^a |
|---------------------|-----------|-----------------------|--|--|
| ph4-1 | A1/D1 | 1.5 | mandatory | 2ILT (ev1.1) |
| | H1.1 | 2.0 | mandatory | 2IRW (ev1.1) |
| | H2.1 | 2.5 | optional | 3CH6 (ev1.1) |
| | H3.1 | 3.0 | optional | 2RBE (ev1.2) 3FRJ (ev1.3) |
| ph4-2 | A2 | 1.5 | optional | |
| | H1.2 | 2.5 | mandatory | 3BYZ (ev2.1) |
| | H2.2 | 2.0 | mandatory | 3CZR (ev2.2) |
| | H3.2 | 3.0 | optional | 3HFG (ev2.3) |

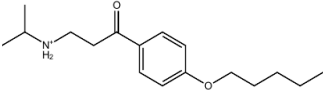
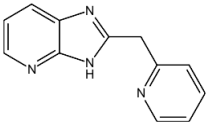
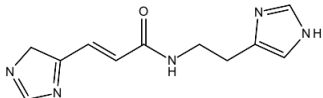
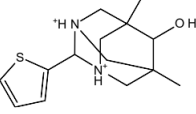
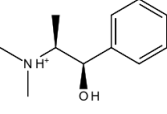
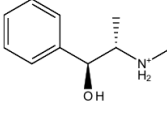
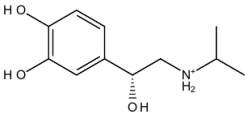
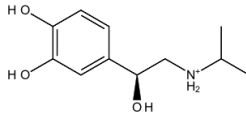
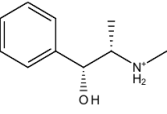
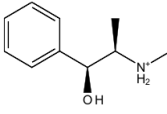
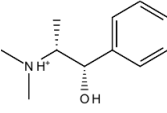
^a The PDB files that were used either to generate the different pharmacophore environments (see Figure 2) or to dock NPs during the protein-ligand docking step of the VS workflow (see Figure 3) are shown in bold.

Table 3. *ET_combo* values for selective 11 β -HSD1 inhibitors predicted by our VS workflow

| ph4-1 | | | | ph4-2 | | | |
|-------------------------|---|-----------|---|-------------------------|---|-----------|---|
| VS hit (cluster) | ET_combo 11β-HSD1 | | ET_combo 11β-HSD2 | VS hit (cluster) | ET_combo 11β-HSD1 | | ET_combo 11β-HSD2 |
| ZINC03643826 (85) | 1.068 | (15-3FRJ) | 0.432 | ZINC03851930 (78) | 1.201 | (6d-3BYZ) | 0.483 |
| ZINC00000491 (85) | 1.066 | (15-3FRJ) | 0.500 | ZINC00099956 (84) | 1.166 | (6d-3BYZ) | 0.496 |
| ZINC00074836 (85) | 1.027 | (15-3FRJ) | 0.461 | ZINC01774351 (11) | 1.079 | (6d-3BYZ) | 0.252 |
| ZINC00074840 (85) | 0.962 | (15-3FRJ) | 0.463 | ZINC00056652 (85) | 0.922 | (6d-3BYZ) | 0.507 |
| ZINC00020259 (85) | 0.954 | (15-3FRJ) | 0.396 | | | | |
| ZINC00056653 (85) | 0.910 | (15-3FRJ) | 0.516 | | | | |
| ZINC03850491 (19) | 0.864 | (3-2RBE) | 0.353 | | | | |

Highest *ET_combo* values for selective 11 β -HSD1 inhibitors predicted by our VS workflow when compared with (1) experimental poses for selective 11 β -HSD1 inhibitors (where the corresponding query pose is shown inside brackets) and (2) the fourth Glide SP pose for **STX351** at the steroid binding site of our 11 β -HSD2 homology model. Data are organized according to the common pharmacophore that retrieved the corresponding VS hits. The cluster where each VS hit was classified is also shown.

