UNIVERSITAT POLITÈCNICA DE CATALUNYA

BARCELONA, 2014-2018

Study of conductive polymers, biomolecules and their hybrids through computational approaches

by

Jordi Triguero Enguídanos

Directors

Prof. Carlos Enrique Alemán Llansó

Dr. David Zanuy Gómara

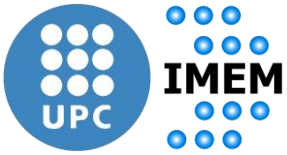
Thesis submitted to obtain the degree of Doctor of Philosophy

Supported by

MINECO: BES-2013-065536

and

FEDER funds: MAT2012-34498 & MAT2015-69367-R



Als meus pares i avis,

els vostres esforços
han permès que avui sigui aquí.

ABSTRACT

In this Thesis, different theoretical approaches designed to study a wide interval of length- and/or time-scales have been used to examine the microscopic properties of chemical systems with varying degrees of complexity and size. Firstly, the conformational preferences and optical properties of a tripeptide derived from the RGD sequence, the unit of a cell adhesive activity domain in adherent proteins, have been analyzed. Calculations on this peptide, which contains an exotic amino acid bearing a 3,4-ethylenedioxythiophene (EDOT) ring as side group, have been performed using DFT and time depending DFT methods. Results indicate that the bioactive characteristics of the RGD sequence become unstable in presence of the new residue because of both the steric hindrance caused by the EDOT side group and the repulsive interactions between the oxygen atoms belonging to the backbone amide groups and the EDOT side group. This information has been used to propose some chemical changes oriented to improve the bioadhesive properties. The interaction of a cell penetrating tetrapeptide, RPAR, adsorbed onto a gold substrate and the deposition of a pre-assembled peptide-polymer conjugate, cyc[(L-Gln-D-Ala-L-Lys-D-Ala)2] coupled with two poly(n-butyl acrylate) blocks, onto a mica substrate have been examined through molecular dynamics (MD). Results indicate that RPAR binds both the (100) and (111) gold surface facets. The conformation of the adsorbed peptide differs considerably from the bioactive conformation. However, the new conformations are not stabilized by strong intramolecular interactions. Accordingly, gold nanoparticles can be considered as suitable vehicles for the transport and targeted delivery of this CendR peptide. For the pre-assembled peptide-polymer conjugate, a theoretical approach that simulates the selective and progressive desolvation of the nanotube-like assembly has been used, demonstrating that the solvent presence during deposition is the main responsible for the unexplained conformational preferences of the acrylate blocks. When the proportion of solvent drops, the loss of many attractive solute-solvent interactions induces a meaningful increase in the number of torsions. MD has been also used to understand the impact of the solvation medium and the action of a detergent in the structure of a representative outer membrane protein (OMP).

Calculations show the destabilization of the protein in water, while in presence of detergent molecules in solution or the bilayer induce a partial and complete protective effect, respectively. Combining stochastic algorithms and MD simulations the atomistic details of polymer coatings deposited over metal surfaces have been modeled to reproduce the experimentally observed topographic features of poly(3,4-ethylenedioxythiophene), PEDOT, deposited onto stainless steel. Results have provided an excellent model system to test the developed modeling strategy. Multi-phasic simulations have been conducted to explain the influence of protein···polymer interactions in the antimicrobial biocapacitors activity of lysozyme (LYZ)-containing PEDOT electrodes through atomistic MD calculations. Two models have been investigated: i) a biphasic system in which the protein was adsorbed onto the surface of PEDOT, PEDOT/LYZ; and ii) a biocomposite in which the LYZ was homogeneously distributed inside the PEDOT matrix, P(EDOT-LYZ). MD simulations have been performed in absence and presence of electric fields, the latter mimicking the one originated by the voltage cell difference in biocapacitors. In PEDOT/LYZ electrodes, the loss of biological activity has been attributed to the anisotropy of the PEDOT···LYZ electrostatic interactions. In contrast, anisotropic effects are minimized in P(EDOT-LYZ), conserving activity. The bactericidal activity of PEDOT/LYZ and P(EDOT-LYZ) biocapacitors is independent of the electric field applied or supplied during charge-discharge processes.

INDEX

INDEX

AGRAIMENTS

Ja han passat 5 anys de tesis. Gairebé res ha anat com m'esperava o imaginava. A tots els nivells. Ha sigut el període més llarg des de la primària, de 6 anys, i no n'acabo de ser conscient.

La part dels agraïments sempre m'ha fet molta por per no saber com expressar-me i sobretot oblidar-me d'algú o d'alguna cosa. Els que em coneixeu sabeu que puc ser despistat, no m'ho tingueu en compte (també el ser curt de paraules).

Vull començar donant les gràcies als meus directors de tesis, el Dr. Carlos Alemán i el Dr. David Zanuy, per donar-me l'oportunitat de dur a terme el meu somni de sempre de ser doctor. És un fet realment important per mi. També als Dr. Eric Perpète i Dra. Catherine Michaux per la seva càlida rebuda durant la meva estança de doctorat a Namur, Bèlgica.

Volia agrair als meus pares tots els esforços i dedicacions que m'han brindat sempre per tal de que pogués complir les meves aspiracions. Però continuo sense endreçar l'habitació. També a ma germana Eira que ha estat quan ho he necessitat.

No em vull deixar de mencionar als altres professors i investigadors que han apostat per mi en tots els meus anys de formació, en especial a Ramón Sancliments† (secundaria i batxillerat), a Bibiana (primària), a Antonio García de Herreros (PRBB) i a Juan Fernández Recio (BSC). M'heu obert moltes portes i m'heu ensenyat molt.

AGRAIMENTS

Al Pau Cabañeros, que ha estat, tot i la distancia, en tots els moments escoltant i aconsellant. Tant se val tots els kilòmetres, amb tu només són una anotació/curiositat/anècdota a peu de pàgina.

A Yi Bo i Marc, que els nostres viatges i passejos mentre conversem ens continuïn portant arreu.

A Jaume Costa, que tot i haver hagut un planeta entre mig el retorn ha sigut èpic.

A tots els meus companys de recerca, que tot i passar-me els dies tancat rere una pantalla al despatx, han sigut motiu de riures i somriures, sobretot els fixes del dinar (especialment les An*a's, Abdel, Elena, Guillem's i Max). Hem de continuar discutint de tot.

Agraïments a tots els que heu estat en un període més que intens per mi, em teniu pel que necessiteu, ja ho sabeu ;)

Sincerament vostre,

Jordi

SYMBOLS AND ABBREVIATIONS

Symbol	Meaning
$(OmpF)_{aq}$	OmpF in aqueous solution
$(OmpF)_{bl}$	OmpF embedded in a bilayer
$(OmpF)_{DDM}$	OmpF in a detergent solution
1LKS	Hen egg white lysozyme nitrate
AFM	Atomic Force Microscopy || microscopio de fuerza atómica
b-b	Backbone⋯backbone interactions
βS	β–strand
CA	Chronoamperometry
CC	Coupled Cluster || Clúster Acoplado
χ_{DMF}	Density of the solvent bulk (DMF)
CendR	C-end Rule
CMC	Critical Micelle Concentration
D	Diffusion coefficients
ΔΔE	Energy gap between two consecutive unique minima
DDM	n-dodecyl β-D-maltoside
ΔE	Relative energies
DFT	Density Functional Theory
DMF	Dimethylformamide
DRX	Difracción de Rayos X
ECP	Electroactive Conducting Polymer
EDD	Electronic Density Difference
EDOT	3,4-ethylenedioxythiophene
ε_p	Hardness for the inert particle
$\overrightarrow{E_z}$	External electric field
fcc	face-centred cubic
Fmoc	N-(fluorenyl-9-methoxycarbonyl)
FsNMs	Free-standing NanoMembranes
G^E || GlE	Exotic amino acid bearing a 3,4-ethylenedioxythiophene as side group attached to an additional methylene group
GTO	Gaussian-Type Orbitals
H	Helices
HF	Hartree-Fock

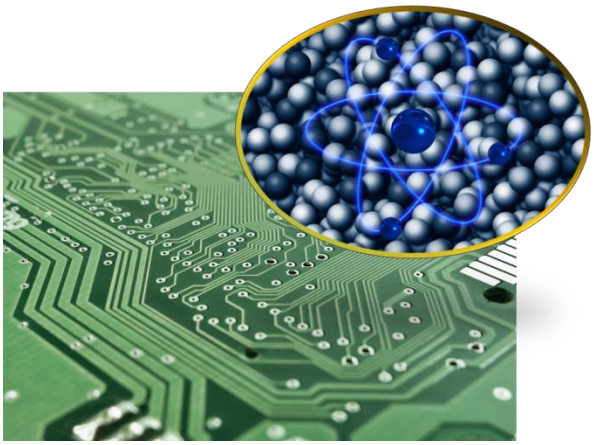
ℓ	Thickness
L+T	Loops and turns
LYZ	Lysozyme
MD	Molecular dynamics
MP	Moller-Plesset
NRP-1	Neuropilin-1
OMP	Outer Membrane Porin
P(EDOT-LYZ)	Lysozyme homogeneously distributed inside the PEDOT matrix
PBD	Protein Data Bank
PCM	Polarizable Continuum Model
PEDOT	poly(3,4-ethylenedioxythiophene)
PEDOT/LYZ	Lysozyme adsorbed onto the surface of PEDOT films
PLA	poly(lactic acid)
PME	Particle Mesh of Eward
P*n*BA	poly(n-butyl acrylate)
PNMPy	poly(N-methylpyrrole)
PPy	polypyrrole
R_a	Average height
RDF	Radial Distribution Function
R_g	Radius of gyration
RGD	Arg-Gly-Asp
R_H	Hydrodynamic radius
RMD	Root Mean Square Deviation
RMN	Resonancia Magnética Nuclear
RMSF	Root Mean Square Fluctuations
r_p	Radius for the inert particle
R_q	Root mean square roughness
RUs	Total numbers of repeating units
SASA	Solvent Accessible Surface Areas
sc-b	side chain⋯backbone interactions
SCRF	Self-Consistent Reaction Field
sc-sc	side chain⋯side chain interactions
SEM	Scanning Electron Microscopy
SPM	Scanning Probing Microscopy || microscopía de barrido por sondeo

STM	Scanning Tunneling Microscope || microscopio de efecto túnel
STO	Slater-Type Orbitals
TD-DFT	Time-Dependent DFT
Th	Thiophene
θ	Polymerization times

FIGURES, TABLES AND SCHEMES

- Capítulo 1 -

Introducción

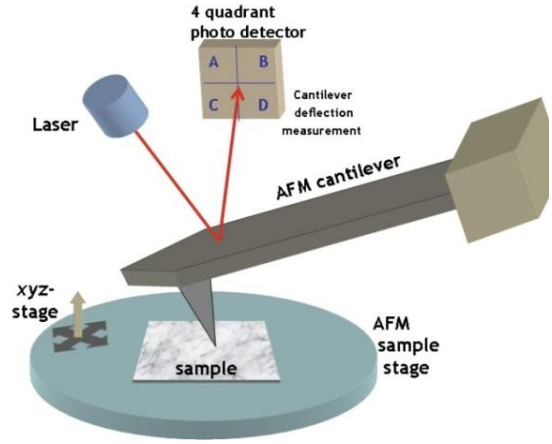


1.1 EXPERIMENTO Y RACIONALIZACION

La ciencia y la técnica se apoyan y complementan para avanzar en el conocimiento del mundo natural. El entender por qué suceden las cosas permite crear ingenios que ayudan a llegar a aquello que no era posible antes para, normalmente, generar nuevas preguntas. La fabricación de lentes pulidas para construir microscopios y telescopios, y el comprender la expansión de los gases para inventar la máquina de vapor o el motor de combustión, son ejemplos de esta mutua retroalimentación. La ciencia permite crear nuevas herramientas mientras que hay herramientas que revolucionan la ciencia.

El estudio de la materia a nivel atómico y molecular, tradicionalmente, ha sido a través de la deducción y racionalización de los fenómenos macroscópicos que se observaban. Actualmente, en las primeras décadas del s. XXI, tenemos instrumentos que nos permiten estudiar la materia a escalas micro- y nanoscópicas, si bien con limitaciones técnicas muy importantes. Uno de los instrumentos más usados que permiten resolución atómica o molecular es el *microscopio de fuerza atómica* (AFM: Atomic Force Microscopy),[1–4] que se incluye entre las técnicas de *microscopía de barrido por sondeo* (SPM: Scanning probing microscopy). El AFM es un instrumento mecano-óptico, capaz de detectar fuerzas del orden de los nanonewtons, que consiste en una punta afilada de forma piramidal o cónica situada en el extremo de una palanca flexible que recorre la superficie de una muestra manteniendo constante una pequeña fuerza de interacción, siendo así capaz de

registrar continuamente su topografía. El movimiento de barrido lo realiza un escáner piezo-eléctrico, y la interacción punta/muestra se monitoriza reflejando un láser en la parte trasera de la palanca, que se recoge en un detector (Figura 1.1). Esta técnica permite obtener el perfil de una superficie, pero con una resolución que habitualmente supera en varios ordenes de magnitud el tamaño de las moléculas que conforman dicho sistema, por lo que resulta imposible obtener información detallada de lo que sucede a nivel molecular/atómico (Figura 1.2).



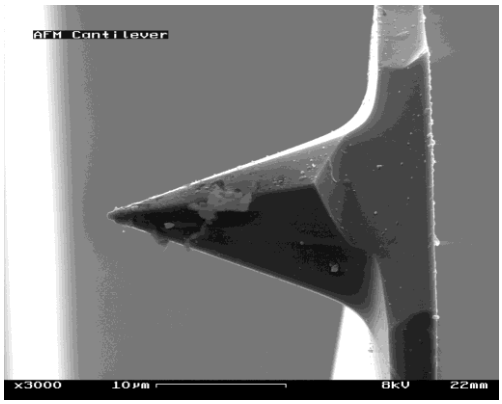


Figura 1.1. (a) Diagrama de un microscopio de fuerza atómica. (b) Ampliación a 3000× de una palanca usada en medidas de AFM.

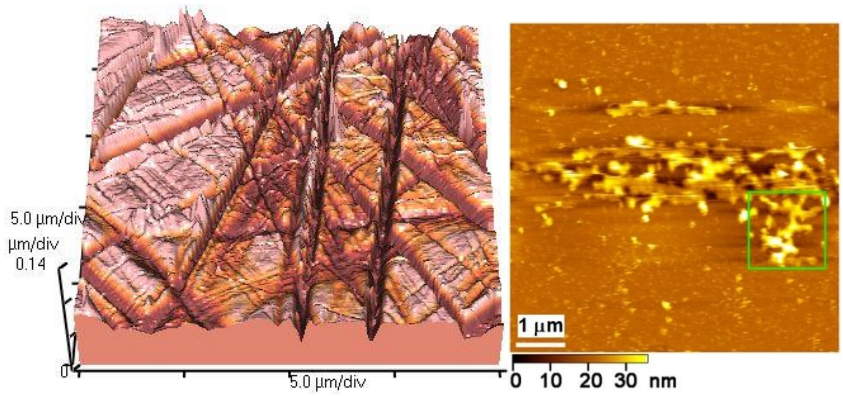


Figura 1.2. (a) Imagen de AFM de una superficie de cristal. (b) Imagen de AMF de una sección del aparato de Golgi obtenido de células HeLA.

Existe una variante del AFM, el microscopio de efecto túnel (STM: *Scanning tunneling microscope*),[5–7] que consiste en tener una diferencia de potencial entre la punta de escaneo y la superficie donde se encuentra la muestra, permitiendo el paso de electrones a través del efecto túnel (Figura 1.3). En este caso, los valores de corriente registrados dependerán del voltaje aplicado y las propiedades electrónicas de los orbitales de la muestra. Para un STM, se considera que una buena resolución es 0.1 nm en medidaslaterales y 0.01 nm en medidas de profundidad. Tanto con el STM como con el AFM se han observado enlaces covalentes de pequeñas moléculas[8–13] (en el caso del STM, el reflejo de las propiedades de los orbitales moleculares enlazantes) (Figura 1.4). Sin embargo, para alcanzar dichas resoluciones, además del tamaño molecular, se requieren condiciones muy concretas, que rara vez se dan en los sistemas de interés químico y/o biológico, como por ejemplo: cierta conductividad (en el caso del STM), superficies

formadas por planos cristalinos perfectos, alto vacío, y temperaturas entre los 5 y 11 K (Figura 1.5).

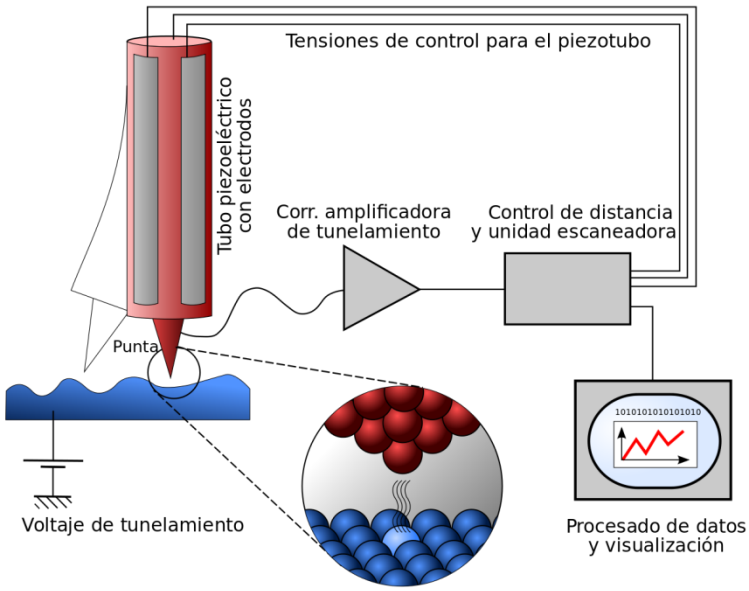


Figura 1.3. Diagrama de microscopio de efecto túnel.

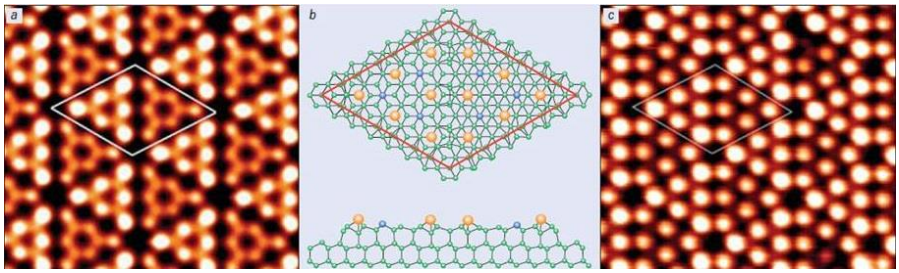


Figura 1.4. (a) y (c).[14] Imágenes recogidas mediante STM de Si(111) – 7×7 con un voltaje de -0.57 y -1.5 V, respectivamente. (b) Las estructuras atómicas calculadas para la superficie de Si(111) – 7×7, vistas superior y lateral, arriba y abajo respectivamente.

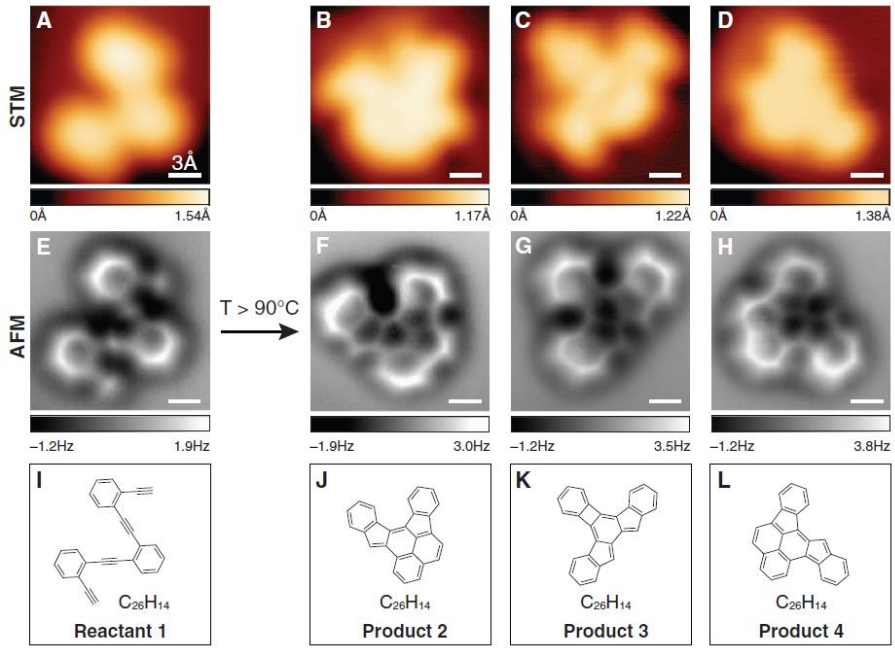


Figura 1.5.[8] Comparación entre imágenes obtenidas a través de STM, AFM y las estructuras tanto del reactante, izquierda, como de sus productos, derecha. Los productos son obtenidos después de calentar a temperaturas superiores a 90 ºC, pero las imágenes son tomadas a temperaturas criogénicas (≤ 7 K).

En resumen, para poder observar experimentalmente la materia a escalas atómica se necesitan, normalmente, unas condiciones extremas y/o unos sistemas con propiedades muy concretas. No existen métodos experimentales que nos permitan observar, e idealmente modificar, los fenómenos que suceden en la escala del Angstrom (Figura 1.6).

También existen otras técnicas como la Resonancia Magnética Nuclear (RMN) y la difracción de rayos (DRX) que permite deducir las interacciones y coordenadas atómicas, si bien habitualmente

requieren del uso de técnicas de modelado molecular. En el caso concreto de DRX, es posible determinar con un elevado grado de precisión las coordenadas atómicas únicamente cuando se trata de moléculas de reducido tamaño y con gran tendencia a formar monocristales.

Dado que para la mayoría de sistemas y condiciones de interés químico y biológico es difícil, sino imposible, obtener información a nivel atómico/molecular de forma directa o con la suficiente resolución, en las últimas décadas las herramientas de simulación molecular y química computacional han jugado un papel muy importante en el avance del conocimiento y racionalización de muchos procesos observados experimentalmente. La química computacional y la simulación molecular han hecho accesible y entendible a nivel atómico procesos y fenómenos tan importantes como los mecanismos de reacción,[15–19] los fenómenos de transporte,[20–25] la formación de nano- y microagregados,[26–30] el autoensamblado,[31–34] el reconocimiento molecular,[35–39] las transiciones térmicas,[40–45] la organización supramolecular de proteínas y ADN,[46–50] y las transiciones de fase[51–55], entre otros muchos.

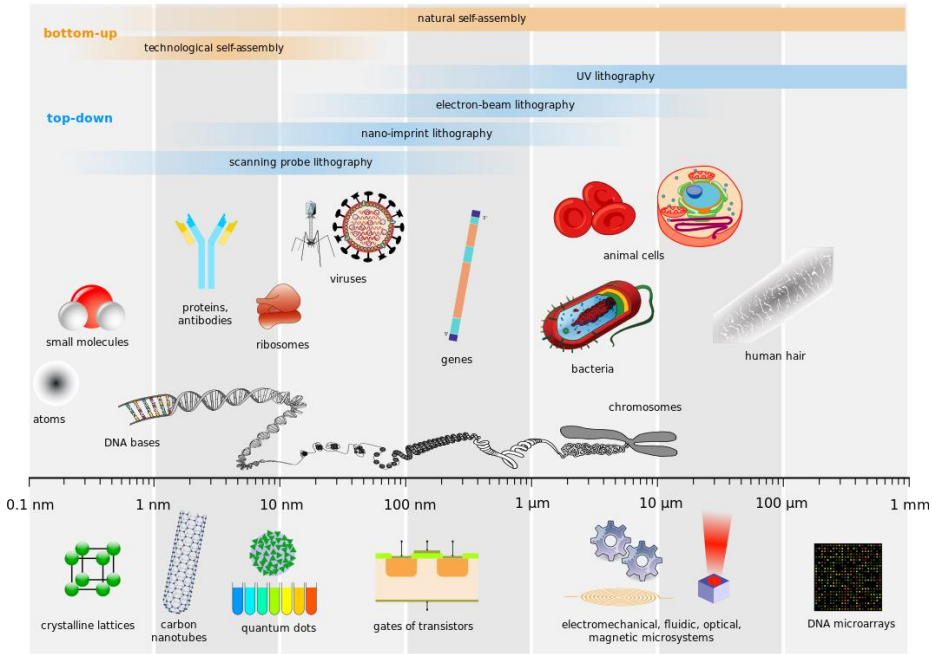


Figura 1.6. Escalas de la materia. Comparación entre objetos naturales y tecnológicos.

Esta tesis ha consistido en el estudio de diferentes sistemas de interés a través del uso de métodos de la química computacional para ayudar a la comprensión de los fenómenos que se observaban experimentalmente. Entonces, ¿qué es la química computacional?

"La química es la ciencia que se ocupa de la composición, transformación y propiedades de las moléculas. La química teórica es el sub-campo donde los métodos matemáticos son combinados con leyes fundamentales de la física para estudiar procesos de relevancia química". "La química computacional es un subcampo

de la química teórica, siendo su foco principal el resolver problemas relacionados con la química a través de cálculos".[56]

Desde que se usó el diseño sistemático de experimentos para corroborar o repudiar teorías y modelos químicos, se incrementó el entendimiento de procesos físicos y químicos con lo que la química vivió un rápido desarrollo.[57] Usando modelos matemáticos, podemos predecir el comportamiento de la materia en ciertas condiciones. Es decir, que para simular lo que sucede en un sistema químico podemos virtualmente crear una distribución representativa de átomos y moléculas, para así poder obtener información de los fenómenos físicos y químicos que probablemente se están dando o se darían en condiciones pre-definidas. Por lo que, una representación lo más precisa posible del sistema químico y el uso de un modelo adecuado que sea lo más cercano a la realidad es muy importante.

1.2 REDUCCION DE LOS REQUERIMIENTOS COMPUTACIONALES

Actualmente, la descripción de la materia está muy avanzada, lo que permite realizar simulaciones muy precisas sobre el comportamiento de los sistemas químicos. El problema surge cuando la resolución de estos modelos matemáticos, aun siendo procesados por computadores muy potentes, podrían requerir un tiempo ridículamente elevado; incluso de miles de años. En consecuencia, es necesario tener en cuenta el factor "tiempo de

cómputo", es decir, cuanto tiempo tarda el ordenador en solucionar todas las instrucciones que le son ordenadas, ya que estos han de concluir en un margen de tiempo razonable y, además, el resultado obtenido debe ser lo significativo posible. Hay diversas aproximaciones para ahorrar tiempo de cálculo sin perder relevancia en los resultados, esto dependerá básicamente de las características de nuestro sistema.

La precisión y aplicabilidad de los modelos en química computacional están limitados por una variedad de factores:[57]

1. Los grados de libertad que gobiernan los procesos químicos son electrónicos, nucleares, atómicos y moleculares. Además todas estas partículas tienen masas y tamaños muy diferentes.
2. En general, las interacciones entre estas partículas están gobernadas por la mecánica cuántica, que son las ecuaciones de Schrödinger o Dirac, o por las ecuaciones clásicas de movimiento, que son las de Newton, Lagrange, Hamilton o Langevin.
3. A temperaturas superiores a 0 K, el movimiento de las partículas está gobernado por la mecánica estadística, siguiendo las estadísticas de Bose-Einstein, Fermi-Dirac o Boltzmann. En consecuencia, es preciso considerar conjuntos de configuraciones, no estructuras únicas.
4. Comparada con la fuerte interacción entre núcleos, las interacciones entre núcleos y electrones son de largo alcance, por lo que involucra y afecta a muchos cuerpos que

hacen que un modelado preciso de las fuerzas involucradas sea computacionalmente muy costoso.

5. La energía o cambios en la energía libre asociada a los fenómenos químicos puede ser muy pequeña en comparación con la energía total de las interacciones entre las partículas involucradas en el sistema.
6. La escala de tiempo de cada proceso químico puede variar en hasta 15 órdenes de magnitud.

Todas estas características de los procesos químicos complican seriamente la formulación de modelos predictivos precisos en química computacional.

Entre las acciones más habituales para reducir el número de elementos en el cálculo computacional, y en consecuencia para disminuir el tiempo de cómputo, está la disminución del número de grados de libertad del sistema. Para ello se seleccionan, para su eliminación, los grados de libertad que son menos relevantes para el fenómeno químico objeto de estudio. Eliminar grados de libertad generalmente lleva a reducir la aplicabilidad de un modelo, pero es algo necesario y una buena elección si los parámetros de interacción clave no se ven alterados por esta acción. La eliminación o reducción de grados de libertad solo está justificado si:

1. Tienen poca importancia en el proceso o propiedades de interés.
2. Son lo suficiente numerosos para conllevar una gran reducción del coste computacional, y así compensar la pérdida de precisión.

3. Las interacciones que gobiernan estos grados de libertad son poco dependientes de las interacciones que gobiernan los otros.
4. Estos grados de libertad pueden ser descritos de forma simplificada a la vez precisa.

Una segunda acción para reducir el tiempo de cómputo es la elección del nivel de descripción de las partículas. Para ello es preciso conocer la importancia de los efectos electrónicos en el fenómeno químico objeto de estudio, siendo posible el uso de dos metodologías bien diferenciadas:

- Mecánica Cuántica: se consideran total o parcialmente los electrones del sistema, cuyas interacciones están gobernadas por primeros principios.
- Dinámica o Mecánica Molecular: se omite cualquier tipo de descripción electrónica y las interacciones entre átomos se representan mediante las leyes de la Mecánica Clásica.

1.3 MECANICA CUANTICA

"Los postulados de la mecánica cuántica forman un fundamento riguroso para la predicción de las propiedades químicas observables a partir de primeros principios. En otras palabras, los postulados fundamentales de la mecánica cuántica afirman que los sistemas microscópicos son descritos mediante "funciones de onda" que

describen todas las propiedades físicas del sistema. En particular, en mecánica cuántica hay un 'operador' para cada observable físico que, cuando es aplicado a la función de onda, permite predecir la probabilidad de que el sistema exhiba un valor particular para esa observable."[58]

Idealmente simularíamos a nivel cuántico cualquier sistema que quisiéramos estudiar, ya que es como se rige el comportamiento químico. Sin embargo, eso no es posible ya que las ecuaciones que gobiernan la teoría cuántica son intratables salvo para el sistema más ideal. Lo que sí se pueden usar son diferentes estrategias para simplificar estas ecuaciones, así como hacer uso de métodos numéricos para resolverlas.

Las funciones de onda, tal y como se mencionaba anteriormente, requieren aproximaciones para su resolución matemática. Además, el grado de aproximación es inversamente proporcional a los recursos informáticos requeridos.

Una de las aproximaciones más famosas y usadas en la química cuántica es la ecuación de Schrödinger desarrollada en 1925, la cual describe la evolución temporal de una partícula subatómica masiva de naturaleza ondulatoria y no relativista, con ecuaciones de onda para todos los electrones y el núcleo:

$$H\Psi = E\Psi \qquad (0.1)$$

Donde, H es el operador Hamiltoniano, Ψ es la función de onda y E la energía del sistema. La ecuación de Schrödinger no puede ser resuelta analíticamente excepto para átomos monoelectrónicos y, como toda aproximación, tiene sus limitaciones.

La forma típica del Hamiltoniano de la ecuación independiente del tiempo de Schrödinger para un sistema poliatómico suele ser:

$$H = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 Z_k Z_l}{r_{kl}} \quad (0.2)$$

donde i y j se refieren a los electrones, k y l a los núcleos, $\hbar$ es la constante de Plank dividida entre 2π, m_e la masa del electrón y m_k la del núcleo, ∇^2 el operador Laplaciano, e la carga del electrón, Z el número atómico, y r_{ab} la distancia entre las partículas a y b. Las cinco contribuciones al Hamiltoniano de la ecuación (1.2) corresponden a las energías cinéticas de los electrones y núcleos, la atracción de los electrones hacia el núcleo, y las repulsiones interelectrónicas e internucleares.

En 1927, Born y Oppenheimer propusieron una estrategia para separar las funciones de onda de los electrones y los núcleos considerando que los primeros se relajan varias órdenes de magnitud más rápido que el núcleo. Esta estrategia, conocida como la aproximación adiabática de Born-Oppenheimer, asume que los electrones siempre se mantienen en su estado fundamental, independientemente de la posición del núcleo ajustándolos adiabáticamente a los movimientos del núcleo.[59] De esta forma, el núcleo se rodea de una nube de carga que corresponde a los electrones, mientras que los electrones sienten a los núcleos como si estos estuvieran estáticos.

Esta aproximación permite reducir increíblemente el tiempo de cálculo de un sistema, si bien presenta algunas limitaciones en

sistemas complejos. Hay que tener en cuenta cada aproximación qué limitaciones tiene y para qué sistemas puede ser usado correctamente.

1.3.1 Hartree-Fock HF

La aproximación de Hartree-Fock (HF) se usa para solucionar la aproximación de Born-Oppenheimer de la ecuación de Schrödinger independiente del tiempo. Esta asunción consiste en que cada electrón ve a los demás como un campo electrostático promedio, algo que simplifica muchísimo los cálculos. Pero el hecho de que no se tenga en cuenta la correlación electrónica implica una serie de limitaciones importantes. Varios métodos han sido explorados, sobretodo con teorías semiempíricas, para tratar de compensar este problema. Si bien vista la clara limitación de no contemplar la correlación electrónica, que hace que las energías HF no suelan ser correctas, este método muchas veces es usado para hacer comparaciones entre sistemas.

1.3.2 Correlación electrónica

Los electrones que forman los orbitales atómicos y moleculares poseen carga y hay interacción coulómbica entre ellos, además existe el principio de exclusión de Pauli donde dos electrones con el mismo spín no pueden cohabitar el mismo orbital. En el método Hartree-Fock estos efectos no se tienen directamente en cuenta o de

forma completa. Esta correlación, muchas veces, es fundamental y el lugar por donde falla el método HF, un claro ejemplo de deficiencia del HF es para tratar las interacciones entre orbitales π. Hay una serie de métodos que buscan corregir estas deficiencias y que se llaman métodos post-HF, como lo son el método Moller-Plesset[60,61] (MP) y el método del Clúster Acoplado[62,63] (CC, *Coupled Cluster*).

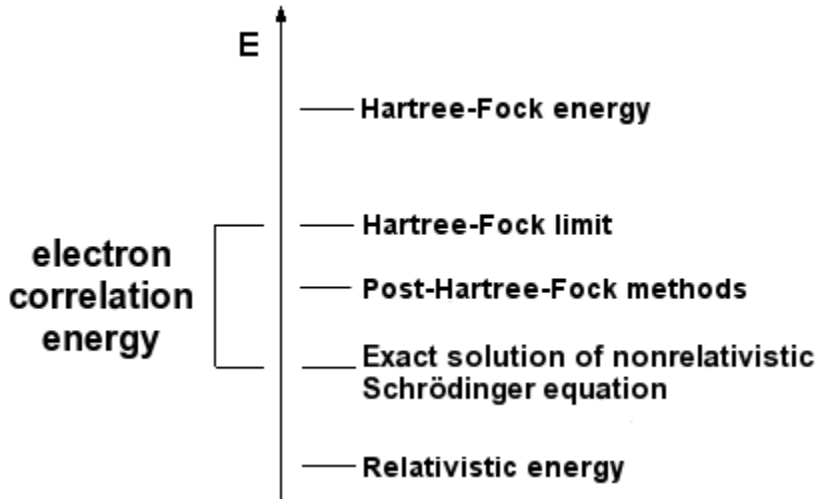


Figura 1.7. Las energías obtenidas a través de HF suelen ser más elevadas de lo que deberían al no tratar correctamente la correlación electrónica.

1.3.3 Teoría del funcional de la densidad DFT

El método de HF y sus derivados usan una función de onda donde los electrones son tenidos en cuenta independientemente, así que la función de onda depende de un spin y tres coordenadas espaciales para cada electrón (asumiendo núcleos fijos). Además, para sistemas de más de un electrón, la función de onda resultante

se vuelve algo muy complicado y poco intuitivo. La teoría del funcional de la densidad (DFT, *Density Functional Theory*) trabaja directamente con densidades electrónicas, por lo que múltiples electrones pueden ser incluidos. Este método reduce las dimensiones del sistema, permitiendo el cálculo de sistemas mayores. Una desventaja es que, salvo los casos más simples, no se conoce de manera exacta el funcional que relaciona esta densidad con la energía del sistema. En la práctica, se usan funcionales que se han comprobado que dan buenos resultados. Este método suele ser el que tiene una mejor relación coste-eficiencia, pero con el problema de encontrar la correlación entre densidad electrónica y una determinada propiedad.

1.3.4 Funciones de Base (Basis Set)

Las funciones de base son la combinación de funciones base con las que se construye la función de onda que define los orbitales moleculares. Idealmente se usaría un conjunto de funciones base infinito para definir perfectamente el orbital molecular y la probabilidad de densidad electrónica, lo que se conoce como el límite de Hartree-Fock. Las funciones base más típicas son los orbitales de Slater (*Slater-Type Orbitals*, STO) y su aproximación con Gaussianas (conocido también como "función de base mínima") y las funciones de tipo Gaussianas (*Gaussian-Type Orbitals*, GTO), también son usadas las (*Correlation-Consistent Polarized Core and Valence (Double/Triple/etc.) Zeta*, cc-pCVDZ)[64]. Además de tener en cuenta el número de funciones

usadas, también se especifica si estas son polarizables o difusas (funciones que ajustan mejor la interacción a mayores distancias), a qué átomos y orbitales se aplican, etc.

Por ejemplo de funciones de Bases Gaussianas, podemos encontrar diferentes combinaciones:[65]

- 3-21G – 3 Gaussianas primitivas para cada núcleo del orbital atómico, los primeros orbitales de valencia estarán definidos por 2 Gaussianas primitivas para la primera función de base y 1 Gaussiana primitiva para la segunda.
- 3-21G* – Añade funciones de polarización a los átomos pesados
- 3-21G** – Añade funciones de polarización a los átomos pesados e hidrógenos
- 3-21+G – Funciones difusas en átomos pesados
- 3-21++G – Funciones difusas en átomos pesados e hidrógenos
- 3-21+G* – Funciones de polarización y difusas en átomos pesados
- 3-21+G** – Funciones de polarización a los átomos pesados e hidrógenos, y también Funciones de polarización y difusas en átomos pesados e hidrógenos

1.4 DINAMICA MOLECULAR CLASICA

Haciendo la asunción de que los electrones son equilibrados instantáneamente durante los movimientos del núcleo, se puede suponer que sus propiedades e interacciones se mantienen lo suficientemente constantes como para establecer unos valores y patrones de interacción que permitan simular el comportamiento de la materia usando ecuaciones de movimiento clásicas. Esto simplifica así el sistema de tal manera que permite simulaciones a escalas atómicas de grandes sistemas (más de 1000 átomos), siendo la esencia de la dinámica molecular clásica. Es usada en sistemas donde los efectos cuánticos tienen un efecto/rol menor en el proceso de interés.

Respecto los principios fundamentales, esta metodología aprovecha que los niveles de energía electrónicos de una molécula están generalmente muy espaciados, por lo que a temperaturas fisiológicas solo los estados base van a estar poblados. Así que en una simulación atomística se considerará que los átomos y moléculas están en su estado fundamental y los grados de libertad electrónicos no se tratan explícitamente.

La dinámica molecular clásica describe las moléculas como “pelotas y muelles” donde los átomos tienen ciertas “durezas”, los enlaces asignadas “elasticidades” y los ángulos entre enlaces sobre un mismo átomo y ángulos diedros tienen también valores y ecuaciones que describen su comportamiento. Los movimientos son regidos por las leyes clásicas del movimiento, como las ecuaciones de Newton. Numéricamente, los átomos son coordenadas en el

espacio conectadas entre sí por enlaces y sus características son determinadas por valores experimentales o por cálculos de mayor nivel, también son tenidas en cuenta las interacciones intermoleculares y de mayor distancia.

El movimiento de los átomos del sistema está regido por la segunda ley de Newton:[66]

$$m_i \frac{d^2 r_i}{dt^2} = f_i = \sum_{\substack{j=i \\ (j \neq i)}}^{N} f_{ij} \qquad (0.3)$$

donde r_i es el vector de posición tridimensional para el átomo i con una masa m_i, f_i la fuerza actuando sobre él en un momento dado. La suma es sobre todos los átomos, excepto él mismo, y la tercera ley de Newton implica que $f_{ij} = -f_{ji}$, por lo que cada para de átomos solo se ha de calcular una vez. Las soluciones a la ecuación son integradas numéricamente.[67]

1.4.1 Campos de fuerza y tipos de átomos

Las ecuaciones y parámetros que rigen la interacción entre los diferentes átomos y moléculas del sistema se llama campo de fuerzas o *Force Field*. Estos campos de fuerza suelen ser desarrollados para reproducir los datos experimentales lo mejor posible. Los átomos tienen diferentes propiedades, que afectan a su capacidad de interaccionar con otros átomos, según cual sea su situación en la molécula y su hibridación.[68–70] De forma más

realista, cada átomo de cada molécula tendría sus propias características, pero eso sería realmente costoso computacionalmente y las diferencias no serían muy relevantes en muchos casos, por lo que estos se agrupan en tipos de átomos o *atom types*.