Table 4. Cluster classification of VS hits

| | | |
|--|---|---|
| <p style="text-align: center;">Cluster 11</p>  <p style="text-align: center;">ZINC01774351</p> | <p style="text-align: center;">Cluster 19</p>  <p style="text-align: center;">ZINC03850491</p> | <p style="text-align: center;">Cluster 78</p>  <p style="text-align: center;">ZINC03851930</p> |
| <p style="text-align: center;">Cluster 84</p>  <p style="text-align: center;">ZINC00099956</p> | <p style="text-align: center;">Cluster 85</p> <div style="display: flex; justify-content: space-around;"> <div data-bbox="694 728 861 840">  <p style="text-align: center;">ZINC00000491</p> </div> <div data-bbox="1045 728 1212 840">  <p style="text-align: center;">ZINC00020259^a</p> </div> </div> <div style="display: flex; justify-content: space-around;"> <div data-bbox="654 918 901 1030">  <p style="text-align: center;">ZINC00056652</p> </div> <div data-bbox="1005 918 1252 1030">  <p style="text-align: center;">ZINC00056653</p> </div> </div> <div style="display: flex; justify-content: space-around;"> <div data-bbox="694 1108 861 1220">  <p style="text-align: center;">ZINC00074836^a</p> </div> <div data-bbox="1045 1108 1212 1220">  <p style="text-align: center;">ZINC00074840^a</p> </div> </div> <div style="display: flex; justify-content: space-around;"> <div data-bbox="694 1310 861 1422">  <p style="text-align: center;">ZINC03643826</p> </div> </div> | |

^a These molecules are not currently classified at the BindingDB database as 11 β -HSD1 inhibitors, but it has been recently reported that they can be used to derive molecules with such bioactivity [78].

References

1. Cragg GM, Newman DJ. Nature: a vital source of leads for anticancer drug development. International Symposium of the Phytochemical-Society-of-Europe. Naples, ITALY: Springer, 2008:313
2. Kumar K, Waldmann H (2009) *Angewandte Chemie-International Edition* 48(18):3224
3. Rollinger JM, Stuppner H, Langer T. Virtual screening for the discovery of bioactive natural products. In: Petersen F, Amstutz R (eds). *Natural Compounds as Drugs. Volume I*. Basel: Birkhäuser Basel, 2008:211
4. Molinari G. Natural Products in Drug Discovery: Present Status and Perspectives. *Pharmaceutical Biotechnology. Volume 655*. Berlin: Springer-Verlag Berlin, 2009:13
5. Rollinger JM, Langer T, Stuppner H (2006) *Curr Med Chem* 13(13):1491
6. Williams PG (2009) *Trends in Biotechnology* 27(1):45
7. Zhou XW, Gong ZH, Su Y, Lin J, Tang KX (2009) *Journal of Pharmacy and Pharmacology* 61(3):279
8. Rollinger JM, Langer T, Stuppner H (2006) *Planta Med* 72(8):671
9. Rollinger JM, Kratschmar DV, Schuster D, Pfisterer PH, Gumy C, Aubry EM, Brandstätter S, Stuppner H, Wolber G, Odermatt A (2010) *Bioorg Med Chem* 18(4):1507
10. Gumy C, Thurnbichler C, Aubry EM, Balazs Z, Pfisterer P, Baumgartner L, Stuppner H, Odermatt A, Rollinger JM (2009) *Fitoterapia* 80(3):200
11. Ricketts ML, Shoesmith KJ, Hewison M, Strain A, Eggo MC, Stewart PM (1998) *Journal of Endocrinology* 156(1):159
12. Tomlinson JW, Walker EA, Bujalska IJ, Draper N, Lavery GG, Cooper MS, Hewison M, Stewart PM (2004) *Endocrine Reviews* 25(5):831
13. Classen-Houben D, Schuster D, Da Cunha T, Odermatt A, Wolber G, Jordis U, Kueenburg B (2009) *J Steroid Biochem Mol Biol* 113(3-5):248
14. Stegk JP, Ebert B, Martin HJ, Maser E (2009) *Mol Cell Endocrinol* 301(1-2):104
15. Galon J, Franchimont D, Hiroi N, Frey G, Boettner A, Ehrhart-Bornstein M, O'Shea JJ, Chrousos GP, Bornstein SR (2002) *FASEB J* 16(1):61
16. Seckl JR WB (2001) *Endocrinology* 142(4):1371
17. Inagaki K, Otsuka F, Miyoshi T, Watanabe N, Suzuki J, Ogura T, Makino H (2004) *Endocr J* 51(2):201
18. Wang MH (2006) *Current Opinion in Investigational Drugs* 7(4):319
19. Wamil M, Seckl JR (2007) *Drug Discov Today* 12(13-14):504