Átomo	Símbolo	Características
Carbono	C	polar (carbonilo, cadena principal péptido)
	CA	aromático
	CC	carbonilo (Asn, Asp, Gln, Glu)
	CPT	entre-anillos del triptófano
	CP1, CP2, CP3	tetraédrico especial en prolina
	CT1	alifático sp3 en CH
	CT2	alifático sp3 en CH2
	CT3	alifático sp3 en CH3
	CN1	carbono carbonílico de ácido nucleico
	CN3	carbono aromático de ácido nucleico
Oxígeno	O	carbonilo
	OC	carboxilato
	OH1	hidroxilo
	ON6	oxígeno de anillo desoxirribosa de ácido nucleico
	ON6B	oxígeno de ribosa de ácido nucleico
Nitrógeno	N	prolina
	NH1	péptido
	NH2	amida
	NN1	nitrógeno amina de ácido nucleico
	NN2	nitrógeno de anillo de ácido nucleico protonado
Hidrógeno	H	polar
	HA	no polar
	HP	aromático
	HS	tiol
	HN1	protón amina de ácido nucleico
	HN2	protón del nitrógeno de anillo de ácido nucleico

Tabla 1.1. Ejemplos de tipos de átomos definidos en CHARMM 22 y 27.[71–74]

El campo de fuerzas puede dividirse en dos tipos de contribuciones según la naturaleza de las interacciones. Estas corresponded a la interacciones enlazantes (contribuciones asociadas a la deformación de la distancia de enlace, el ángulo de enlace y el ángulo diedro formado por cuatro átomos consecutivos) y a las interacciones no enlazantes (contribuciones asociadas a las interacciones electrostáticas y de van der Waals). En general, las expresiones que emplean los diferentes campos de fuerzas para describir cada una de las contribuciones a los términos enlazantes y no enlazantes son:

$$E = E_{\text{bond}} + E_{\text{bang}} + E_{\text{tor}} + E_{\text{vdW}} + E_{\text{coul}} \tag{0.4}$$

$$E_{\text{bond}} = \sum_{i,j \in S_B} S_{ij} \left(r_{ij} - \bar{r}_{ij} \right)^2 \tag{0.5}$$

$$E_{\text{bang}} = \sum_{i,j,k \in S_{BA}} K_{ijk} \left(\cos\theta_{ijk} - \cos\bar{\theta}_{ijk} \right)^2 \tag{0.6}$$

$$E_{\text{tor}} = \sum_{ijkl \in S_{DA}} \sum_{n} \left(\frac{Vn_{ijkl}}{2} \left[1 \pm \cos\left(n\tau_{ijkl} \right) \right] \right) \tag{0.7}$$

$$E_{\text{vdW}} = \sum_{i,j \in S_{NB}} \left(\frac{-A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} \right) \tag{0.8}$$

$$E_{\text{coul}} = \sum_{i,j \in S_{NB}} \left(\frac{q_i q_j}{\varepsilon\left(r_{ij} \right) r_{ij}} \right) \tag{0.9}$$

Donde E_{bond} es el potencial asociado a la deformación de la distancia de enlace, E_{bang} el potencial que describe la deformación del ángulo de enlace, E_{tor} el potencial referente de torsión, E_{vdW} el potencial que modela las interacciones de

van der Waals, y E_{coul} el potencial que representa las interacciones electrostáticas entre los átomos que están separados por tres o más enlaces. En estas expresiones, S_B, S_{BA} y S_{DA} denotan el conjunto de enlaces químicos, ángulos de enlace y ángulos diedros. El conjunto, S_{NB}, típicamente incluye todos los, pares de átomos separados por tres o más enlaces, $(i, j), i < j$. Los parámetros de distancia de enlace y ángulo de enlace con una barra superior indican los valores de equilibrio considerados en el campo de fuerzas.

El cálculo de las interacciones entre pares de átomos es sin duda la parte más costosa computacionalmente de la dinámica molecular clásica. Por lo que reducir el número de cálculos a efectuar puede incrementar drásticamente el tiempo de cómputo. Como las partículas que están más alejadas van a tener una interacción muy pequeña en comparación con las que están más cerca, puede establecerse un límite de distancia (cut-off) a partir de la cual pueden despreciarse las interacciones no enlazantes, en particular las interacciones de van der Waals. Las interacciones electrostáticas pueden tener una mayor influencia, incluso a grandes distancias, por lo que se suele usar el *Particle Mesh of Eward* (PME).[78]

El algoritmo SHAKE[79] mantiene las distancias de enlace entre los átomos de hidrógeno y átomos pesados en su valor de equilibrio, lo cual también resulta beneficioso para reducir el tiempo de cálculo.

1.4.2 Condiciones Periódicas de Contorno

Una simulación molecular siempre involucra un número relativamente pequeño de atómos o moléculas, comparado con el número de Avogadro, que es el orden de magnitud de constituyentes de una muestra macroscópica. El volumen de la caja de simulación ha de elegirse adecuadamente para que las densidades se correspondan con las experimentales, Por tanto, si se quiere reproducir lo que ocurre en el seno de un sistema químico, es necesario escoger condiciones de frontera que reproduzcan un entorno infinito de partículas rodeando el sistema, minimizando así los efectos de borde de la caja de simulación. Las condiciones periódicas de contorno generan exactamente este efecto; se eliminan los bordes imaginando que el sistema, celda principal, se encuentra rodeado por un número infinito de sistemas idénticos al de interés, celdas imagen. La Figura 1.8 ilustra esquemáticamente un sistema bidimensional con condiciones periódicas, la caja central representa la celda principal y las cajas adyacentes las celdas imagen, de manera que cuando una partícula sale de la celda principal por la derecha, su imagen reaparece por la izquierda. De esta manera, cada partícula con coordenadas ***ri*** en la celda principal poseerá un número infinito de imágenes periódicas ***ri***(n).

En general, la energía de un sistema periódico tiene en cuenta las interacciones de todas las imágenes periódicas de la celda principal, por lo tanto la energía se convierte en una suma infinita. Afortunadamente, en interacciones de corto alcance el potencial V se anula para valores mayores al valor de corte, r_c. Por tanto, el cálculo de la energía será sólo sobre las imágenes de la celda

principal que cumplan esta condición. Si el radio de corte no rebasa la longitud más corta de la celda principal dividida entre dos, la aproximación es conocida como convención de mínima imagen.

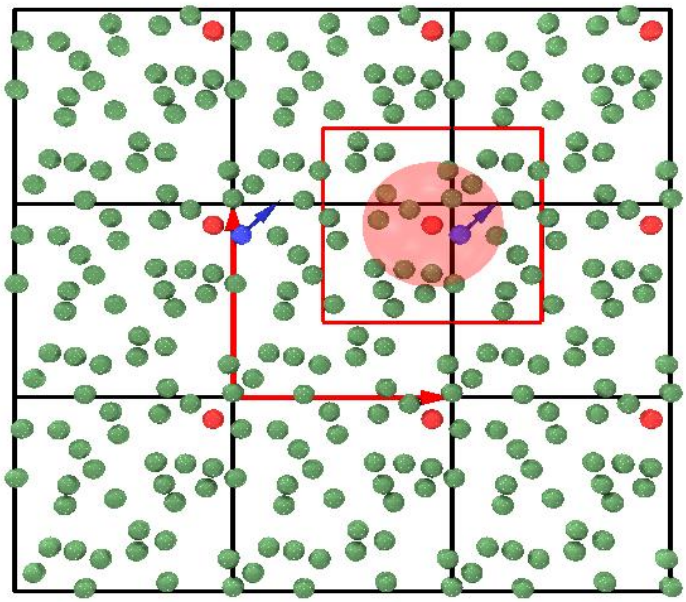


Figura 1.8. Representación de las condiciones periódicas de contorno y el criterio de mínima imagen.

1.4.3 Condiciones Fisicoquímicas de Simulación

Al modelar un sistema químico se genera una caja de simulacioón que contiene todas las partículas que conforman el mimo. Al margen de las ya mencionadas condiciones periódicas de contorno, las condiciones fisicoquímicas en las que se llevará a cabo la simulación vienen definidas por las variables que definen su estado termodinámico. Ejemplos de estas condiciones son:

- Condiciones canónicas (NVT), donde el número de partículas N, el volumen V y la temperatura T se mantienen constantes.
- Condiciones isobáricas-isotérmicas (NPT), donde el número de partículas N, la presión P y la temperatura T se mantienen constantes.
- Condiciones isobáricas-isoentálpicas (NPH), donde el número de partículas N, la presión P y la entalpia H se mantienen constantes.
- Condiciones microcanónicas (NVE), corresponde a un sistema cerrado o aislado donde la energía E, presión P y temperatura T se mantienen constantes.
- Condiciones gran-canónicas (μVT), donde el potencial químico μ, el volumen V y la temperatura T se mantienen constantes.

Habitualmente, las simulaciones de sistemas químicos se realizarán en condiciones NVT o NPT, ya que son las más frecuentes. Son las que han sido usadas en esta tesis.

1.5 BIBLIOGRAFIA

1. Binnig, G., Quate, C. F. & Gerber, C. Atomic Force Microscope. *Phys. Rev. Lett.* **56,** 930–933 (1986).
2. Lang, K. M. *et al.* Conducting atomic force microscopy for nanoscale tunnel barrier characterization. *Rev. Sci. Instrum.* **75,** 2726–2731 (2004).
3. Cappella, B. & Dietler, G. Force-distance curves by atomic force microscopy. *Surf. Sci. Rep.* **34,** 1–104 (1999).
4. Roiter, Y. & Minko, S. AFM single molecule experiments at the solid-liquid interface: In situ conformation of adsorbed flexible polyelectrolyte chains. *J. Am. Chem. Soc.* **127,** 15688–15689 (2005).
5. Binnig, G. & Rohrer, H. Scanning tunneling microscopy. *Surf. Sci.* **126,** 236–244 (1983).
6. Bonnell, D. A. *Scanning probe microscopy and spectroscopy : theory, techniques, and applications*. (Wiley-VCH, 2001).
7. Bai, C. *Scanning Tunneling Microscopy and Its Application*. (Springer, 2000).
8. De Oteyza, D. G. *et al.* Direct imaging of covalent bond structure in single-molecule chemical reactions. *Science (80-.).* **340,** 1434–1437 (2013).
9. Onoda, J., Yabuoshi, K., Miyazaki, H. & Sugimoto, Y. High-resolution imaging of silicene on an Ag(111) surface by atomic force microscopy. *Phys. Rev. B* **96,** 1–5 (2017).
10. Cocker, T. L., Peller, D., Yu, P., Repp, J. & Huber, R. Tracking the ultrafast motion of a single molecule by femtosecond orbital imaging. *Nature* **539,** 263–267 (2016).
11. Mohn, F., Gross, L., Moll, N. & Meyer, G. Imaging the charge distribution within a single molecule. *Nat. Nanotechnol.* **7,** 227–231 (2012).
12. Gross, L. *et al.* High-resolution molecular orbital imaging using a p-wave STM tip. *Phys. Rev. Lett.* **107,** 1–4 (2011).
13. Iwata, K. *et al.* Chemical structure imaging of a single molecule by atomic force microscopy at room temperature. *Nat. Commun.* **6,** 1–7 (2015).
14. Wang, Y. L. *et al.* Tip size effect on the appearance of a STM image for complex surfaces: Theory versus experiment for $Si(111)-(7\times7)$. *Phys. Rev. B* **70,**

073312 (2004).
15. Dryer, F. L. & Westbrook, C. K. Simplified Reaction Mechanisms for the Oxidation of Hydrocarbon Fuels in Flames. *Combust. Sci. Technol.* **27,** 31–43 (1981).
16. Kazemi, M. & Åqvist, J. Chemical reaction mechanisms in solution from brute force computational Arrhenius plots. *Nat. Commun.* **6,** 7293 (2015).
17. Davies, D. L., Donald, S. M. A. & Macgregor, S. A. Computational study of the mechanism of cyclometalation by palladium acetate. *J. Am. Chem. Soc.* **127,** 13754–13755 (2005).
18. Cheng, G. J., Zhang, X., Chung, L. W., Xu, L. & Wu, Y. D. Computational organic chemistry: Bridging theory and experiment in establishing the mechanisms of chemical reactions. *J. Am. Chem. Soc.* **137,** 1706–1725 (2015).
19. Yang, T. & Ehara, M. Computational Studies on Reaction Mechanism and Origins of Selectivities in Nickel-Catalyzed (2 + 2 + 2) Cycloadditions and Alkenylative Cyclizations of 1,6-Ene-Allenes and Alkenes. *J. Org. Chem.* **82,** 2150–2159 (2017).
20. Ravi A. Vijayendran, †, Frances S. Ligler, ‡ and & Deborah E. Leckband†, §,*. A Computational Reaction−Diffusion Model for the Analysis of Transport-Limited Kinetics. (1999). doi:10.1021/AC990672B
21. Slepchenko, B. M., Schaff, J. C., Macara, I. & Loew, L. M. Quantitative cell biology with the Virtual Cell. *Trends Cell Biol.* **13,** 570–576 (2003).
22. Roux, B. & Schulten, K. Computational Studies of Membrane Channels. *Structure* **12,** 1343–1351 (2004).
23. Berning, T., Lu, D. M. & Djilali, N. Three-dimensional computational analysis of transport phenomena in a PEM fuel cell. *J. Power Sources* **106,** 284–294 (2002).
24. Clark, D. E. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 1. Prediction of intestinal absorption. *J. Pharm. Sci.* **88,** 807–814 (1999).
25. Beard, D. A. A Biophysical Model of the Mitochondrial Respiratory System and Oxidative Phosphorylation. *PLoS Comput. Biol.* **1,** e36 (2005).
26. Whitesides, G. M. *et al.* Noncovalent Synthesis: Using Physical-Organic Chemistry to Make Aggregates. *Acc.*

Chem. Res. **28,** 37–44 (1995).
27. Stathopulos, P. B. *et al.* Sonication of proteins causes formation of aggregates that resemble amyloid. *Protein Sci.* **13,** 3017–3027 (2008).
28. Marrink, S. J., Lindahl, E., Edholm, O. & Mark, A. E. Simulation of the spontaneous aggregation of phospholipids into bilayers [23]. *J. Am. Chem. Soc.* **123,** 8638–8639 (2001).
29. Ma, B. & Nussinov, R. Molecular dynamics simulations of alanine rich β-sheet oligomers: Insight into amyloid formation. *Protein Sci.* **11,** 2335–2350 (2009).
30. Daniel, M. C. & Astruc, D. Gold nanoparticles: Assemble, supramolecular chemistry, quantum-size-related properties, and applications toward biology, catalysis, and nanotechnology. *Chem. Rev.* **104,** 293–346 (2004).
31. Olenyuk, B., Whiteford, J. A., Fechtenkötter, A. & Stang, P. J. Self-assembly of nanoscale cuboctahedra by coordination chemistry. *Nature* **398,** 796–799 (1999).
32. Whitesides, G. M., Mathias, J. P. & Seto, C. T. Molecular self-assembly and nanochemistry: a chemical strategy for the synthesis of nanostructures. *Science* **254,** 1312–9 (1991).
33. GAZIT, E. A possible role for pi-stacking in the self-assembly of amyloid fibrils. *FASEB J.* **16,** 77–83 (2002).
34. Zeng, F. & Zimmerman, S. C. Dendrimers in supramolecular chemistry: From molecular recognition to self-assembly. *Chem. Rev.* **97,** 1681–1712 (1997).
35. Vakser, I. A. & Aflalo, C. Hydrophobic docking: A proposed enhancement to molecular recognition techniques. *Proteins Struct. Funct. Bioinforma.* **20,** 320–329 (1994).
36. Lamb, M. L. & Jorgensen, W. L. Computational approaches to molecular recognition. *Curr. Opin. Chem. Biol.* **1,** 449–457 (1997).
37. Meyer, E. A., Castellano, R. K. & Diederich, F. Interactions with aromatic rings in chemical and biological recognition. *Angew. Chemie - Int. Ed.* **42,** 1210–1250 (2003).
38. Klärner, F. G. & Kahlert, B. Molecular Tweezers and Clips as Synthetic Receptors. Molecular Recognition and Dynamics in Receptor-Substrate Complexes. *Acc. Chem. Res.* **36,** 919–932 (2003).
39. Jiménez-García, B. *et al.* LightDock: A new multi-scale approach to protein-protein docking. *Bioinformatics* **34,** 49–55 (2018).

40. Vyazovkin, S. Evaluation of activation energy of thermally stimulated solid-state reactions under arbitrary variation of temperature. *J. Comput. Chem.* **18,** 393–402 (1997).
41. Keck, J. C. Variational Theory of Reaction Rates. in 85–121 (Wiley-Blackwell, 2007). doi:10.1002/9780470140154.ch5
42. Frenkel, D., Mooij, G. C. A. M. & Smit, B. Novel scheme to study structural and thermal properties of continuously deformable molecules. *J. Phys. Condens. Matter* **4,** 3053–3076 (1992).
43. Houk, K. N., Li, Y. & Evanseck, J. D. Transition Structures of Hydrocarbon Pericyclic Reactions. *Angew. Chemie Int. Ed. English* **31,** 682–708 (1992).
44. Banerjee, S. *et al.* An experimental and computational investigations of photosensitive Schottky barrier diode property of an azobenzene based small organic molecule. *New J. Chem.* (2018). doi:10.1039/C8NJ02193G
45. Frandsen, B. N., Skov, A. B., Cacciarini, M., Nielsen, M. B. & Kjaergaard, H. G. Computational and Experimental Evidence of Two Competing Thermal Electrocyclization Pathways for Vinylheptafulvene. *Chem. - An Asian J.* (2018). doi:10.1002/asia.201800437
46. Lehn, J.-M. Supramolecular Chemistry—Scope and Perspectives Molecules, Supermolecules, and Molecular Devices(Nobel Lecture). *Angew. Chemie Int. Ed. English* **27,** 89–112 (1988).
47. Lehn, J. M. Towards complex matter: Supramolecular chemistry and self-organization. *Eur. Rev.* **17,** 263–280 (2009).
48. Periole, X., Knepp, A. M., Sakmar, T. P., Marrink, S. J. & Huber, T. Structural determinants of the supramolecular organization of G protein-coupled receptors in bilayers. *J. Am. Chem. Soc.* **134,** 10959–10965 (2012).
49. Luca, S., Yau, W. M., Leapman, R. & Tycko, R. Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR. *Biochemistry* **46,** 13505–13522 (2007).
50. Mirkin, C. A., Letsinger, R. L., Mucic, R. C. & Storhoff, J. J. A DNA-based method for rationally assembling nanoparticles into macroscopic materials. *Nature* **382,** 607–609 (1996).
51. Flory, P. J. & Volkenstein, M. Statistical mechanics of chain

molecules. *Biopolymers* **8,** 699–700 (1969).
52. Dellago, C., Bolhuis, P. G. & Geissler, P. L. Transition Path Sampling. in *Advances in Chemical Physics* 1–78 (Wiley-Blackwell, 2003). doi:10.1002/0471231509.ch1
53. David, W. I. F., Ibberson, R. M., Dennis, T. J. S., Hare, J. P. & Prassides, K. Structural Phase Transitions in the Fullerene C_{60}. *Europhys. Lett.* **18,** 219–225 (1992).
54. Milman, V. *et al.* Electronic Structure, Properties, and Phase Stability of Inorganic Crystals: A Pseudopotential Plane-Wave Study. *Int. J. Quantum Chem.* **77,** 895–910 (2000).
55. Giovambattista, N., Rossky, P. J. & Debenedetti, P. G. Phase transitions induced by nanoconfinement in liquid water. *Phys. Rev. Lett.* **102,** 050603 (2009).
56. Jensen, F. *Introduction to Computational Chemistry Computational Chemistry*. (2007). doi:10.1007/s00214-013-1372-6
57. Meier, K. *et al.* Multi-resolution simulation of biomolecular systems: A review of methodological issues. *Angew. Chemie - Int. Ed.* **52,** 2820–2834 (2013).
58. Cramer, C. J. *Essentials of Computational Chemistry Theories and Models*. *Essentials of Computational Chemistry* **42,** (J. Wiley, 2004).
59. Gooneie, A., Schuschnigg, S. & Holzer, C. A review of multiscale computational methods in polymeric materials. *Polymers (Basel).* **9,** (2017).
60. Møller, C. & Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **46,** 618–622 (1934).
61. Binkley, J. S. & Pople, J. A. Møller–Plesset theory for atomic ground state energies. *Int. J. Quantum Chem.* **9,** 229–236 (1975).
62. Coester, F. & Kümmel, H. Short-range correlations in nuclear wave functions. *Nucl. Phys.* **17,** 477–485 (1960).
63. Paldus, J., Íek, J. & Shavitt, I. Correlation problems in atomic and molecular systems. IV. Extended coupled-pair many-electron theory and its application to the BH3 molecule. *Phys. Rev. A* **5,** 50–67 (1972).
64. Woon, D. E. & Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *J. Chem. Phys.* **103,** 4572–4585 (1995).

65. Ditchfield, R., Hehre, W. J. & Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **54,** 724–728 (1971).
66. Frenkel, D. and Smit, B. Understanding Molecular Simulation: From Algorithms to Applications. (2001).
67. Rapaport, D. C., Blumberg, R. L., McKay, S. R. & Christian, W. *The Art of Molecular Dynamics Simulation*. *Computers in Physics* **10,** (Cambridge university press, 1996).
68. Brooks, B. R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4,** 187–217 (1983).
69. MacKerell, A. D. *et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins †. *J. Phys. Chem. B* **102,** 3586–3616 (1998).
70. Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **7,** 230–252 (1986).
71. Foloppe, N. & MacKerell, A. D. All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **21,** 86–104 (2000).
72. MacKerell, A. D. & Banavali, N. K. All-Atom Empirical Force Field for Nucleic Acids: II. Application to Molecular Dynamics Simulations of DNA and RNA in Solution. *J. Comput. Chem.* **21,** 105–120 (2000).
73. Mackerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation. *J. Comput. Chem.* **25,** 1400–1415 (2004).
74. MacKerell, A. D. *et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins †. *J. Phys. Chem. B* **102,** 3586–3616 (1998).
75. Schlick, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide*. **21,** (Springer Science & Business Media, 2010).
76. Verlet, L. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* **159,** 98–103 (1967).
77. Allen, M. P. & Tildesley, D. J. *Computer simulation of*

liquids. (Clarendon Press, 1989).
78. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089–10092 (1993).
79. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23,** 327–341 (1977).
80. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81,** 3684–3690 (1984).
81. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52,** 255–268 (1984).
82. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31,** 1695–1697 (1985).

- Chapter 2 -

Objectives

2 OBJECTIVES

In this Thesis, different theoretical approaches designed to study a wide interval of length- and/or time-scales have been used to study the microscopic properties of chemical systems with varying degrees of complexity and size.

1) On the utilization of first principle calculations: Conformational and electronic studies of small and medium size compounds. DFT calculations combined with a smart build-up strategy have been applied to study the conformational profile of a small tripeptide derived from the RGD sequence. In addition, the optical properties of such compound have been predicted by applying time-dependent DFT (TD-DFT) calculations on the resulting DFT geometries.

2) On the utilization of classical MD simulations for studying small and medium size peptides: Influence of metallic substrates in the conformational preferences of individual and self-assembled peptides adsorbed onto rigid substrates. Atomistic MD simulations have been used to investigate different phenomena that involve small and medium size peptides. Firstly, the conformational preferences of a cell penetrating tetrapeptide, RPAR, adsorbed onto a gold substrate have been examined. More specifically, the effects of two gold surface facets, the (100) and the (111), have been exhaustively compared through clustering and conformational analyses. After this, the deposition of a pre-assembled peptide-polymer conjugate onto a mica substrate has been examined to

understand the structural phenomena experimentally observed. This system consists of a cyclic syndiotactic octapeptide, cyc[(L-Gln-D-Ala-L-Lys-D-Ala)$_2$], with a poly(*n*-butyl acrylate) blocks tethered to each Lys side chains that self-assembles into nanotube like structures. Particular attention has been given to rationalize the influence of the de-solvation process in the conformational contraction experienced by the acrylate side chains.

3) On the utilization of classical MD simulations for studying proteins in biological environment of increasing complexity: Influence of environment in the structure of re-naturalized outer membrane porins. The impact of solvation medium in the structure of a representative outer membrane porin (OMP) have been examined by using atomistic MD simulations. More specifically, a β-barrel OMP bearing 340 amino acids (OmpF) has been simulated considering three different environments: (i) the protein in aqueous solution at pH= 7; (ii) the protein in aqueous solution with a number of detergent molecules corresponding to the critical micelle concentration; and (iii) the protein embedded in a detergent bilayer, which mimics the lipid environment typically found in cellular membranes.

4) On the combination of stochastic algorithms and MD simulations to model the atomistic details of polymers coatings deposited over metal surfaces. An atomistic modeling strategy able to reproduce the experimentally observed topographic features of poly(3,4-ethylenedioxythiophene), PEDOT, deposited onto stainless

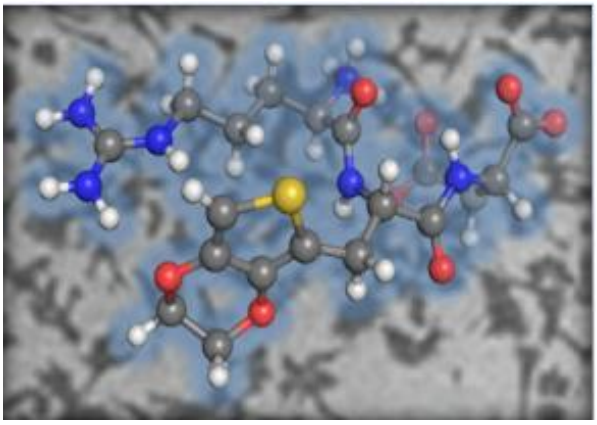
steel have been developed by combining available mathematical algorithms with classical force-fields simulations. This strategy, which has been validated by comparing experimental and predicted topographic magnitudes, is necessary for the study of bioinspired devices based on merging PEDOT and biomolecules, such as proteins.

5) Multi-phasic simulations on antimicrobial biocapacitors made of protein adsorbed onto or embedded into a conducting polymer matrix. Atomistic MD simulations have been conducted to explain the influence of protein···polymer interactions in the antimicrobial activity of lysozyme (LYZ)-containing PEDOT electrodes used for biocapacitors. Simulations have been conducted on two models: (1) a biphasic system in which the protein was simply adsorbed onto the surface of PEDOT; and (2) a biocomposite in which the LYZ was homogeneously distributed inside the PEDOT matrix. These systems have been constructed using the computational strategy developed in the objective 4. MD simulations have been performed in absence and presence of electric fields, the latter mimicking the one originated by the voltage cell difference in biocapacitors.

CHAPTER 2

- Chapter 3 -

Conformational and electronic analysis of a modified RGD adhesive sequence



3.1 CONFORMATIONAL ANALYSIS OF A MODIFIED RGD ADHESIVE SEQUENCE[§]

3.1.1 ABSTRACT

The conformational preferences of the Arg-GlE-Asp sequence, where GlE is a engineered amino acid bearing a 3,4-ethylenedioxythiophene (EDOT) ring as side group, have been determined combining density functional theory calculations with a well-established conformational search strategy. Although the Arg-GlE-Asp sequence was designed to prepare a conducting polymer-peptide conjugate with excellent electrochemical and bioadhesive properties, the behavior of such hybrid material as adhesive biointerface is improvable. Results obtained in this work prove that the bioactive characteristics of the parent Arg-Gly-Asp sequence become unstable in Arg-GlE-Asp because of both the steric hindrance caused by the EDOT side group and the repulsive interactions between the oxygen atoms belonging to the backbone amide groups and the EDOT side group. Detailed analyses of the conformational preferences identified in this work have been used to re-engineer the Arg-GlE-Asp sequence for the future development of a new electroactive conjugate with improved bioadhesive properties. The preparation of this new conjugate is in progress.

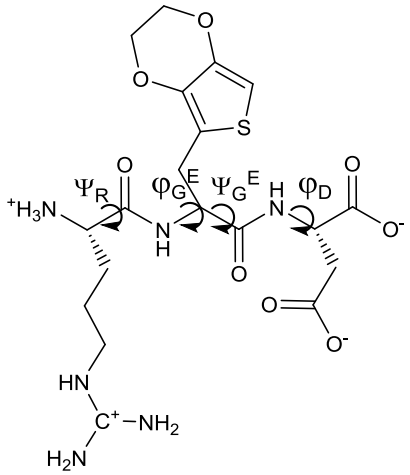
[§] This work has been published as Triguero, J., Zanuy, D. & Alemán, C. Conformational analysis of a modified RGD adhesive sequence. *J. Pept. Sci.* **23,** 172–181 (2017).

3.1.2 INTRODUCTION

Polymer-peptide conjugates result from the combination of a synthetic polymer block with a peptide.[1] This type of hybrid macromolecules usually merges the stability and properties of synthetic polymers with the unique functionality that come from the precise chemical structure of peptides. In spite of this, for a few decades research on polymer-peptide conjugates has been focused on the combination of conventional polymers, as for example poly(ethylene oxide) and poly(lactic acid), with peptides.[2–6] However, with the recent explosion of bionanotechnology, the design and fabrication of polymer-peptide conjugates with electrical and/or electrochemical responses have become expanding fields of interdisciplinary research in the last decade.[7–15]

We recently reported a new strategy for the preparation of conducting polymer-peptide conjugates,[16,17] which avoids the presence of long aliphatic linkers usually employed to connect the peptide and polymer blocks. This approach, which was initially tested with small model peptides, is based on the design of exotic synthetic amino acids bearing the main chemical groups of the polymer to facilitate the covalent conjugation between the components directly. The omission of linkers allows that the electrochemical and electrical capacities of the conducting polymer reach the surface of the biomaterial (where the peptide is allocated), which is a limitation typically found in conducting polymer-peptide conjugates containing linkers as well as in polymer/peptide biocomposites.

In a recent study we designed, synthesized and evaluated the applicability as bioadhesive interface of an electroactive RGD-based conducting polymer-peptide conjugate.[12] For this purpose, the chemical similarity strategy was used to replace the Gly residue of the RGD sequence by an exotic amino acid (denoted GlE) bearing a 3,4-ethylenedioxythiophene as side group attached to an additional methylene group. The resulting peptide, denoted RG^{E}D (Scheme 3.1), was covalently linked to the ends of poly(3,4-ethylenedioxythiophene) (PEDOT) molecules through such side group, forming the PEDOT-RG^{E}D conjugate.[12] The electrochemical and electrical properties of PEDOT were maintained in the conjugate because of the small size of the peptide but also by the utilization of the EDOT ring as a connector between the two blocks of the conjugate, both factors contributing to preserve the transport of charge. In addition, cell attachment and spreading on PEDOT-RG^{E}D was significantly promoted with respect to PEDOT. In spite of this, the conformational preferences are still unknown and, therefore, the cell recognition abilities of the RG^{E}D motif could be improved by introducing additional constraints in the design.



Scheme 3.1. Chemical structure of the RG^{E}D peptide

In this work we use a theoretical approach based on the combination of Density Functional Theory (DFT) calculations and a procedure inspired in the build-up method early developed by Scheraga and coworkers[18] to characterize the conformational profile of the RG^{E}D peptide motif and identify the corresponding bioactive conformation, which is responsible of the cell recognition function. For this purpose, minimum energy conformations have been clustered considering different criteria, as for example the dihedral angles and the interactions. In particular, we have been focused on the impact of the Gly-by-GlE replacement in the interactions, which may produce drastic alterations with respect to the conformational profile of the parent RGD peptide. Accordingly, results obtained in this work for RG^{E}D have been compared with those previously reported for RGD.[19–26]

3.1.3 METHODS

DFT calculations at the B3LYP/6-31+G(d,p) level[27,28] were performed in aqueous solution. It should be noted that the B3LYP functional describes very satisfactorily the geometries and relative energies of conformations stabilized by electrostatic or hydrogen bonding interactions.[29] The solvent was described as a dielectric medium using a well stablished Self-Consistent Reaction Field (SCRF) method, that is, the Polarizable Continuum Model (PCM) of Tomasi and co-workers.[30]

The conformational potential energy surface of the RG^ED peptide was systematically explored using a procedure inspired in the build-up method of Scheraga and coworkers.[18] This approach is based on the assumption that short-range interactions play a dominant role in determining the conformation of a given peptide. Accordingly, the accessible conformations of a given peptide result from the combination of *N* independent rotamers (*i.e.* the rotational isomeric approximation),[31] in which each independent rotational state corresponds to the most favored conformation of each residue. In practice, accessible starting geometries of RG^ED were constructed by combining all the minima identified for each of the three involved residues (*i.e.* Arg, GlE and Asp). Specifically, the minima of the GlE, Asp and Arg individual residues were taken from the literature.[17,32,33] All Quantum mechanics computations have been realized with Gaussian09 software.[34]

The radius of gyration (R_g) and the hydrodynamic radius (R_H) have been computed using the HYDROPRO program.[35,36] With respect to the analysis of the interactions, the non-directional (*i.e.*

unspecific) interaction between two charged groups have been recognized as salt bridge when the distance between their main atoms is lower than 4.5 Å.[37,38] In contrast, directional (*i.e.* specific) N–H···O hydrogen bonding interactions have been defined to occur when the H···O distance is ≤ 3.0 Å and the ∠N–H···O angle is ≥ 120º.[39,40] Finally, the formation of stabilizing N–H···π interaction has been considered when the distance between the hydrogen atom of the N–H moiety and the center of the thiophene (Th) ring, H···Th, is ≤ 3.5 Å.[29]

3.1.4 RESULTS

Geometry optimization of all the starting structures constructed using the build-up strategy described above led to 433 different minimum energy conformations. Two minima were considered different when they differ in at least one of their backbone or side chain dihedral angles by more than 5º. Figure 3.1a shows the distribution of energies, which are relative to that of the global minimum, for the different minimum energy conformations considering intervals of 1 kcal/mol. As it can be seen, the highest energy borders is 26 kcal/mol. However, the most striking finding of Figure 3.1a corresponds to the bimodal shape suggested by this distribution. Thus, the bimodal distribution, which is schematically illustrated by profile with two Gaussian functions, indicates that minima can be categorized in two groups. Intramolecular interactions are significantly stronger for the minima associated to the Gaussian centered at the lower energy interval than for those contained in the highest Gaussian peak (see below).

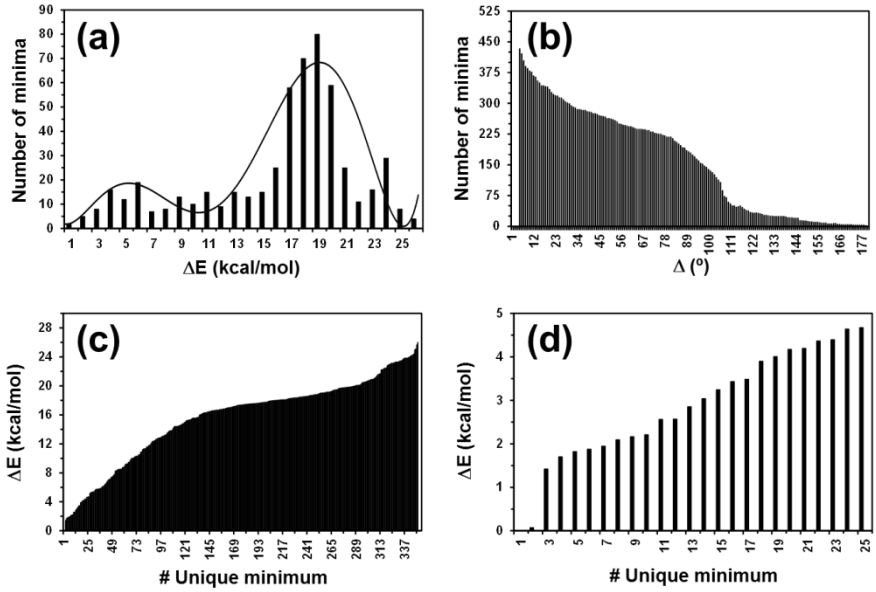


Figure 3.1. Minima were considered unique when they differ in at least one of their backbone or side chain dihedral angles by more than Δº. (a) Distribution of relative energies (ΔE) for the unique minimum energy conformations found for the $RG^{E}D$ sequence considering Δ= 5º. Represented ΔE values are relative to the global minimum. The profile has been approximated to schematically illustrate that minima follow a bimodal distribution. (b) Variation of the number of unique minima against Δ, which has been varied from 5º to 180º in steps of 1º. (c) ΔE of each of the unique minima defined using Δ= 15º against the position of such minima in the energy ranking list. The same representation is provided in (d) but considering the unique minima with ΔE < 5 kcal/mol only.

In order to evaluate the differences among the 433 minima represented in Figure 3.1a, structures have been re-classified as different minima considering that they differ in at least one of their backbone or side chain dihedral angles by more than a value defined as Δ (in degrees). Figure 3.1b displays the number of minima that differ in at least one of their dihedral angles by more

than Δ, where Δ has been varied between 5º and 180º in steps of 1º. As it can be seen, the number of unique minima decreases with increasing Δ following an almost linear behavior. However, the slope associated to such linear tendency experiences successive reductions when Δ reaches 15º, 30º and 50º, suggesting that the effectivity of the clustering process decreases at such Δ values (*i.e.* clusters with conformational differences lower than the corresponding Δ value disappear). When Δ= 80º the slope sharply increases, indicating that Δ is too large and, therefore, conformations with marked differences are considered equivalent. According to this result, unique minimum energy conformations were defined considering Δ= 15º, the number of structures decreasing from 433 for Δ= 5º to 349 for Δ= 15º.

The list of unique minima with Δ= 15º was organized by ranking them in an increasing order of energy. Figure 3.1c represents the relative energy of each minimum (ΔE), which was calculated with respect to the global minimum, against its position within the list. Although the gap between two consecutive unique minima in the list (ΔΔE) is in average 0.07 kcal/mol only, ΔΔE fluctuates between < 0.1 and 1.34 kcal/mol. Interestingly, the latter value corresponds to the difference between the relative energies of the second and third minima in the list, whereas the difference between the relative energy energies of the first (lowest) and the second minimum is approximately the above mentioned average value. These features are reflected in Figure 3.1d, which shows that only 25 minima remained within a threshold interval of ΔE < 5 kcal/mol.

Analysis of root mean square deviation (RMSD) was carried out to examine the main differences between the 349 unique conformations defined with Δ= 15º. The RMSD was calculated with respect to the global minimum considering the C^{α}, C^{β} or C^{γ} atoms of the three residues to evaluate if differences come from the backbone (C^{α}) or the side chain (C^{β} and C^{γ}) dispositions. Results, which have categorized considering groups of conformations with RMSD ≤ 0.25, 0.50, 0.75 and 1.00 Å, are displayed in Figure 3.2a. As it can be seen, many of the unique conformations present RMSD values lower than 0.25 Å (64%, 60% and 55% for C^{α}, C^{β} or C^{γ}, respectively), which is not surprising due to both the small size of the peptide and the charged nature of two of the residues. In contrast, the number of unique conformations with RMSD values comprised between 0.75 and 1.00 Å is of only 4-6%. This behavior suggests that the differences among the unique conformations mainly correspond to local re-arrangements that promote the formation of a relatively reduced number of well-defined interactions, probably those involving the charged side groups of Arg and/or Asp.

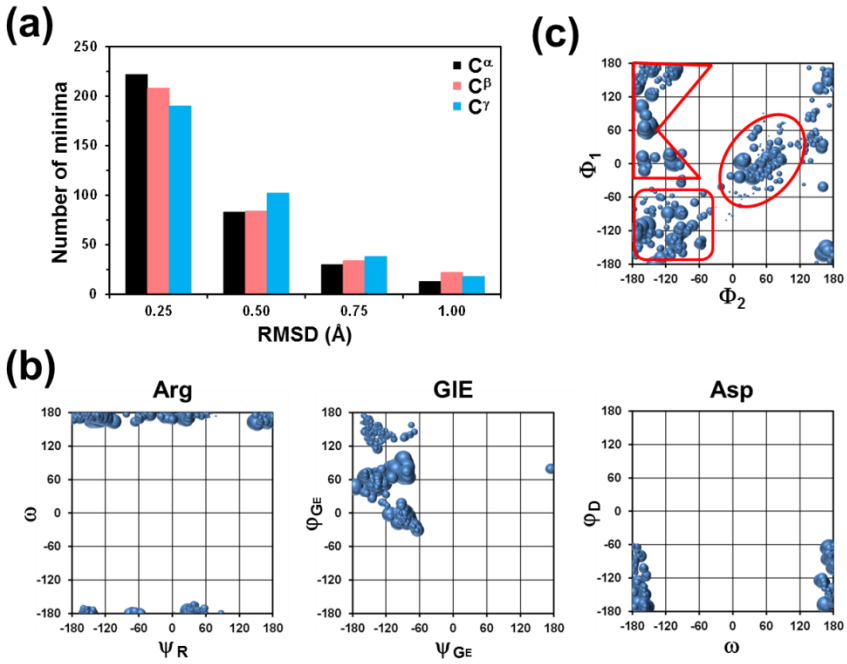


Figure 3.2. (a) Classification of the 349 unique minima identified with Δ= 15º according to their root mean square deviation (RMSD) with respect to the global minimum. The RMSD has been calculated considering the C^{α}, C^{β} or C^{γ} atoms of the three residues, conformations being categorized in groups of conformations with RMSD ≤ 0.25, 0.50, 0.75 and 1.00 Å. (b) {ω,ψ_R}, {$\varphi_G{}^E$,$\psi_G{}^E$} and {φ_D,ω} maps constructed the 349 minima identified for the RG^ED sequence. Flexible dihedral angles are defined in Scheme 3.1, while ω refers to the H-N-C^{α}(Arg)-C(=O) or N-C^{α}(Asp)-C(=O)-O dihedral angles (left and right, respectively). (c) {Φ_1,Φ_2} map constructed the 349 minima identified for the RG^ED sequence. Φ_1 and Φ_2 correspond to the virtual dihedral angles defined by C^{ζ}(Arg)–C^{α}(Arg)–C^{α}(Asp)–C^{γ}(Asp) and C^{β}(Arg)–C^{α}(Arg)–C^{α}(Asp)–C^{β}(Asp), respectively. The size of the balls decreases with increasing ΔE in (b) and (c).