20. Zhu Y, Olson SH, Hermanowski-Vosatka A, Mundt S, Shah K, Springer M, Thieringer R, Wright S, Xiao J, Zokian H, Balkovec JM (2008) *Bioorg Med Chem Lett* 18(11):3405
21. Rosenstock J, Banarer S, Fonseca VA, Inzucchi SE, Sun W, Yao W, Hollis G, Flores R, Levy R, Williams WV, Seckl JR, Huber R, Investigators I-P (2010) *Diabetes Care* 33(7):1516
22. Tiwari A (2010) *IDrugs* 13(4):266
23. Jean DJS, Yuan C, Bercot EA, Cupples R, Chen M, Fretland J, Hale C, Hungate RW, Komorowski R, Veniant M, Wang MH, Zhang XP, Fotsch C (2007) *Journal of Medicinal Chemistry* 50(3):429
24. Schnackenberg CG (2008) *Curr Opin Investig Drugs* 9(3):295
25. St Jean DJ, Wang M, Fotsch C (2008) *Curr Top Med Chem* 8(17):1508
26. Ge R, Huang Y, Liang G, Li X (2010) *Curr Med Chem* 17(5):412
27. Boyle CD, Kowalski TJ (2009) *Expert Opin Ther Pat* 19(6):801
28. Yang HY, Dou W, Lou J, Leng Y, Shen JH (2008) *Bioorganic & Medicinal Chemistry Letters* 18(4):1340
29. Yang H, Shen Y, Chen J, Jiang Q, Leng Y, Shen J (2009) *Eur J Med Chem* 44(3):1167
30. Schuster D, Maurer E, Laggner C, Nashev L, Wilckens T, Langer T, Odermatt A (2006) *J Med Chem* 49(12):3454
31. Feng Y, Huang S-l, Dou W, Zhang S, Chen J-h, Shen Y, Shen J-h, Leng Y (2010) *Br J Pharmacol* 161(1):113
32. Sorensen B, Winn M, Rohde J, Shuai Q, Wang J, Fung S, Monzon K, Chiou W, Stolarik D, Imade H, Pan L, Deng X, Chovan L, Longenecker K, Judge R, Qin W, Brune M, Camp H, Frevert EU, Jacobson P, Link JT (2007) *Bioorg Med Chem Lett* 17(2):527
33. Patel JR, Shuai Q, Dinges J, Winn M, Pliushchev M, Fung S, Monzon K, Chiou W, Wang J, Pan L, Wagaw S, Engstrom K, Kerdesky FA, Longenecker K, Judge R, Qin W, Imade HM, Stolarik D, Beno DWA, Brune M, Chovan LE, Sham HL, Jacobson P, Link JT (2007) *Bioorg Med Chem Lett* 17(3):750
34. Yuan C, St Jean DJ, Liu Q, Cai L, Li A, Han N, Moniz G, Askew B, Hungate RW, Johansson L, Tedenborg L, Pyring D, Williams M, Hale C, Chen M, Cupples R, Zhang J, Jordan S, Bartberger MD, Sun Y, Emery M, Wang M, Fotsch C (2007) *Bioorg Med Chem Lett* 17(22):6056
35. Johansson L, Fotsch C, Bartberger MD, Castro VM, Chen M, Emery M, Gustafsson S, Hale C, Hickman D, Homan E, Jordan SR, Komorowski R, Li A, McRae K, Moniz G, Matsumoto G, Orihuela C, Palm G, Veniant M, Wang M, Williams M, Zhang J (2008) *J Med Chem* 51(10):2933