The restricted conformational flexibility of the RG^ED is confirmed in Figure 3.2b, which represents the {ω,ψ_R}, {$\varphi_G{}^E$,$\psi_G{}^E$} and {φ_D,ω} map for the Arg, GlE and Asp residue (Scheme 3.1),

respectively, considering the 349 unique minima identified with Δ= 15º. As it can be seen, although the flexible ψ_R angle is able to adopt practically all possible values, the values of dihedral angles associated to the GlE (φ_{G^E} and ψ_{G^E}) are confined around the β-sheet, γ-turn and α_R regions. Similarly, the values of φ_D are restricted in the range between -60º and -180º. On the other hand, the relative orientation of the charged ends of the Arg and Asp side chains has been determined by representing the Φ_1,Φ_2 map (Figure 3.2c), where Φ_1 and Φ_2 correspond to the virtual dihedral angles defined by C^{ζ}(Arg)–C^{α}(Arg)–C^{α}(Asp)–C^{γ}(Asp) and C^{β}(Arg)–C^{α}(Arg)–C^{α}(Asp)–C^{β}(Asp), respectively. It is worth noting that Φ_1 defines the relative orientation of the two charged side ends and, therefore, is important for optimizing the interactions in which such groups participate, while Φ_2 indicates the regions of the conformational space available to the side chains. Three relative large regions are clearly identified as the more populated (marked with red lines) evidencing several interaction patterns with major differences. Probably, such differences are due to changes in the groups participating in such interactions (*e.g.* two side groups or a side group interacting with a charged terminus).

In order to understand the conformational preferences of RG^ED, the formation of three different families of interactions have been analyzed for the unique minima with Δ= 15º: (*i*) salt bridges involving charged group located at the ends of the side groups and/or the backbone; (*ii*) N–H···O hydrogen bonds; and (*iii*) N–H···π interactions.

Salt bridges were identified in 122 unique minima, which represent a 35% of the unique minima. Results derived from the analysis of these salt bridges are displayed in Figure 3.3. These interactions have been grouped in different categories, which result from the combination of the four charged groups identified in the $RG^{E}D$ sequence: the N-terminus, the C-terminus, the guanidinium group of Arg and the carboxylate side group of Asp, hereafter denoted Arg(bb), Asp(bb), Arg(sc) and Asp(sc), respectively (*i.e.* bb= backbone, sc= side chain). The Arg(bb)···Asp(sc) salt bridge is detected in 71% of the minima, while 16%, 11% and 3% of the unique conformation exhibit Arg(sc)···Asp(sc), Arg(bb)···Asp(bb) and Arg(sc)···Asp(bb) interactions, respectively. Figure 3.3b displays the variation of the number of unique conformations with salt bridges of each type against ΔE. The Arg(bb)···Asp(sc) and the Arg(sc)···Asp(sc) are only salt bridges detected in unique minima with $\Delta E \leq 5$ kcal/mol. Figure 3.3c, which depicts the structure of lowest ΔE for each group of salt bridges, reflects the unspecific nature of the salt bridge, in which groups with opposite charges interact electrostatically, independently of their relative orientation.

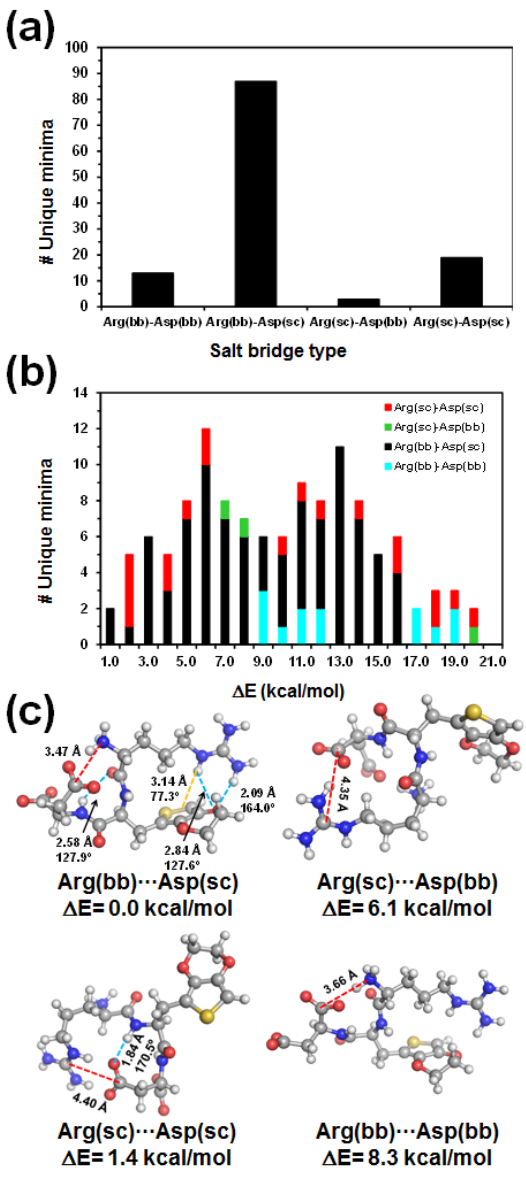


Figure 3.3. Analysis of the salt bridges interactions in the 349 unique minima identified with Δ= 15º. (a) Number of structures, from a total of 122, with Arg(bb)···Asp(bb), Arg(bb)···Asp(sc), Arg(sc)···Asp(bb) and Arg(sc)···Asp(sc) salt bridges. (b) Distribution of the number of structures with a given type of salt bridge against ΔE. (c) Lowest energy structure for each of the four types of salt bridges mentioned in (a) and (b). The distance between the interacting groups (in Å) and the ΔE (in kcal/mol) are displayed for each structure. Red, blue and yellow dashed lines correspond to salt bridges, hydrogen bonds and N–H···π interactions, respectively.

Investigation of specific N–H···O hydrogen bonding interactions, revealed the presence of eight different types:

(1) Arg(bb)N–H···O(sc)GlE; (2) Arg(sc)N^{ε}–H···O(sc)GlE;
(3) Arg(sc)$N^{\eta 1}$–H···O(bb)Arg; (4) Arg(sc)$N^{\eta 2}$–H···O(sc)GlE;
(5) GlE(bb)N–H···O(sc)GlE; (6) GlE(bb)N–H···O(sc)Asp;
(7) Asp(bb)N–H···O(bb)Arg; and (8) Asp(bb)N–H···O(sc)Asp.

N^{ε}, $N^{\eta 1}$ and $N^{\eta 2}$ correspond to the three nitrogen atoms of the guanidinium group of the Arg side chain, (bb)N and O(bb) refers to nitrogen and oxygen atoms of the backbone amide group of the indicated residue, and O(ss) is the oxygen atom of the carboxylate group in Asp. Figure 3.4a displays the structure of lowest energy for each of such eight hydrogen bond types. Interestingly, 329 of the 349 unique minimum energy conformations (94%) present stabilizing hydrogen bonds, types 8 and 6 being the most frequent ones (Figure 3.4b). The population of the other types of hydrogen bond is significantly lower than those of the ones involving the Asp(sc), which is consistent with the fact that the conformational preferences of the $RG^{E}D$ sequence is dominated by salt bridges. Amazingly, the three hydrogen bonds found in the global minimum (*i.e.* types 2, 3 and 7) and the one detected in second minimum *(i.e.* type 4), with ΔE= 0.1 kcal/mol, are rarely detected in the other minima (Figure 3.4b). However, due to the Boltzmann factor such conformations, these interaction are expected to be predominant.

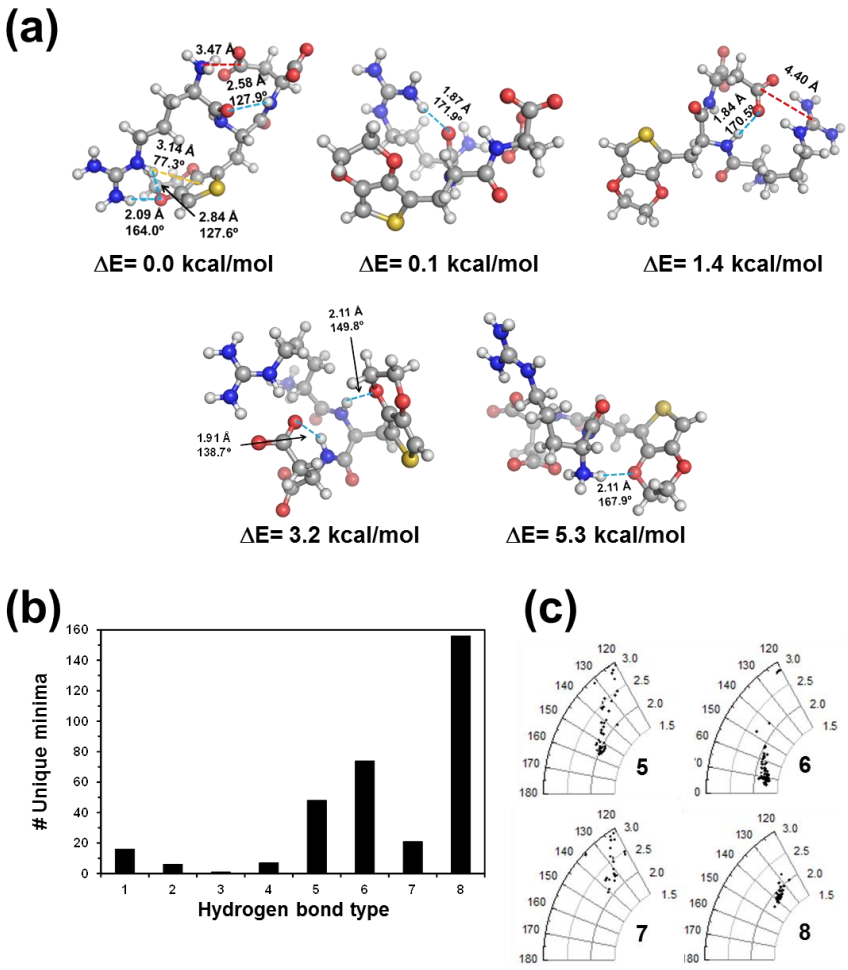


Figure 3.4. Analysis of the hydrogen bonds in the 349 unique minima identified with Δ= 15º. The eight types of hydrogen bonds detected have been labelled from 1 to 8 (see text for detailed description). (a) Lowest energy structure for each of the eight types of hydrogen bonds described in the text. Top-left: Arg(sc)N^{ε}–H···O(sc)GlE, Arg(sc)$N^{\eta 1}$–H···O(bb)Arg and Asp(bb)N–H···O(bb)Arg (labelled as 2, 3 and 7, respectively). Top-middle: Arg(sc)$N^{\eta 2}$–H···O(sc)GlE (labelled as 4). Top-right: GlE(bb)N–H···O(sc)Asp (labelled as 6). Bottom-left: GlE(bb)N–H···O(sc)GlE and Asp(bb)N–H···O(sc)Asp (labelled as 6 and 8, respectively). Bottom-right: Arg(bb)N–H···O(sc)GlE (labelled as 1). Red, blue and yellow dashed lines correspond to salt bridges, hydrogen bonds and N–H···π interactions, respectively. (b) Number of structures, from a total of 329, with at

least one hydrogen bond. (c) Polar diagrams of ∠N–H···O angles (in º) versus H···O distances (in Å) for types of hydrogen bonds with a population higher than 5%: from 5 to 8.

The hydrogen bonding geometry has been analyzed by examining the H···O distance and the ∠N–H···O angle. Figure 3.4c represents the distribution of these geometric parameters for the four hydrogen bonds types with a frequency higher than 5% (*i.e.* 15%, 23%, 6% and 47% for types 5, 6, 7 and 8, respectively). The smaller H···O distances (≤ 2 Å) correspond to hydrogen bonds of types 6 and 8 since the negative charge delocalized into carboxylate group of Asp(sc) enhances the stability of the interaction. Besides, the interaction of type 6 exhibits a geometry with ∠N–H···O values typically comprised between 160º and 170º, whereas the hydrogen bonding angle is usually ≤ 140º for types 7 and 8. However, the origin of the non-linear behavior in the latter is different: 8 is an intra-residue interaction and, therefore, the relative orientation of the two interacting groups is restricted by the bonding geometry of the own residue geometry, while in 7 the relative orientation of the two interacting backbone groups is restricted by the bulky side groups of the GlE residue.

Regarding to hydrogen bonds of type 8, an early theoretical study devoted to compare the conformational preferences of linear *vs.* cyclic RGD peptides evidenced that the capability of the Asp carboxylate side group to approach the backbone amide bond affects the stability of the peptide.[19] Thus, since the site of chemical degradation of the RGD peptide is the Asp residue, the formation of

a Asp(bb)N–H···O(sc)Asp hydrogen bond was found to promote such degradation in the linear peptide. With respect to the $RG^{E}D$ sequence, our results indicate that the bulky side group of GlE favors the formation of this kind of interaction. This has been attributed to the fact that, one of the oxygen atoms of the negatively charged carboxylate group of Asp forms a salt bridge, while the other interacts through hydrogen bond with the nearest N–H group.

N–H···π interactions involving the side thiophene (Th) group of GlE and one backbone amide group have been only detected in 7 of the 349 unique conformations with Δ= 15º. Although the N–H···π interaction only appears in the lowest energy minimum, as it displayed in Figures 3.1.3c and 3.1.4a, its population is expected to be very high because in terms of weight factor associated to such structure. In addition of the H···Th distance, the angle defined by the N–H bond and the plane of the Th ring (∠N–H···Th), is necessary to characterize N–H···π interactions. The ∠N–H···Th angle found for the lowest energy minimum is 42º, which corresponds to a tilted arrangement halfway between the parallel and perpendicular dispositions (0º and 90º, respectively).

On the basis of the results for the different interactions, the unique minima with Δ= 15º have been classified using clustering analyses with criteria based on the existence of stabilizing hydrogen bonds and salt bridges. Two structures were considered different if they differ in at least one of the interactions counted. Results, which are represented in Figure 3.5, indicate that the 349 conformers can be classified in 30 different clusters. The distribution of hydrogen bonds and salt bridges in those clusters, which have been ranked

using the lowest ΔE of the most stable conformation contained in each one, is provided in Figure 3.5a. The maximum number of interactions in a cluster is five (*i.e.* cluster 12 exhibits 4 hydrogen bonds and 1 salt bridge), whereas three clusters present four interactions (*i.e.* clusters 1, 18 and 25 exhibit 3 hydrogen bonds and 1 salt bridge). Also, there are thirteen clusters that only differ in the hydrogen bonding pattern (*i.e.* they do not present stabilizing salt bridges). The only clear relationship between the number of interactions and the stability of the cluster (Figures 3.1.5a and 3.1.5b) is that salt bridges contribute more than hydrogen bonds to the conformational, as it was expected.

Figure 3.5c represents the number of structures contained in each cluster. Cluster 23, which involves structures stabilized exclusively by Asp(bb)N–H···O(sc)Asp hydrogen bond, contains 39.5% of the analyzed minima. The second cluster in terms of population (4.9%) is the 7 that exhibits the structures stabilized by two hydrogen bonds and the Arg(sc)···Asp(sc) salt bridge. Interestingly, only the first 9 clusters display at least one structure with $\Delta E < 5$ kcal/mol. These clusters contain a total of 72 structures that corresponds to 20.6% of the clustered conformations.

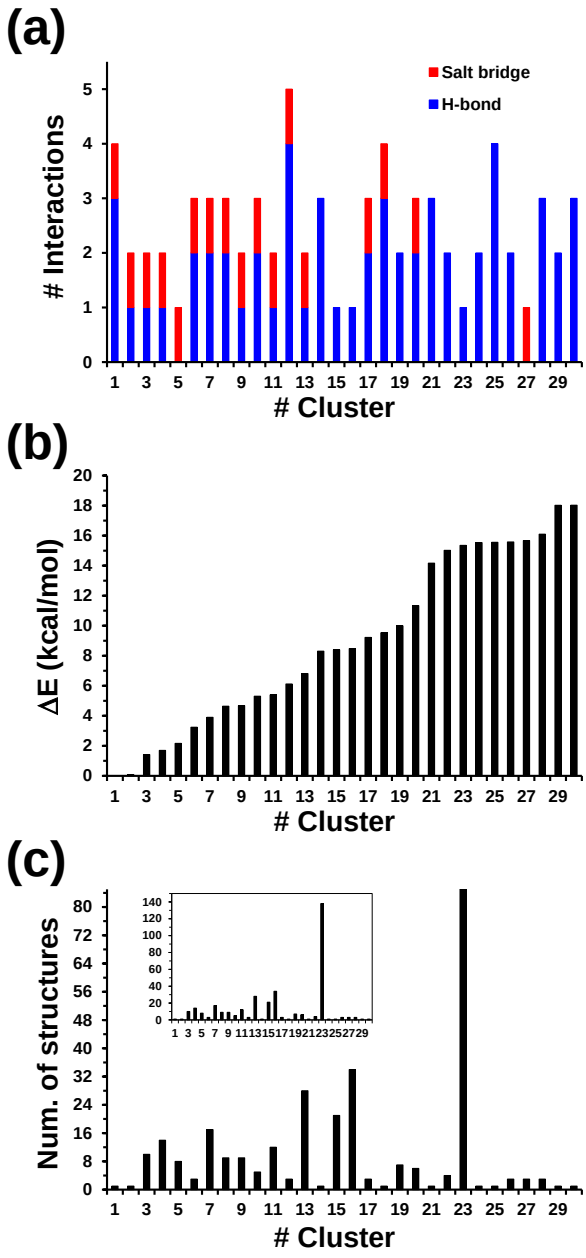


Figure 3.5. Clustering in 30 clusters according to interactions (hydrogen bonds and salt bridges) of the 349 unique minima identified with Δ= 15º. (a) Number of hydrogen bonds and salt bridges contained in each of 30 resulting clusters. (b) ΔE of the lowest energy structure contained in each cluster. (c) Number of structures contained in each cluster.

It is worth noting that the clustering analysis is consistent with the bimodal shape displayed by the distribution of ΔE values (Figure 3.1a). Thus, the Gaussian function associated to lower ΔE values correspond the unique minima containing a stabilizing salt bridge, in addition of hydrogen bonds (Figure 3.5a), while the Gaussian function at higher ΔE values involve the minima stabilized by hydrogen bonds (*i.e.* without salt bridges) and the minima with the less stabilizing salt bridges. Overall, this relationship agrees with the conclusions reached by different authors:[41–43] the single-modal gaussian distribution identified for the minima associated to simple rotational isomers transforms into multi-modal distributions when strong and well-defined non-covalent interactions perturb such ideal behavior.

The R_g and R_H values are compared in Figure 3.6. More specifically, the average, minimum and maximum values of R_g and R_H are displayed in Figure 3.6a, while the two radius are compared for each unique conformation ranked with increasing order of energy in Figure 3.6b. As it was expected, the magnitude of both structural parameters is similar, even though R_g is slightly lower than $R_{H.}$ This is because the former is defined as the weighted average distance from the core of the molecule to each mass element, while latter is defined as the radius of an equivalent hard sphere diffusing at the same rate as the molecule under observation that, in practice, means that R_H closely reflects the apparent size adopted by the solvated tumbling molecule. The R_g and R_H values displayed in Figure 3.6 (averages: R_g= 5.82±0.03 Å and R_H= 7.09±0.02 Å) reflects a proportionality constant of 0.82±0.02

between the two parameters, which is very close to the value reported for globular proteins, 0.77.[44]

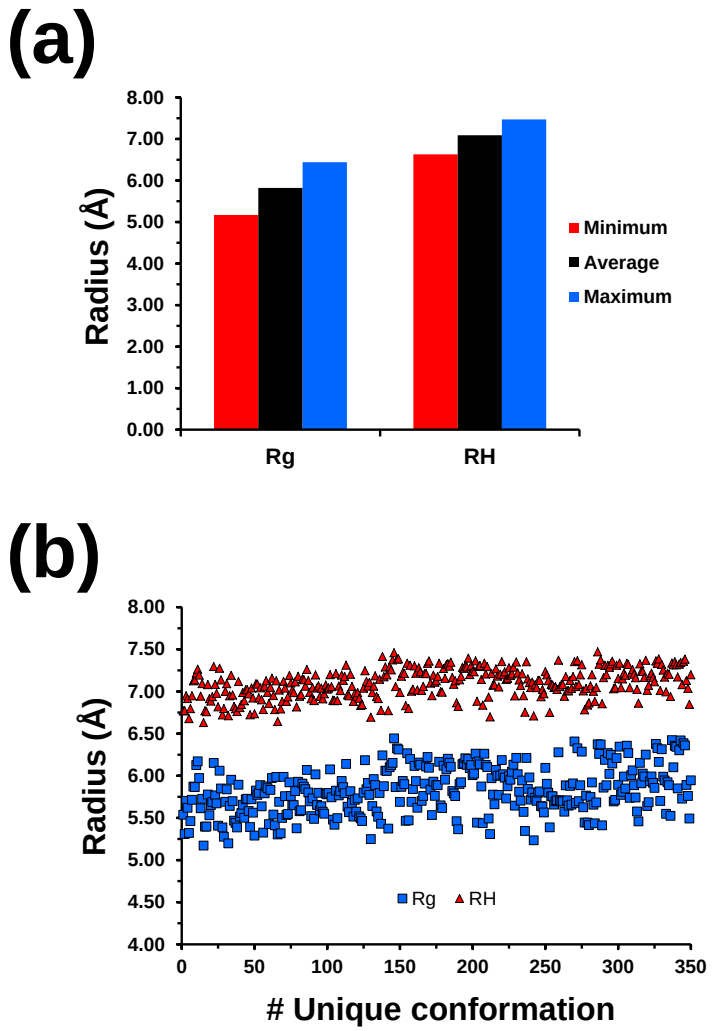


Figure 3.6. (a) Maximum, average and minimum R_g and R_H values for RG^ED. (b) R_g and R_H values for each of the 349 unique minima identified with Δ= 15º. Unique minima have been ranked according to the corresponding ΔE values.

3.1.5 DISCUSSION

In the early 1970s, Ruoslashti and Piersbacher first characterized the RGD peptide sequence, which adopts a conserved hairpin loop structure of integrin binding proteins such as fibronectin, fibrinogen and vitronectin.[45] This seminal work inspired a great number of chemists to develop scaffolds by rigidifying the conserved tripeptide motif in the desired state. For example, Kessker and co-workers[46] were able to develop a successful class of molecules by screening conformationally cyclic RGD mimics containing N-methylated residues of D-amino acids. Although investigation in bioactive cyclic RGD analogues has continued so far,[47] cell-binding linear peptides containing the RGD sequence are still frequently used. Within this context, we recently reported the synthesis, characterization and properties as bioadhesive interface of the PEDOT-RG^{E}D conjugate.[12] In this hybrid, electroactive PEDOT chains were covalently conjugated to the EDOT ring of the RG^{E}D tripeptide (*i.e.* in the latter the Gly of the RGD was replaced by GlE), the omission of linkers improving the surface properties of the system.

The conformational preferences of RGD-containing peptides have been investigates using a wide range of theoretical and experimental methods. In an early study, NMR experiments on RGD and $_{D}$RGD (*i.e.* the L-Arg of the former was replaced by D-Arg in the latter) suggested a type II' β-turn in solution, but with fast internal motions.[20] Unfortunately, determination of the alternative conformations was not possible due to the time scale of the internal motions. Complementary MD simulations reflected a

turn conformation in solution, in agreement with NMR results, even though a highly populated extended conformation was also identified.[21] Further analysis showed that the extended conformation was nearly indistinguishable from the β-turn conformation with respect to NMR interpretation and that the best fit to experimental data was from a mixture of a turn and the extended conformations.[21] On the other hand, conformational studies using circular dichroism and folding algorithms on RGDX peptides with X= Y (Tyr), W (Trp), F (Phe), L (Leu), V (Val), C (Cys), Q (Gln) and S (Ser) evidenced that the population of β-turn structure defined by the Gly-Asp sequence increases with the hydrophobicity of the X-residue.[22] In a recent study, Kulkarni and Ojha[23] combined spectroscopic ^{1}H-NMR and MD methods to study the conformational behaviour of the RRGDSPC peptide, which exhibited a two-folded state of β-turns, type II (4→1) and type II' (7→4) between Gly1 and Asp4, and Asp4 and Cys7, respectively. However, residues Asp4, Ser5 and Cys7 imparted flexibility to such two-folded β-turn state, tending to fold into additional conformations such as β-turn type III (7→4), reverse β-turn type III (4→7), γ turns and inverse γ-turns.

Besides, the RGD motif has been incorporated into peptide amphiphiles to study the self-assembly[48,49]. Among many examples, those consisting in the combination of the RGD tripeptide with an aromatic moiety are particularly relevant because of their similarity with the RG^{E}D peptide (*i.e.* the main difference between the latter and the RGD is the aromatic EDOT group). Hamley and co-workers combined the hydrophilic RGD and RGDS sequences with the

hydrophobic N-(fluorenyl-9-methoxycarbonyl) (Fmoc) blocking group to fabricate the Fmoc-RGD and Fmoc-RGDS amphiphiles. Both peptides were found to form fibrillary structures based on the β-sheet self-assembly[50,51]. Computer simulation studies evidenced that in these systems peptide molecules adopt pseudo-extended conformations because the assembly process was dominated by the π-π interactions between hydrophobic Fmoc groups.[52,53]

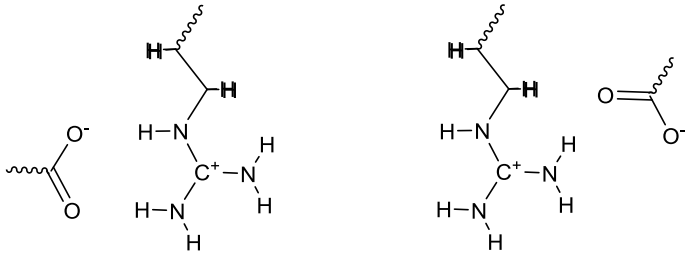
The Gly-by-GlE substitution provokes drastic changes in the conformational preferences of the RGD sequence. Detailed inspection of the $RG^{E}D$ unique conformations with $\Delta E < 5$ kcal/mol reveals the complete absence of both β-turn and extended conformations. Indeed, the γ-turn of the global minimum (Figure 3.4a) is the only folded state identified in such set of representative unique minima. The bulky EDOT side group alters the interaction pattern of the peptide as follows: *(i)* the formation of backbone···backbone N–H···O hydrogen bonds is difficult because of the size of the EDOT ring, which reduces considerably the stability of the folded conformations; *(ii)* extended and pseudo-extended conformations, including the β-strand observed in Fmoc-RGD and Fmoc-RGDS amphiphiles, are destabilized by the repulsive interaction between the oxygen atoms of the Arg carbonyl moiety and the EDOT dioxane ring; and *(iii)* the oxygen atoms of the dioxane ring and the aromatic thiophene ring interact repulsively with both the backbone carbonyl and the carboxylate side group of the Asp, destabilizing many conformations energetically accessible to the RGD sequence. Consequently, the conformational preferences of $RG^{E}D$ are mainly dictated by

inaccessibility of the interactions typically found in RGD-containing peptides, which also affects the Arg(sc)···Asp(sc) salt bridge related with the biological activity (see below).

Stote[24] investigated the bioactive conformation of the RGD sequence by comparing the conformational profiles of RGDW and $_{D}$RGDW, as obtained from MD simulations, to the conformation of the RGD sequence in cell adhesion proteins. Interestingly, the biological activity of the RGD fragment in these system was related with two aspects of the side chain orientation: *(i)* the Arg and Asp side chains are on the same side of the peptide in an almost parallel alignment; and *(ii)* the relative orientation of the charged termini of the Arg and Asp side groups is defined by a well-defined region in the Φ_1-Φ_2 map, bounded by -150º and 90º for Φ_1 and between -120º and 100º for Φ_2. The latter region, which corresponds to the sum of the regions marked by an ellipsoid and a square in Figure 3.2c, is preserved when Gly is replaced by GlE. However, a new region appears in the top-left side of the Φ_1-Φ_2 map (marked by an irregular polyhedron in Figure 3.2c) evidencing that the bulky EDOT ring perturbs the orientation of the Arg and Asp side chains. Moreover, in the global and first local minimum, the Arg and Asp side groups are not on the same side, suggesting that these cannot be associated with the bioactive conformation. Although some representative conformations with $\Delta E < 5$ kcal/mol fulfil the activity conditions identified by Stote[24], the bioactive conformations seems to be destabilized in RG^{E}D with respect to the parent RGD sequence.

Besides, Upadhyay[25] modeled the conformational preferences of the RGD motif in the PRGDMP and RIPRGDMP peptide sequences using multiple MD simulations. This author found that the presence of flanking residues around the RGD motif modulates the conformational sampling of the peptides, forcing them to adopt a non-bioactive conformation at room temperature. However, flanked peptides were able to sample the bioactive conformation when the temperature was increased to 42 ºC. Thus, the particular characteristics of the flanking residues (*i.e.* the cyclic constraint of Pro and the bulky side group of Met) restricted the conformational space of the RGD motif owing to steric hindrance and forced the system to require more energy to visit the bioactive conformation.[25] A similar feature apparently occurs when GlE replaces Gly in RG^{E}D, even though in this case the GlE residue also affects the backbone conformation.

A theoretical investigation[54] of salt bridges formed between Arg and Asp side chains revealed that the geometry with the most favoured binding energy, which is displayed in Scheme 3.2, corresponds to the symmetric "*side-on*" disposition that was also identified in the bioactive conformation of the RGD fragment.



Scheme 3.2. Symmetric side-on (left) and staggered-like (right) geometries of the Arg(sc)···Asp(sc) salt bridge.

Geometric analyses of the Arg(sc)···Asp(sc) salt bridges in $RG^{E}D$, which was displayed by 19 (5%) of the 349 unique minima with Δ= 15º (Figure 3.3a), indicated that only 9 minima (2.6%) displayed the bioactive side-on geometry (*i.e.* 7 of such 9 minima exhibited $\Delta E < 5$ kcal/mol). The other 10 minima bearing a Arg(sc)···Asp(sc) salt bridge, which were $\Delta E \geq 5$ kcal/mol, exhibited a staggered-like geometry (Scheme 3.2). Interestingly, the side-on geometry was detected in intermolecular Arg(sc)···Asp(sc) salt bridges formed by Fmoc-RGDS molecules belonging to neighboring β-sheets in 3D fibrils.[53] The most frequent salt bridge in $RG^{E}D$ corresponds to the Arg(bb)···Asp(sc) (Figure 3.3a). The preference towards this interaction, which is detected in 19 minima with $\Delta E < 5$ kcal/mol (including the global minimum), has been attributed to the destabilization of the β-turn, pseudo-extended and extended conformations discussed above.

Asakura and co-workers[26] studied the dynamics of the RGD sequence in a recombinant silk-like protein, $SLPF_{10}$, consisting of the repeated silk fibroin sequence $(AGSGAG)_3$ and the ASTGRGDSPA peptide using solution NMR. The ^{1}H, ^{15}N and ^{13}C chemical shifts indicated that the RGD moiety, as well as the silk fibroin sequence, took a random coil form with high mobility in aqueous solution. Moreover, the random coil state of the RGD fragment was maintained in dry and swollen state also. Similar features were observed for the case of RGD moiety in fibronectin. Our results indicate that this flexibility decreases when Gly is replaced by GlE. Although the two conformations of lowest energy are very different and practically isoenergectic (*i.e.* the global and

the first local minima are compared in Figure 3.4a), the ΔE of the second local minimum increases to 1.4 kcal/mol and the number of unique minima with ΔE < 5 kcal/mol is 25 only.

On the other hand, the Asp residue of RGD peptides is highly susceptible to chemical degradation, resulting in a loss of the biological activity.[19,55] The extent of this degradation depends on the degree of flexibility of the peptide, which should be restricted to preclude the Asp side chain to orient itself in the appropriate position for attack on the peptide backbone through the Asp(bb)N–H···O(sc)Asp hydrogen bond. In the case of the RG^{E}D sequence, the latter interaction is largely favored.

Overall, the RG^{E}D peptide can be re-designed to improve the cell adhesion properties of the PEDOT-RG^{E}D hybrid system. As mentioned above, the main role of the Gly-by-GlE substitution was to facilitate the covalent attachment of the PEDOT chains to the peptide through the EDOT side group of GlE, eliminating the undesirable effects induced by linkers in conducting polymer-peptide conjugates.[12] Nevertheless, our results prove that the bulky EDOT side group drastically reduces the stability of the β-turn backbone conformation and precludes the formation of the Arg(sc)···Asp(sc) salt bridge, which are both responsible of the RGD cell adhesive properties. In order to promote the stability of the pocket-like bioactive conformation in the RG^{E}D conformation the following chemical modifications are proposed: *(i)* reverse the chirality of the GlE residue; and *(ii)* replace the Arg residue by α,β-dehydro arginine ($\Delta^{\alpha,\beta}$Arg; with an unsaturation between the C$^{\alpha}$ and

C^{β} atoms) or β,γ-dehydro arginine ($\Delta^{\beta,\gamma}$Arg; with an unsaturation between the C^{β} and C^{γ} atoms).

Reverse the chirality of GlE in $RG^{E}D$ (*i.e.* L-GlE by D-GlE substitution) is expected not only to reduce the unfavorable steric effects provoked by the EDOT side group, but also to decrease the repulsive interactions between the oxygen atoms of the backbone amide bond and the EDOT side group. These two features will contribute to favor the formation of the Arg(sc)···Asp(sc) salt bridge and to increase the stability of both the β-turn and pseudo-extended conformations. On the other hand, the stiffening of the Arg side group by introducing a chemical constraint (*i.e.* Arg-by-$\Delta^{\alpha,\beta}$Arg or Arg-by-$\Delta^{\beta,\gamma}$Arg substitution) is expected to promote the mobility of the Asp side group for the formation of the Arg(sc)···Asp(sc) salt bridge, which will contribute to eliminate the undesirable Asp(bb)N–H···O(sc)Asp interactions.

3.1.6 CONCLUSIONS

In conclusion, here we have demonstrated that, although the $RG^{E}D$ sequence was engineered to facilitate the elimination of the linker in the electroactive PEDOT-$RG^{E}D$ conjugate,[12] the bioadhesive properties of the latter shoud be significantly improved by introducing some additional changes. Thus, the exhaustive conformational search performed using DFT calculations indicates that the main bioactive characteristics of the parent RGD sequence are not preferred by the $RG^{E}D$ sequence. This fact has been attributed to both the steric hindrance caused by the bulky EDOT side group and the repulsive O···O interactions between the

backbone amide groups and the EDOT ring. The conformational preferences observed for the $RG^{E}D$ sequence have been employed to propose the following changes, which are oriented towards the stabilization of a bioactive conformation similar to that of the RGD peptide. These are: *(i)* reverse the chirality of the GlE residue; and *(ii)* increase the stiffness of the Arg residue by introducing one of the following constraints $C^{\alpha}=C^{\beta}$ or $C^{\beta}=C^{\gamma}$, the choice obviously depending on the synthetic difficulties encountered during the preparation of such dehydro-Arg residues. Our next step will be focused on the exploitation of these results by preparing the hybrid material derived from the conjugation of PEDOT with the re-engineered sequence. We expect that the bioadhesive properties of this re-engineered material will be significantly better than those of PEDOT-$RG^{E}D$, evidencing the important role of in silico molecular engineering strategies.

3.2 ELECTRONIC PROPERTIES OF RG^{E}D[§,**]

3.2.2 INTRODUCTION

The RGD (Arg-Gly-Asp) amino acid sequence is the unit of a cell adhesive activity domain in adherent proteins (*e.g.* fibronectin, fibrin and vitronectin).[45,56,57] This cell adhesion motif has been widely used in the development of synthetic materials for bioengineering, different approaches being applied for such purpose (*e.g.* integration in larger peptide sequences,[50,58–60] surface functionalized substrates[11,61–64] and polymer-peptide conjugates[11,65–67]). In all cases RGD-containing biomaterials increased the cellular adhesion with respect to the control when the adhesion motif was immobilized maintaining its biological activity.

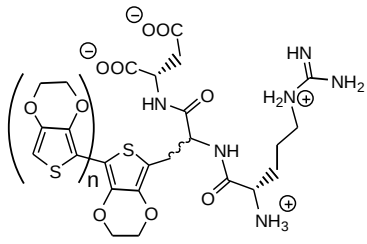
Among advanced organic biomaterials, polymer-peptide conjugates, which result from the covalent integration of a peptide with a synthetic polymer block, are especially attractive because this kind of hybrid macromolecules combines unique properties that come from the precise chemical structure and functionality of peptides and the stability, functions and processability of synthetic polymers.[5,68–70] In order to take advantage of the electrochemical properties of electroactive conducting polymers (ECPs), different ECP-peptide conjugates have been prepared during the last decade. For example, the first oligothiophene directly conjugated to a β-sheet pentapeptide

[§] This work has been published as Maione, S. *et al.* Electroactive polymer-peptide conjugates for adhesive biointerfaces. *Biomater. Sci.* **3,** 1395–1405 (2015).

was reported in 2004,[7] while the modification of the β-position on polypyrrole (PPy) to create strong disulfide bonds with the Cys of the RGDS sequence (Arg–Gly–Asp–Cys) was described in earlier works.[71,72] Within this context, we have recently reported a new strategy for the preparation of ECP-peptide conjugates that is based on the chemical similarity of their two components.[16,73] This approach is based on the design of exotic synthetic amino acids bearing the main chemical groups of the polymer to facilitate the covalent conjugation between the components. Such methodology avoids the presence of long aliphatic linkers, which are typically used to join the two components of the polymer-peptide conjugate. The absence of linkers is expected to promote the extension of the polymer properties towards the peptide surface region, which is an important limitation typically found in ECP-peptide conjugates.

In this work we have used a strategy based on chemical similarity to design an electroactive RGD-based ECP-peptide conjugate. The polymer selected for this purpose is poly(3,4-ethylenedioxythiophene), hereafter abbreviated PEDOT, which is among the most successful ECPs due to its excellent electrochemical and thermal properties, high conductivity, good environmental stability in its doped state, mechanical flexibility, relative ease of preparation, and fast doping-undoping process.[74–76] Regarding to the peptide, the Gly residue of the RGD sequence has been replaced by GlE, which consists of an amino acid bearing a 3,4-

ethylenedioxythiophene as side group attached to an additional methylene group. The resulting sequence, hereafter denoted $RG^{E}D$, was attached to end of PEDOT chains forming the PEDOT-$RG^{E}D$ conjugate (Scheme 1). The chemical and physical characterization of the conjugate as well as its behavior as soft bioelectroactive support for cell attachment have been investigated by Dr. Silvana Maione along her PhD Thesis.[77] Therefore, only the results obtained in this work by the author of this Thesis, which exclusively involve the $RG^{E}D$ peptide, are presented in this Chapter.