36. Rew Y, McMinn DL, Wang Z, He X, Hungate RW, Jaen JC, Sudom A, Sun D, Tu H, Ursu S, Villemure E, Walker NP, Yan X, Ye Q, Powers JP (2009) *Bioorg Med Chem Lett* 19(6):1797
37. Sun D, Wang Z, Di Y, Jaen JC, Labelle M, Ma J, Miao S, Sudom A, Tang L, Tomooka CS, Tu H, Ursu S, Walker N, Yan X, Ye Q, Powers JP (2008) *Bioorg Med Chem Lett* 18(12):3513
38. Wang H, Ruan Z, Li JJ, Simpkins LM, Smirk RA, Wu SC, Hutchins RD, Nirschl DS, Van Kirk K, Cooper CB, Sutton JC, Ma Z, Golla R, Seethala R, Salyan ME, Nayeem A, Krystek SR, Sheriff S, Camac DM, Morin PE, Carpenter B, Robl JA, Zahler R, Gordon DA, Hamann LG (2008) *Bioorg Med Chem Lett* 18(11):3168
39. Rohde JJ, Pliushchev MA, Sorensen BK, Wodka D, Shuai Q, Wang J, Fung S, Monzon KM, Chiou WJ, Pan L, Deng X, Chovan LE, Ramaiya A, Mullally M, Henry RF, Stolarik DF, Imade HM, Marsh KC, Beno DW, Fey TA, Droz BA, Brune ME, Camp HS, Sham HL, Frevert EU, Jacobson PB, Link JT (2007) *J Med Chem* 50(1):149
40. Wan ZK, Chenail E, Xiang J, Li HQ, Ipek M, Bard J, Svenson K, Mansour TS, Xu X, Tian X, Suri V, Hahm S, Xing Y, Johnson CE, Li X, Qadri A, Panza D, Perreault M, Tobin JF, Saiah E (2009) *J Med Chem* 52(17):5449
41. Julian LD, Wang Z, Bostick T, Caille S, Choi R, DeGraffenreid M, Di Y, He X, Hungate RW, Jaen JC, Liu J, Monshouwer M, McMinn D, Rew Y, Sudom A, Sun D, Tu H, Ursu S, Walker N, Yan X, Ye Q, Powers JP (2008) *J Med Chem* 51(13):3953
42. Beseda I, Czollner L, Shah PS, Khunt R, Gaware R, Kosma P, Stanetty C, Del Ruiz-Ruiz MC, Amer H, Mereiter K, Da Cunha T, Odermatt A, Classen-Houben D, Jordis U (2010) *Bioorg Med Chem* 18(1):433
43. Walker EA, Stewart PM (2003) *Trends Endocrinol Metab* 14(7):334
44. Ferrari P (2010) *Biochim Biophys Acta* 1802(12):1178
45. Sala, E., Guasch L, Mulero M, Valls C, Salvadó MJ, Arola-Arnal A, Garcia-Vallvé S, Pujadas G (2011) *J Comput Aided Mol Des* (*Submitted needing revision*)
46. Oppermann U, Filling C, Hult M, Shafqat N, Wu XQ, Lindh M, Shafqat J, Nordling E, Kallberg Y, Persson B, Jornvall H (2003) *Chemico-Biological Interactions* 143:247
47. Persson B, Kallberg Y, Oppermann U, Jornvall H (2003) *Chemico-Biological Interactions* 143:271
48. Bray JE, Marsden BD, Oppermann U (2009) *Chemico-Biological Interactions* 178(1-3):99
49. Oppermann UCT, Filling C, Jornvall H (2001) *Chemico-Biological Interactions* 130(1-3):699
50. Kavanagh KL, Jornvall H, Persson B, Oppermann U (2008) *Cell Mol Life Sci* 65(24):3895