PEDOT-$RG^{E}D$

Scheme 3.3. Chemical structure of the PEDOT-$RG^{E}D$ conjugate

3.2.3 METHODS

3.2.3.1 Conformational analyses

For the conformational analyses, Density Functional Theory (DFT) calculations at the B3LYP/6-31+G(d,p) level were performed in aqueous solution. The solvent was described as a dielectric medium using a well stablished Self-Consistent Reaction Field (SCRF), that is, the Polarizable Continuum Model (PCM) of Tomasi and co-workers.[30]

The conformational potential energy surface of the RG^ED peptide was systematically explored using a procedure inspired in the build-up method early developed by Scheraga and coworkers.[18] This approach is based on the assumption that short-range interactions play a dominant role in determining the conformation of a given peptide. Accordingly, the accessible conformations of a given peptide result from the combination of *N* independent rotamers (*i.e.*, the rotational isomeric approximation),[31] in which each independent rotational state corresponds to the most favored conformations of each residue. In practice, accessible starting geometries of RG^ED were constructed by combining all the minima identified for each of the three involved residues (*i.e.* Arg, GlE and Asp). Specifically, the minima of GlE and Asp were taken from the literature[17,32] while those of Arg were explicitly calculated in this work. For this purpose, starting geometries for calculations on Arg were constructed using the minima previously obtained for *homo*-Arginine, that is, the Arg homologue containing one more methylene unit.[78]

Starting from the five conformers computed with PCM(water)-B3LYP/6-31++G(d,p) and lying in a 5 kcal/mol energetic gap, a re-optimization has been performed in the experimental condition (in acetonitrile) using a PCM(acetonitrile)-ωB97XD/6-311++G(d,p) level of theory that has already shown satisfying results to study conformational preferences of dipeptides[79] and EDOT residues.[17] No negative frequencies were detected. For each conformer the forty first excited states were evaluated with the CAM-B3LYP/6-311++G(d,p) level of theory and the electronic density difference

(EDD) between the excited state and the ground state have been plotted to qualitatively understand the variation of the maximum absorption wavelength (λ_{max}) depending on the conformation. All Quantum mechanics computations have been realized with Gaussian09 software.[34]

3.2.3.2 UV-vis spectroscopy

UV-vis absorption spectra were obtained using a UV-vis-NIR Shimadzu 3600 spectrophotometer equipped with a tungsten halogen visible source, a deuterium arc UV source, a photomultiplier tube UV-vis detector, and a InGaAs photodiode and cooled PbS photocell NIR detectors. Spectra were recorded in the absorbance mode using the integrating sphere accessory (model ISR-3100), the range wavelength being 200-900 nm. The interior of the integrating sphere was coated with a highly diffuse $BaSO_4$ reflectance standard. Single-scan spectra were recorded at a scan speed of 60 $nm \cdot min^{-1}$. Measurements, data collection and data evaluation were controlled by the computer software UVProbe version 2.31.

3.2.3 RESULTS AND DISCUSSION

The conformational potential energy surface of the RG^ED peptide was systematically explored using a procedure inspired in the build-up method early developed by Scheraga and coworkers.[18] This method uses the geometries of smaller peptide segments to build the structures of larger segments. Accordingly, 559 starting conformations were constructed by combining the most stable minima identified for each of the

involved residues. The geometry of all these structures was optimized at the B3LYP/6-31+G(d,p) level of theory in aqueous solution, solvent effects being modeled by the Polarizable Continuum Model (PCM). The resulting 552 minima were grouped in a list of 26 unique minimum energy conformations, which was constructed by assessing the similarities among the different structures (clustering of structures described in the Methods section). Figure 3.7a displays the distribution of such unique minima in the maps constructed using the backbone dihedral angles of Arg (ω,ψ), GlE (φ,ψ) and Asp (ω,φ) residues. As it can be seen, Arg and Asp tend to adopt an extended arrangement with ψ≈ 180º and φ≈ 180º, respectively, even though the latter also shows three minima with φ≈ 60º. Despite this apparent backbone homogeneity, both Arg and Asp present significant conformational variability at their side chains giving place to different networks of side chain···backbone interactions (Figure 3.7b). In contrast, the GlE residue exhibits remarkable conformational variability at the backbone displaying minima in the α_R (φ,ψ≈ -70º,-20º), C_{7eq} (φ,ψ≈ -80º,80º) and β (φ,ψ≈ -150º, 60º) regions of the φ,ψ-map.

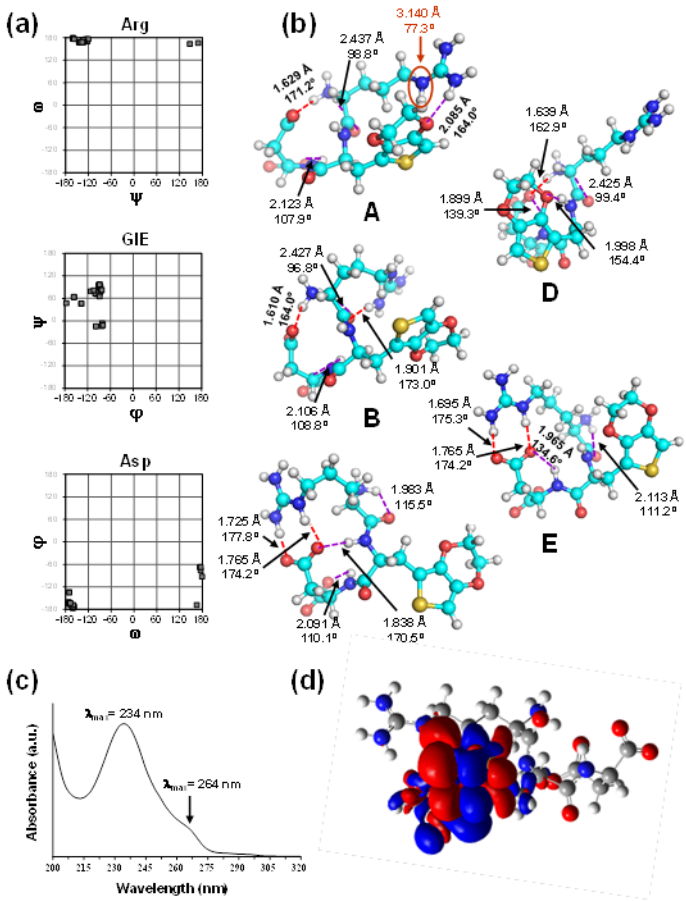


Figure 3.7. (a) Conformational maps constructed considering the backbone dihedral angles of the maps Arg (ω,ψ), GlE (φ,ψ) and Asp (ω,φ) residues, using the 25 unique minimum energy conformations identified in the clustering analysis for $RG^{E}D$. (b) Representation of the five low-energy structures (labeled from A to E in increasing order of energy) found in the clustering analysis for $RG^{E}D$/B in PBS. Salt bridges and hydrogen bonds are indicated by red and violet dashed lines, respectively. The red circle in A indicates that the N–H group involved in the N–H···π interaction with the thiophene group. Distances and angles associated to each of these interactions are given in Å and degrees, respectively. (c) UV-vis spectrum recorded for acetonitrile dilute $RG^{E}D$ solution (37 μM). (d) Electronic density difference (excited – ground) corresponding to the first electronic

transition ($\lambda^{(1)}_{max}$). Blue (red) regions indicate an increase (decrease) of electron density upon electronic transition.

Figure 3.7b depicts the five unique minimum energy conformations with relative energy (ΔE) lower than 5 kcal/mol, which have labeled with letters A-E in increasing order of energy. The lowest energy minimum (A) is stabilized by a complex network of hydrogen bonds that include: (i) two backbone···backbone (b-b) interactions involving the ionized N- and C-terminal group and the C=O and N-H moieties of the adjacent amide groups (*i.e.* known as C_5 interactions because atoms involved in such known define a five-membered ring); (ii) one side chain····backbone (sc-b) salt bridge between the carboxylate side group of Asp and the ionized N-terminus; and (iii) one side chain···side chain (sc-sc) interaction between the guanidium group of Arg and the dioxane ring of GlE. In addition, this structure is stabilized by a N–H···π interaction between the guanidium group of Arg and the thiophene ring of GlE. Such interaction has not been identified in any other unique minimum with ΔE≤ 5 kcal/mol. The next minimum (B), which is practically isoenergetic to A (ΔE < 0.1 kcal/mol), exhibits the same b-b and sc-b interactions. Thus, the only difference between A and B refers to the sc-sc interaction that, in the latter, consists of a hydrogen bond between the guanidium group and the C=O of Arg.

The most characteristic interaction in the next minimum (C), with ΔE= 1.5 kcal/mol, corresponds to a double sc-sc salt bridge between the guanidium and carboxylate groups of Arg

and Asp, respectively. As a consequence of such interaction, charged side groups are less exposed to the solvent in C than in A and B. This conformer also presents b-b interactions between the charged ammonium-end group and the amide of Arg as well as between the carboxylate-end group and the amide of Asp. In minimum D (ΔE= 3.9 kcal/mol) the Arg side chain adopts a fully extended arrangement that precludes the participation of the guanidium group in the formation of sc-sc and sc-b intramolecular interactions. Thus, interactions observed in D are similar to those described for A-C. Finally, minimum E, with ΔE= 4.7 kcal/mol, exhibits a double sc-sc salt bridge between Arg and Asp as well as two sc-b interactions involving the charged end groups. These are similar to the interactions identified for C.

Results displayed in Figure 3.7b indicate that, in general, the overall shape of the unique minimum energy conformations with $\Delta E < 5$ kcal/mol is appropriate to form intermolecular interactions with the integrin receptor. Thus, the general shape adopted by these conformations can be simply described as a small and deformed hairpin with the charged groups pointing outwards. As PEDOT chains are connected to $RG^{E}D$ through the side group of the GlE residue, all these conformations are not expected to be affected by the formation of the conjugate. Accordingly, the functionality of the RGD adhesive sequence is expected to the preserved in the peptide of PEDOT-$RG^{E}D$ conjugates.

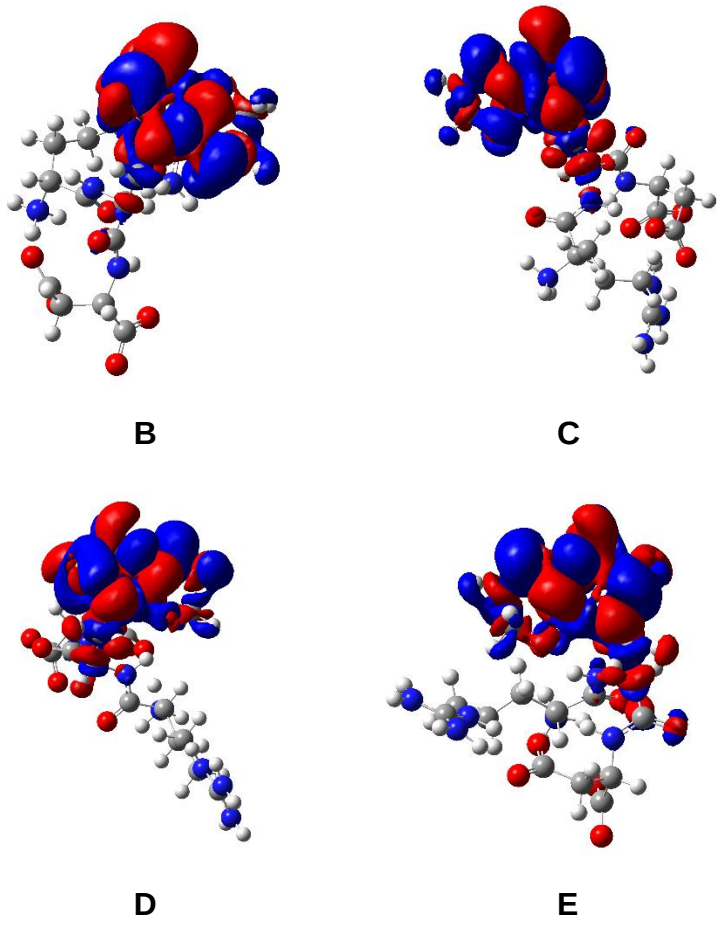


Figure 3.8. Electronic density difference (excited - ground) corresponding to the first and second electronic transition ($\lambda^{(1)}_{max}$) of conformation B-E. Blue (red) regions indicate an increase (decrease) of electron density upon electronic transition.

While EDOT absorption spectrum exhibits a broad peak characterized by a double hump (264 nm and 255 nm), the spectrum of $RG^{E}D$ in similar conditions (diluted in acetonitrile solution) exhibits a clear transition maximum appears at 234 nm with a small shoulder at 264 nm (see Figure 3.7c). In a previous work[17] on EDOT, we showed the importance of accounting for several conformations of such flexible structure

to reproduce the experimental data with computational means. In the same vein, optical properties were here simulated in acetonitrile for A-E conformations starting from the five above-mentioned structures. All conformations present similar spectra with two close maximum absorption wavelengths ($\lambda^{(1)}_{max}$ and $\lambda^{(2)}_{max}$) at 234 nm and 232 nm (associated oscillator strengths > 0.1). No significant variation was observed when selecting one conformation or another. Looking at the electronic density difference of these transitions of A conformation, it is crystal clear that the electronic transition occurs on the G^E moiety and is therefore not affected by Arg and Asp flexibility (Figure 3.7d, see Figure 3.8 for similar plots for B-D). Comparing the computational results with experiment, the computed peak at 234 nm reproduces the experimental main peak while the observed experimental shoulder at 264 nm could not be resolved using such methodology. The shoulder is therefore not the result of conformational effects but is related to the fine structure of G^E.

3.1.8 REFERENCES

1. Lutz, J. F. & Börner, H. G. Modern trends in polymer bioconjugates design. *Prog. Polym. Sci.* **33,** 1–39 (2008).
2. Eckhardt, D., Groenewolt, M., Krause, E. & Börner, H. G. Rational design of oligopeptide organizers for the formation of poly(ethylene oxide) nanofibers. *Chem. Commun.* **0,** 2814–2816 (2005).
3. Hentschel, J., Krause, E. & Börner, H. G. Switch-peptides to trigger the peptide guided assembly of poly(ethylene oxide)-peptide conjugates into tape structures. *J. Am. Chem. Soc.* **128,** 7722–7723 (2006).

4. Hamley, I. W., Krysmann, M. J., Castelletto, V. & Noirez, L. Multiple lyotropic polymorphism of a poly(ethylene glycol)-peptide conjugate in aqueous solution. *Adv. Mater.* **20,** 4394–4397 (2008).
5. Morell, M. & Puiggalí, J. Hybrid block copolymers constituted by peptides and synthetic polymers: An overview of synthetic approaches, supramolecular behavior and potential applications. *Polymers (Basel).* **5,** 188–224 (2013).
6. Murase, S. K. *et al.* Molecular characterization of l-phenylalanine terminated poly(l-lactide) conjugates. *RSC Adv.* **4,** 23231–23241 (2014).
7. Klok, H. A., Rösler, A., Götz, G., Mena-Osteritz, E. & Bäuerle, P. Synthesis of a silk-inspired peptide-oligothiophene conjugate. *Org. Biomol. Chem.* **2,** 3541–3544 (2004).
8. Guo, Y. *et al.* Electroactive oligoaniline-containing self-assembled monolayers for tissue engineering applications. *Biomacromolecules* **8,** 3025–3034 (2007).
9. Sanghvi, A. B., Miller, K. P. H., Belcher, A. M. & Schmidt, C. E. Biomaterials functionalization using a novel peptide that selectively binds to a conducting polymer. *Nat. Mater.* **4,** 496–502 (2005).
10. Tovar, J. D. Supramolecular construction of optoelectronic biomaterials. *Acc. Chem. Res.* **46,** 1527–1537 (2013).
11. Oyman, G. *et al.* Peptide-modified conducting polymer as a biofunctional surface: Monitoring of cell adhesion and proliferation. *RSC Adv.* **4,** 53411–53418 (2014).
12. Maione, S. *et al.* Electroactive polymer-peptide conjugates for adhesive biointerfaces. *Biomater. Sci.* **3,** 1395–1405 (2015).
13. Akbulut, H., Endo, T., Yamada, S. & Yagci, Y. Synthesis and characterization of polyphenylenes with polypeptide and poly(ethylene glycol) side chains. *J. Polym. Sci. Part A Polym. Chem.* **53,** 1786–1793 (2015).
14. Sanders, A. M., Magnanelli, T. J., Bragg, A. E. & Tovar, J. D. Photoinduced Electron Transfer within Supramolecular Donor-Acceptor Peptide Nanostructures under Aqueous Conditions. *J. Am. Chem. Soc.* **138,** 3362–3370 (2016).
15. Besar, K., Ardoña, H. A. M., Tovar, J. D. & Katz, H. E. Demonstration of Hole Transport and Voltage Equilibration in Self-Assembled π-Conjugated Peptide Nanostructures

Using Field-Effect Transistor Architectures. *ACS Nano* **9,** 12401–12409 (2015).
16. Fabregat, G. *et al.* An electroactive and biologically responsive hybrid conjugate based on chemical similarity. *Polym. Chem.* **4,** 1412–1424 (2013).
17. Fabregat, G. *et al.* Design of hybrid conjugates based on chemical similarity. *RSC Adv.* **3,** 21069–21083 (2013).
18. Gibson, K. D. & Scheraga, H. A. Revised algorithms for the build-up procedure for predicting protein conformations by energy minimization. *J. Comput. Chem.* **8,** 826–834 (1987).
19. Bogdanowich-Knipp, S. J., Jois, D. S. S. & Siahaan, T. J. The effect of conformation on the solution stability of linear vs. cyclic RGD peptides. *J. Pept. Res.* **53,** 523–529 (1999).
20. Kieffer, B., Mer, G., Mann, A. & Lefevre, J. F. Structural studies of two antiaggregant RGDW peptides by 1H and 13C NMR. *Int J Pept Protein Res* **44,** 70–79 (1994).
21. Stote, R. H., Dejaegere, a P., Lefevre, J. F. & Karplus, M. Multiple conformations of RGDW and DRGDW: A theoretical study and comparison with NMR results. *J. Phys. Chem. B* **104,** 1624–1636 (2000).
22. Park, H. S., Kim, C. & Kang, Y. K. Preferred conformations of RDGX tetrapeptides to inhibit the binding of fibrinogen to platelets. *Biopolymers* **63,** 298–313 (2002).
23. Kulkarni, A. K. & Ojha, R. P. Combined ^{1} H-NMR and Molecular Dynamics Studies on Conformational Behavior of a Model Heptapeptide, GRGDSPC. *Chem. Biol. Drug Des.* **84,** 567–577 (2014).
24. Stote, R. H. Analysis of the RGD sequence in protein structures: Comparison to the conformations of the RGDW andDRGDW peptides determined by molecular dynamics simulations. *Theor. Chem. Acc.* **106,** 128–136 (2001).
25. Upadhyay, S. K. Molecular dynamics simulations of certain RGD-based peptides from kistrin provide insight into the higher activity of REI-RGD34 protein at higher temperature. *Chem. Biol. Drug Des.* **83,** 600–609 (2014).
26. Asakura, T. *et al.* NMR analysis of the fibronectin cell-adhesive sequence, Arg-Gly-Asp, in a recombinant silk-like protein and a model peptide. *Biomacromolecules* **12,** 3910–3916 (2011).
27. Becke, A. D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **98,** 1372–1377

(1993).
28. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37,** 785–789 (1988).
29. Alemán, C., Jiménez, A. I., Cativiela, C., Perez, J. J. & Casanovas, J. Influence of the phenyl side chain on the conformation of cyclopropane analogues of phenylalanin. *J. Phys. Chem. B* **106,** 11849–11858 (2002).
30. Tomasi, J. & Persico, M. Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. *Chem. Rev.* **94,** 2027–2094 (1994).
31. Flory, P. J. & Volkenstein, M. Statistical mechanics of chain molecules. *Biopolymers* **8,** 699–700 (1969).
32. Koo, J. C. P. *et al.* Conformational dependence of the intrinsic acidity of the aspartic acid residue sidechain in N-acetyl-L-aspartic acid-N′-methylamide. *J. Mol. Struct. THEOCHEM* **620,** 231–255 (2003).
33. Schlund, S., Müller, R., Gramann, C. & Engels, B. Conformational analysis of arginine in gas phase-A strategy for scanning the potential energy surface effectively. *J. Comput. Chem.* **29,** 407–415 (2008).
34. Frisch, M. J. *et al.* Gaussian 09, revision A. 1. *Gaussian Inc. Wallingford CT* **27,** 34 (2009).
35. García De La Torre, J., Huertas, M. L. & Carrasco, B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* **78,** 719–730 (2000).
36. Ortega, A., Amorós, D. & García De La Torre, J. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101,** 892–898 (2011).
37. Zanuy, D., Casanovas, J. & Alemán, C. Thermal Stability of the Secondary Structure of Poly(α,L-glutamate) in Self-Assembled Complexes as Studied by Molecular Dynamics in Chloroform Solution. *J. Am. Chem. Soc.* **126,** 704–705 (2004).
38. Zanuy, D., Alemán, C. & Muñoz-Guerra, S. A molecular dynamics study of the stoichiometric complex formed by poly (α, L-glutamate) and octyltrimethylammonium ions in chloroform solution. *Biopolymers* **63,** 151–162 (2002).
39. Teixeira-Dias, B. *et al.* Binding of 6-mer single-stranded

homo-nucleotides to poly(3,4- ethylenedioxythiophene): Specific hydrogen bonds with guanine. *Soft Matter* **7,** 9922–9932 (2011).