51. Kim KW, Wang Z, Busby J, Tsuruda T, Chen M, Hale C, Castro VM, Svensson S, Nybo R, Xiong F, Wang M (2006) *Biochim Biophys Acta* 1764(4):824
52. Odermatt A, Atanasov AG, Balazs Z, Schweizer RA, Nashev LG, Schuster D, Langer T (2006) *Mol Cell Endocrinol* 248(1-2):15
53. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C (2002) *Acta Crystallographica Section D-Biological Crystallography* 58:899
54. Kleywegt GJ, Harris MR, Zou JY, Taylor TC, Wählby A, Jones TA (2004) *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2240
55. Wallace AC, Laskowski RA, Thornton JM (1995) *Protein Eng* 8(2):127
56. Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML, Thornton JM (1997) *Trends Biochem Sci* 22(12):488
57. Guex N, Peitsch MC (1997) *Electrophoresis* 18(15):2714
58. Wolber G, Langer T (2005) *J Chem Inf Model* 45(1):160
59. Irwin JJ, Shoichet BK (2005) *Journal of Chemical Information and Modeling* 45(1):177
60. Lagorce D, Sperandio O, Galons H, Miteva MA, Villoutreix BO (2008) *BMC Bioinformatics* 9:396
61. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) *Adv Drug Deliv Rev* 46(1-3):3
62. Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA (2006) *J Comput Aided Mol Des* 20(10-11):647
63. Zsoldos Z, Reid D, Simon A, Sadjad BS, Johnson AP (2006) *Curr Protein Pept Sci* 7(5):421
64. Zsoldos Z, Reid D, Simon A, Sadjad SB, Johnson AP (2007) *J Mol Graph Model* 26(1):198
65. Naylor E, Arredouani A, Vasudevan SR, Lewis AM, Parkesh R, Mizote A, Rosen D, Thomas JM, Izumi M, Ganesan A, Galione A, Churchill GC (2009) *Nat Chem Biol* 5(4):220
66. Rush TS, 3rd, Grant JA, Mosyak L, Nicholls A (2005) *J Med Chem* 48(5):1489
67. Sala E, Guasch L, Mulero M, Valls C, Salvadó MJ, Arola-Arnal A, Garcia-Vallvé S, Pujadas G (2011) *J Comput Aided Mol Des*
68. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) *J Med Chem* 47(7):1739

69. Potter BVL, Purohit A, Reed MJ, Vicker N; Glycyrrhetic acid derivatives, progesterone and progesterone derivatives and their use for the manufacture of a medicament to inhibit 11beta-hydroxysteroid dehydrogenase activity. 2002
70. Liu TQ, Lin YM, Wen X, Jorissen RN, Gilson MK (2007) *Nucleic Acids Research* 35:D198
71. Duan J, Dixon SL, Lowrie JF, Sherman W (2010) *J Mol Graph Model* 29(2):157
72. Kelley LA, Gardner SP, Sutcliffe MJ (1996) *Protein Eng* 9(11):1063
73. Hosfield DJ, Wu Y, Skene RJ, Hilgers M, Jennings A, Snell GP, Aertgeerts K (2005) *J Biol Chem* 280(6):4639
74. Hale C, Véniant M, Wang Z, Chen M, McCormick J, Cupples R, Hickman D, Min X, Sudom A, Xu H, Matsumoto G, Fotsch C, St Jean DJ, Wang M (2008) *Chem Biol Drug Des* 71(1):36
75. Tu H, Powers JP, Liu JS, Ursu S, Sudom A, Yan XL, Xu HD, Meininger D, DeGraffenreid M, He X, Jaen JC, Sun DQ, Labelle M, Yamamoto H, Shan B, Walker NPC, Wang ZL (2008) *Bioorganic & Medicinal Chemistry* 16(19):8922
76. McMinn DL, Rew Y, Sudom A, Caille S, Degraffenreid M, He X, Hungate R, Jiang B, Jaen J, Julian LD, Kaizerman J, Novak P, Sun D, Tu H, Ursu S, Walker NP, Yan X, Ye Q, Wang Z, Powers JP (2009) *Bioorg Med Chem Lett* 19(5):1446
77. Tu H, Powers JP, Liu J, Ursu S, Sudom A, Yan X, Xu H, Meininger D, Degraffenreid M, He X, Jaen JC, Sun D, Labelle M, Yamamoto H, Shan B, Walker NP, Wang Z (2008) *Bioorg Med Chem* 16(19):8922
78. Moniz GA, Frizzle MJ, Bernard C, Martinelli M, Faul M, Larrow J, Hansen KB, Klingensmith L; Process for making substituted 2-amino-thiazolones. 2010
79. Sala E, Guasch L, Iwaszkiewicz J, Mulero M, Salvadó MJ, Pinent M, Zoete V, Grosdidier A, Garcia-Vallvé S, Michielin O, Pujadas G (2011) *PLoS ONE* 6(2): e16903. doi:10.1371/journal.pone.0016903

IV. SUMMARIZING DISCUSSION

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

Finding new drugs are undoubtedly one of the most challenging and exciting tasks in science today¹⁻³. Fortunately, during the last two decades, virtual screening workflows have emerged that increase the chance of finding small molecules that show the desired bioactivity against a therapeutic target, thus significantly decreasing the cost associated with bringing a new drug to the market⁴. In the present thesis, the same *in silico* tools that are routinely used by the pharmaceutical industry to develop new drugs⁴⁻⁷ have been used to screen natural product databases in order to search for molecules that inhibit IKK-2 and 11 β -HSD1 and that, therefore, can be used as a bioactive ingredients to design functional foods for the prevention of metabolic syndrome and its related pathologies (such as cardiovascular disease).

A fundamental step for the development of both IKK-2 and 11 β -HSD1 structure-based virtual screenings shown in this report was building homology models for IKK-2 and 11 β -HSD2. This task was not trivial, because of the low sequence similarity between these two proteins and the templates that are available at the Protein Data Bank (<http://www.pdb.org>). In spite of these difficulties, the homology models derived for both enzymes have successfully allowed the generation of bioactive ligand conformations by protein-ligand docking on the corresponding protein binding sites that, subsequently, were used to either derive one structure-based pharmacophore (for IKK-2) or to find the query pose that better distinguished between active and non-active molecules for 11 β -HSD1. Obviously, the performance of the virtual screening workflows shown in the present PhD thesis would greatly benefit from the publication of IKK-2 and 11 β -HSD2 experimental structures in the PDB, but unfortunately, this is out of our control.

The suitability of the virtual screening methodology developed in this thesis in the functional-food design field (but also for drug discovery purposes) is especially well reflected in the work performed with IKK-2. Thus, it has been demonstrated that a stricter version of the original IKK-2 workflow is able to find hits that are present in a relatively large number of natural extracts with proven anti-inflammatory properties. Moreover, the same strict protocol predicts that there are 414 other extracts with undescribed anti-inflammatory activities that contain at least one IKK-2 inhibitor. Consequently, this work allows the discovery of new anti-inflammatory extracts of natural origin that could be of use, for example, in the design of functional foods aimed at preventing diseases that are the result of chronic inflammation. In this context, the characterization of such extracts merits further attention, and current work in this regard is underway in our research group. In the case of drug discovery, the results show that the workflow is able to identify IKK-2 inhibitors in very challenging conditions (*i.e.*, not structurally related to any known synthetic molecule that inhibits hIKK-2 and never reported to have anti-inflammatory activity),

and therefore, it is suitable to find new chemical scaffolds for further the development of potent and specific IKK-2 inhibitors.

The second part of this thesis was focused on the 11 β -HSD1 target. Thus, initially, we created an innovative, robust, easy-to-use and accurate *in silico* approach that is able to discern in a molecular database between molecules that inhibit 11 β -HSD2 and those that do not. This discriminatory filter, used as the last virtual screening step when looking for 11 β -HSD1 inhibitors, will allow for an increased chance of finding 11 β -HSD1 selective inhibitors and, consequently, is expected to significantly reduce the budget required to experimentally prove 11 β -HSD1 selectivity. However, in spite of these findings, it is advisable to experimentally test the performance of the methodology (that is currently only tested by *in silico* methods).

The results of this thesis also show that the combination of computational techniques (*i.e.*, protein-ligand docking, pharmacophore modeling, electrostatic and similarity searches) during the same virtual screening based on different approaches have a very positive impact on the performance of such screenings. At this point, it is important to emphasize the fundamental contribution made by the EON program during the screening workflow because it is very effective in using known bioactive molecules to find new lead-hopping candidates in molecular databases by smoothing differences in chemical structures and translating them into criteria important for their intermolecular interactions with the ligand-binding site⁸.

Last but not least, I would like to remark that the different tools developed during this PhD thesis are available upon request, and therefore, it is expected that they could contribute to the development of cheaper and better anti-inflammatory drugs, hopefully returning the funds spent in my training and education to society.

REFERENCES

1. Knehans, T.; Schüller, A.; Doan, D. N.; Nacro, K.; Hill, J.; Güntert, P.; Madhusudhan, M. S.; Weil, T.; Vasudevan, S. G. Structure-guided fragment-based in silico drug design of dengue protease inhibitors. *J Comput Aided Mol Des* **2011**.
2. Munteanu, C. R.; Fernández-Blanco, E.; Seoane, J. A.; Izquierdo-Novo, P.; Rodríguez-Fernández, J. A.; Prieto-González, J. M.; Rabuñal, J. R.; Pazos, A. Drug discovery and design for complex diseases through QSAR computational methods. *Curr Pharm Des* **2010**, *16*, 2640-55.
3. Mullard, A. 2010 FDA drug approvals. *Nat Rev Drug Discov* **2011**, *10*, 82-5.
4. Sousa, S. F.; Cerqueira, N. M.; Fernandes, P. A.; Ramos, M. J. Virtual screening in drug design and development. *Comb Chem High Throughput Screen* **2010**, *13*, 442-53.
5. Langer, T.; Krovat, E. M. Chemical feature-based pharmacophores and virtual library screening for discovery of new leads. *Current Opinion in Drug Discovery & Development* **2003**, *6*, 370-376.
6. Ripphausen, P.; Nisius, B.; Bajorath, J. State-of-the-art in ligand-based virtual screening. *Drug Discov Today* **2011**.
7. Zhang, S. Computer-aided drug discovery and development. *Methods Mol Biol* **2011**, *716*, 23-38.
8. Sala, E.; Guasch, L.; Iwaszkiewicz, J.; Mulero, M.; Salvadó, M. J.; Pinent, M.; Zoete, V.; Grosdidier, A.; Garcia-Vallvé, S.; Michielin, O.; Pujadas, G. Identification of Human IKK-2 Inhibitors of Natural Origin (Part I): Modeling of the IKK-2 Kinase Domain, Virtual Screening and Activity Assays. *PLoS ONE* **2011**, in press.

V. CONCLUSIONS

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

1. The ATP-binding site of the homology model for human IKK-2 (hIKK-2) has been correctly modeled. Protein-ligand docking calculations with known IKK-2 inhibitors are able to reproduce docked poses with the interactions described in the literature as important for kinase inhibitors. These knowledge-based coherent poses can be used to derive a structure-based pharmacophore for hIKK-2.
2. The success rate of the predictions done by the hIKK-2 virtual screening workflow has been proven in very challenging conditions. Thus, it has been experimentally shown that this virtual-screening protocol is able to identify hIKK-2 inhibitors that (a) are not structurally related to any known synthetic molecule that inhibits hIKK-2 and (b) have never been reported to have anti-inflammatory activity.
3. When a stricter version of the hIKK-2 virtual screening workflow is applied to an in-house database of 29,779 natural products that are annotated with their natural source, the screen predicted that 274 molecules (isolated from 453 different natural extracts) inhibit hIKK-2. An exhaustive bibliographic search reveals that anti-inflammatory activity has been described for (a) 36 out of these 453 extracts and (b) 17 out of 30 virtual screening hits present in these 36 extracts. Only one of the remaining 13 hit molecules in these extracts shows chemical similarity with known synthetic hIKK-2 inhibitors. Therefore, it is reasonable to suspect that a significant portion of the remaining 12 hit molecules are lead candidates of natural origin that can be used to develop new hIKK-2 inhibitors.
4. The 3D-QSAR model developed for IKK-2 inhibition allows one to predict the activity of molecules with a pyridine scaffold.
5. The virtual screening protocol of 11 β -HSD1 successfully identifies the selective inhibitors of natural origin with no chemical-structure similarities to synthetic 11 β -HSD1 inhibitors, and therefore, it is suitable for scaffold hopping on this target.
6. The 11 β -HSD2 homology model demonstrated its ability to bind known inhibitors in a flipped orientation on the steroid-binding site, which, in turn, allows the use of these poses in performing structure-based approaches.
7. We have developed a robust, easy-to-use and accurate *in silico* approach that is able to discern in a molecular database between molecules that inhibit 11 β -HSD2 and those that do not inhibit this target.

8. The results of this thesis emphasize the importance of using EON during the calculation of the similarity between two poses because this program has been shown to be very effective in using known bioactive molecules for finding new lead-hopping candidates in molecular databases by smoothing differences in chemical structures and translating them into criteria important for their intermolecular interactions with the ligand-binding site.

9. Our work also opens the door to the discovery of new anti-inflammatory natural compounds and/or extracts that can be of use, for instance, in the design of functional foods focused toward the prevention of diseases that are the result of chronic inflammation.

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011

UNIVERSITAT ROVIRA I VIRGILI

IN SILICO METHODOLOGIES FOR THE DESIGN OF FUNCTIONAL FOODS THAT CAN PREVENT CARDIOVASCULAR DISEASES

Esther Sala Argüello

ISBN:/DL:T. 1030-2011