40. Alemán, C. On the ability of modified peptide links to form hydrogen bonds. *J. Phys. Chem. A* **105,** 6717–6723 (2001).
41. Perez, J. J., Villar, H. O. & Loew, G. H. Characterization of low-energy conformational domains for Met-enkephalin. *J. Comput. Aided. Mol. Des.* **6,** 175–190 (1992).
42. Pérez, J. J., Villar, H. O. & Arteca, G. A. Distribution of conformational energy minima in molecules with multiple torsional degrees of freedom. *J. Phys. Chem.* **98,** 2318–2324 (1994).
43. Alemán, C. & Curcó, D. Simulation of amorphous nylon 6 using SuSi, a strategy based on generation-relaxation algorithms. *Macromol. Theory Simulations* **13,** 345–354 (2004).
44. Burchard, W., Schmidt, M. & Stockmayer, W. H. Information on Polydispersity and Branching from Combined Quasi-Elastic and Integrated Scattering. *Macromolecules* **13,** 1265–1272 (1980).
45. Ruoslahti, E. & Pierschbacher, M. D. New perspectives in cell adhesion: RGD and integrins. *Science (80-.).* **238,** 491–497 (1987).
46. Chatterjee, J., Gilon, C., Hoffman, A. & Kessler, H. N-methylation of peptides: A new perspective in medicinal chemistry. *Acc. Chem. Res.* **41,** 1331–1342 (2008).
47. Wakefield, A. E., Wuest, W. M. & Voelz, V. A. Molecular simulation of conformational pre-organization in cyclic RGD peptides. *J. Chem. Inf. Model.* **55,** 806–813 (2015).
48. Hamley, I. W. Self-assembly of amphiphilic peptides. *Soft Matter* **7,** 4122–4138 (2011).
49. Hersel, U., Dahmen, C. & Kessler, H. RGD modified polymers: Biomaterials for stimulated cell adhesion and beyond. *Biomaterials* **24,** 4385–4415 (2003).
50. Cheng, G., Castelletto, V., Jones, R. R., Connon, C. J. & Hamley, I. W. Hydrogelation of self-assembling RGD-based peptides. *Soft Matter* **7,** 1326–1333 (2011).
51. Castelletto, V. *et al.* Self-assembly of Fmoc-tetrapeptides based on the RGDS cell adhesion motif. *Soft Matter* **7,** 11405–11415 (2011).
52. López-Pérez, D. E., Revilla-López, G., Hamley, I. W. &

Alemán, C. Molecular insights into aggregates made of amphiphilic Fmoc-tetrapeptides. *Soft Matter* **9,** 11021–11032 (2013).
53. Zanuy, D., Poater, J., Solà, M., Hamley, I. W. & Alemán, C. Fmoc-RGDS based fibrils: Atomistic details of their hierarchical assembly. *Phys. Chem. Chem. Phys.* **18,** 1265–1278 (2016).
54. Barril, X., Alemán, C., Orozco, M. & Luque, F. J. Salt bridge interactions: Stability of the ionic and neutral complexes in the gas phase, in solution, and in proteins. *Proteins Struct. Funct. Genet.* **32,** 67–79 (1998).
55. Geiger, T. & Clarke, S. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. *J. Biol. Chem.* **262,** 785–794 (1987).
56. Pierschbacher, M. D. & Ruoslahti, E. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature* **309,** 30–33 (1984).
57. Hynes, R. O. Integrins: Versatility, modulation, and signaling in cell adhesion. *Cell* **69,** 11–25 (1992).
58. Li, A. *et al.* A bioengineered peripheral nerve construct using aligned peptide amphiphile nanofibers. *Biomaterials* **35,** 8780–8790 (2014).
59. Fukunaga, K., Tsutsumi, H. & Mihara, H. Self-assembling peptide nanofibers promoting cell adhesion and differentiation. *Biopolymers* **100,** 731–737 (2013).
60. Galler, K. M., Aulisa, L., Regan, K. R., D'Souza, R. N. & Hartgerink, J. D. Self-assembling multidomain peptide hydrogels: Designed susceptibility to enzymatic cleavage allows enhanced cell migration and spreading. *J. Am. Chem. Soc.* **132,** 3217–3223 (2010).
61. Gil, E. S. *et al.* Helicoidal multi-lamellar features of RGD-functionalized silk biomaterials for corneal tissue engineering. *Biomaterials* **31,** 8953–8963 (2010).
62. Wu, J. *et al.* Corneal stromal bioequivalents secreted on patterned silk substrates. *Biomaterials* **35,** 3744–3755 (2014).
63. Kim, B. J., Choi, Y. S. & Cha, H. J. Reinforced multifunctionalized nanofibrous scaffolds using mussel adhesive proteins. *Angew. Chemie - Int. Ed.* **51,** 675–678 (2012).
64. Dennes, T. J., Hunt, G. C., Schwarzbauer, J. E. & Schwartz,

J. High-yield activation of scaffold polymer surfaces to attach cell adhesion molecules. *J. Am. Chem. Soc.* **129,** 93–97 (2007).

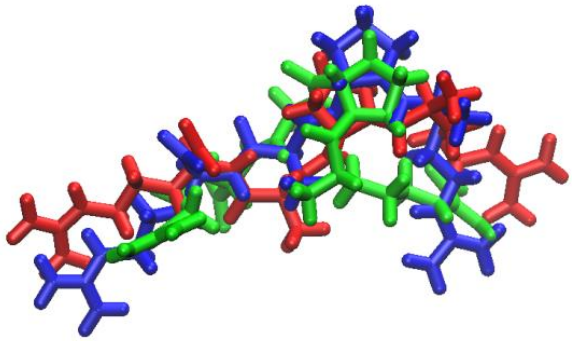
65. Castelletto, V., Gouveia, R. J., Connon, C. J. & Hamley, I. W. Self-assembly and bioactivity of a polymer/peptide conjugate containing the RGD cell adhesion motif and PEG. *Eur. Polym. J.* **49,** 2961–2967 (2013).
66. Viswanathan, P. *et al.* Controlling surface topology and functionality of electrospun fibers on the nanoscale using amphiphilic block copolymers to direct mesenchymal progenitor cell adhesion. *Biomacromolecules* **16,** 66–75 (2015).
67. Yang, Z. *et al.* Chitosan microsphere scaffold tethered with RGD-conjugated poly(methacrylic acid) brushes as effective carriers for the endothelial cells. *Macromol. Biosci.* **14,** 1299–1311 (2014).
68. Whitesides, G. M. Nanoscience, nanotechnology, and chemistry. *Small* **1,** 172–179 (2005).
69. Zhang, S. Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol.* **21,** 1171–1178 (2003).
70. Liu, T. Y. *et al.* Self-adjuvanting polymer-peptide conjugates as therapeutic vaccine candidates against cervical cancer. *Biomacromolecules* **14,** 2798–2806 (2013).
71. De Giglio, E., Sabbatini, L. & Zambonin, P. G. Development and analytical characterization of cysteine-grafted polypyrrole films electrosynthesized on ptand Ti-substrates as precursors of bioactive interfaces. *J. Biomater. Sci. Polym. Ed.* **10,** 845–858 (1999).
72. De Giglio, E., Sabbatini, L., Colucci, S. & Zambonin, G. Synthesis, analytical characterization, and osteoblast adhesion properties on RGD-grafted polypyrrole coatings on titanium substrates. *J. Biomater. Sci. Polym. Ed.* **11,** 1073–1083 (2000).
73. Fabregat, G. *et al.* Design of hybrid conjugates based on chemical similarity. *RSC Adv.* **3,** 21069–21083 (2013).
74. Groenendaal, L., Zotti, G., Aubert, P. H., Waybright, S. M. & Reynolds, J. R. Electrochemistry of poly(3,4-alkylenedioxythiophene) derivatives. *Adv. Mater.* **15,** 855–879 (2003).
75. Groenendaal, L., Jonas, F., Freitag, D., Pielartzik, H. &

Reynolds, J. R. Poly(3,4-ethylenedioxythiophene) and its derivatives: past, present, and future. *Adv. Mater.* **12,** 481–494 (2000).

76. Kirchmeyer, S. & Reuter, K. Scientific importance, properties and growing applications of poly(3,4-ethylenedioxythiophene). *J. Mater. Chem.* **15,** 2077–2088 (2005).
77. Maione, S. Polymers and peptide-polymer conjugates as bioactive platforms and carriers for the encapsulation of biomolecules. *TDX (Tesis Dr. en Xarxa)* (2017).
78. Zanuy, D. *et al.* In silico molecular engineering for a targeted replacement in a tumor-homing peptide. *J. Phys. Chem. B* **113,** 7879–7889 (2009).
79. Kang, Y. K. & Byun, B. J. Assessment of density functionals with long-range and/or empirical dispersion corrections for conformational energy calculations of peptides. *J. Comput. Chem.* **31,** 2915–2923 (2010).

- Chapter 4 -

Modeling of a C-End Rule Peptide Adsorbed onto Gold Nanoparticles



4 MODELING OF A C-END RULE PEPTIDE ADSORBED ONTO GOLD NANOPARTICLES[§]

4.1 ABSTRACT

The RPAR peptide, a prototype C-end Rule (CendR) sequence that binds to neuropilin-1 (NRP-1), has potential therapeutic uses as internalization trigger in anticancer nanodevices. Recently, the functionalization of gold nanoparticles with CendR peptides has been proved to be a successful strategy to target the NRP-1 receptor in prostate cancer cells. In this work, we investigate the influence of two gold surface facets, (100) and (111), on the conformational preferences of RPAR using molecular dynamics simulations. Both clustering and conformational analyses revealed that the peptide backbone becomes very rigid upon adsorption onto gold, which is a very fast and favored process, the only flexibility being attributed to the side chains of the two Arg residues. Thus, the different components of RPAR tend to adopt an elongated shape, which is characterized by the pseudo-extended conformation of both the backbone and the Arg side chains. This conformation is very different from the already known bioactive conformation, indicating that RPAR is drastically affected by the substrate. Interestingly, the preferred conformations of the peptide adsorbed onto gold facets are not stabilized by salt bridges and/or specific intramolecular hydrogen bonds, which represent an important difference with respect to the conformations found in other environments (*e.g.* the peptide in solution and interacting with NRP-1 receptor). However,

[1] This work has been published as Triguero, J., Flores-Ortega, A., Zanuy, D. & Alemán, C. Modeling of a C-end rule peptide adsorbed onto gold nanoparticles. *J. Pept. Sci.* **24,** e3057 (2018).

the conformational changes induced by the substrate are not in detriment of the use of gold nanoparticles as appropriated vehicles for the transport and targeted delivery of the RPAR. Thus, once their high affinity for the NRP-1 receptor induces the targeted delivery of the elongated peptide molecules from the gold nanoparticles, the lack of intramolecular interactions facilitates their evolution towards the bioactive conformation, increasing the therapeutic efficacy of the peptide.

4.2 INTRODUCTION

Delivering drugs to neoplastic tissues entails major challenges, as for example tumor homing specificity and the internalization of the drug inside the tumor cells.[1,2] Ruoslahti and coworkers[3] reported a class of peptides that were effective in entering into cultured cells. Peptides that induced cell- and tissue-penetration activity exhibited a strict requirement at the C-terminal consensus motif, R/KXXR/K (where X represents any amino acid), which was termed the C-end rule (CendR).[3] Among CendR peptides, the prototype sequence RPAR binds to neuropilin-1 (NRP-1) on the target cells.[3] NRP-1 is a cell surface receptor with multiple ligands that have a major role in angiogenesis and cardiovascular development,[4,5] and acts as a co-receptor of a tumor-secreted vascular permeability factor, named VEGF, during induction of vascular permeability. CendR peptides are particularly relevant to cancer drug delivery, as NRP-1 is frequently overexpressed in tumor cells. Thus, Sugahara *et al.*[6] suggested that cancer therapeutic approaches based on peptides with an RPAR terminus can be useful for the penetration of drug into

tumors since the binding of such peptides to NRP-1 promotes the penetration of co-administered peptides.

Unfortunately, the crystal structure of NRP-1 complexed with a RPAR, or any other R/KXXR/K peptide, has not been reported yet. In contrast, the structure of the binding domain of NRP-1 with a tetrapeptide named tuftsin (TKPR), which mimic the VEGF terminus, was elucidated.[7] We used the structure of NRP-1/tufsin complex as a template to model the NRP-1/RPAR complex using atomistic molecular dynamics (MD) simulations.[8] Results showed that RPAR binds NRP-1 through specific interactions of the RPAR C-terminus: three hydrogen bonds and a salt bridge anchor the ligand in the receptor pocket. Also, modeling results were employed as the starting point for a systematic computational study of new RPAR analogues based on chemical modifications.[8,9] The potential therapeutic use of the RPAR peptide as an internalization trigger in anticancer nanodevices requires major understanding of how its conformation and its ability to form networks of interactions change with the environment. For example, previous MD-based studies on CREKA, a linear pentapeptide that selectively binds to tumor vasculature,[10] allowed to determine the bioactive conformation,[11,12] to increase the resistance of the peptide versus proteases[13] and to ascertain the influence of inert gold surfaces when the peptide is adsorbed onto nanoparticles.[14] The aim of this work is to investigate the influence of gold surfaces on the energy landscape of RPAR adsorbed onto gold nanoparticles. Thus, in recent years gold nanoparticles have emerged as attractive candidates for delivery of therapeutic biomolecules into their targets since they are selectively

taken by human cells without detriment of cellular functions.[15–19] Accordingly, gold nanoparticles coated with biomolecules have many applications in biomedical imaging, clinical diagnosis and therapy. Indeed, in a recent study gold nanoparticles were successfully functionalized with a R/KXXR/K (RGDK) peptide to target the NRP-1 receptor in prostate cancer cells.[20] This approach evidenced an enhancement of both cell penetration and internalization efficiency.

In this work we use atomistic MD simulations to investigate the influence of the gold substrate in the conformational preferences of RPAR, which can be considered as a very representative cell penetrating sequence. The structures found for the RPAR peptide when it is physically adsorbed onto the (100) and (111) gold surface facets have been exhaustively characterized through clustering and conformational analyses, and the formed interactions have been evaluated. Comparison with the bioactive conformation, which corresponds to that adopted when bound to the NRP-1 receptor, reveals that the metallic substrate induces significant conformational changes. In spite of those expected changes, our MD results prove that gold nanoparticles can be considered as suitable vehicles for the targeted delivery of cell penetrating peptides.

4.3 METHODS

MD simulations were carried out to obtain microscopic information about how the conformational preferences of RPAR are affected by its adsorption onto two different gold surfaces: (100)

and (111). RPAR was built and represented using the standard parameters of the AMBER force-field.[21] The peptide titratable groups were set to their ionized forms at pH = 7.0.

The gold surfaces were represented using a *n*-layered slab model with dimensions 51.91 × 51.91 $Å^2$, representing the cubic (100) and the rhomboid (111) facets, respectively, of the FCC gold unit cell. Thus, each layer contained 324 atoms forming an 18 × 18 supercell, the modeled (100) and (111) facets displaying 3564 (*n*= 11 layers) and 3240 (*n*= 10 layers) gold atoms, respectively. The thickness (*i.e.* extension in the *c*-direction) of this sheet model was 21.96 and 20.39 Å for the (100) and (111) facets, respectively. The van der Waals parameters used to compute the interaction between the Au particles and the rest of species were compatible with the AMBER force-field.[22] The crystallographic parameter for gold was *a*= 2.93 Å. Although this parameter is 1.4% larger than the experimental one (*a*= 2.89 Å), it was chosen for consistency with the force-field parameters used to simulate the Au particles, which predicted *a*= 2.93 Å.[22] In order to avoid the bending of the gold sheet, the positions of the Au atoms at the surface was kept fixed during the MD simulations.

Water molecules, which were represented using the TIP3P model,[23] were added to the simulation box to fill up the space not occupied by the surface and the peptide. This increased the total *z*-axis length up to 89.0 and 71.7 Å for systems constructed with the (100) and (111) surface, respectively. These model systems involved a total of 5661 and 3473 water molecules, respectively. It is worth noting that the (111) facet is lower in energy than that of

(100) one.[24] All simulations were performed using the NAMD 2.9 program.[25]

The atom pair cut-off distance was set at 14.0 Å to compute the van der Waals interactions. In order to avoid discontinuities in the potential energy function, non-bonding energy terms were forced to slowly converge to zero, by applying a smoothing factor from a distance of 12.0 Å. Beyond the cut off distance, electrostatic interactions were calculated by using Particle Mesh of Ewald (PME).[26] The real space term was defined by the van der Waals cutoff (14.0 Å), while the reciprocal space was computed by interpolation of the effective charge into a charge mesh with a grid thickness of 1 points per volume unit. Due to its fast convergence, the weak coupling method[27] was used to heat the system and to rapidly equilibrate its pressure and temperature around 1 bar and 298 K, respectively. The relaxation times used for the coupling were 2 and 10 ps for temperature and pressure, respectively. For final relaxation and for all production runs, both temperature and pressure were controlled by the Nose–Hoover piston.[28] Pressure was kept at 1.01325 bars, the oscillation period was set at 400 fs while the piston decay time was set at 100 fs. The piston temperature was set at the same value as the thermostat control, 298 K, which used a damping coefficient of 2 ps. The numerical integration step was set at 2 fs. Bond lengths involving hydrogen atoms were constrained using the Rattle algorithm.[29]

Before the relaxation and production steps described in the Results section, models were thermalized and equilibrated using a four-step process. Such processes, which applied keeping the

peptide frozen, can be summarized as follows: 1) 1×10^4 steps of energy minimization; 2) 0.2 ns of NVT MD at 800 K; 3) 0.3 ns of NVT MD at 500 K; and 4) 0.5 ns of $NP_{zz}T$ (*i.e.* only the box length in the *z*-direction was allowing to move such that the P_{zz} component of the pressure tensor was equal to the imposed pressure) at 298 K and 1.01325 bars.

The next section provides details about the analysis of the structural evolution of the peptide during the MD production runs, for which the equilibrated models were used as starting points of 100 ns simulations. Snapshots were stored every 20 ps (*i.e.* 5000 structures were saved for analyses per production run). In order to ensure our results reproducibility, additional production MD simulations were performed considering as starting points models obtained using the same procedures described above but introducing small changes at the equilibration step. The length of these additional production trajectories was 50 ns. As results were very similar to those described in the Results and Discussion section, these production trajectories were not enlarged.

4.4 RESULTS AND DISCUSSION

4.4.1 Adsorption of RPAR onto gold

To gain information into the adsorption of RPAR onto gold, two independent simulations (each one by duplicate) were performed considering the (100) and (111) facets. In these simulations, hereafter denoted MD_{100} and MD_{111}, respectively, the peptide was initially set at a distance of ~7 Å from the surface. According to this starting geometry, the solvated peptide was at a distance close to the

surface but without adsorption. The temporal evolution of the distances between the center of masses of each peptide residue and the gold surface, which were calculated using the normal vectors, showed that the peptide was rapidly adsorbed by the surface (*i.e.* after 70 and 40 ps for MD_{100} and MD_{111}, respectively).

	E_{p-p}	E_{p-Au}	E_{p-wat}	E_{Au-wat}
MD_{100}	-62±8	-90±5	-532±24	-31165±143
MD_{111}	-88±10	-96±4	-469±24	-30826±109

Table 4.1. Average energy contributions (in kcal/mol) extracted from MD_{100} and MD_{111} production trajectories (see text).

In order to understand the peptide adsorption process, the interactions involved in four main energy contributions have been analyzed: (*i*) the intramolecular peptide bonding (*i.e.* stretching, bending and torsional strain) and non-bonding (*i.e.* electrostatic and van der Waals beyond three consecutively bonded atoms) interactions, E_{p-p}, which take into account the changes induced in the geometry of the peptide; (*ii*) the interactions of the peptide with the gold surfaces, E_{p-Au}; (*iii*) the interactions of the peptide with the solvent molecules, E_{p-wat}; and (*iv*) the interactions between the gold surface and solvent molecules, E_{Au-wat}. The average energy contributions obtained for MD_{100} and MD_{111} are listed in Table 4.1. The sum of E_{p-p}, E_{p-Au} and E_{p-wat} favor the adsorption of RPAR onto the two gold surfaces. However, the lowest sum corresponds to the (100), which is favored with respect to the (111) by ~30 kcal/mol. Although intramolecular peptide interactions (ΔE_{p-p}= -26 kcal/mol)

and peptide-gold interactions ($\Delta E_{p\text{-}Au}$= -6 kcal/mol) favor the adsorption onto the (111) surface, peptide-water interactions ($\Delta E_{p\text{-}wat}$= +63 kcal/mol) are more favorable when RPAR is adsorbed onto the (100) surface. On the other hand, inspection of the temporal evolution of the different contributions reflects the stability of the adsorption peptide (Figure 4.1).

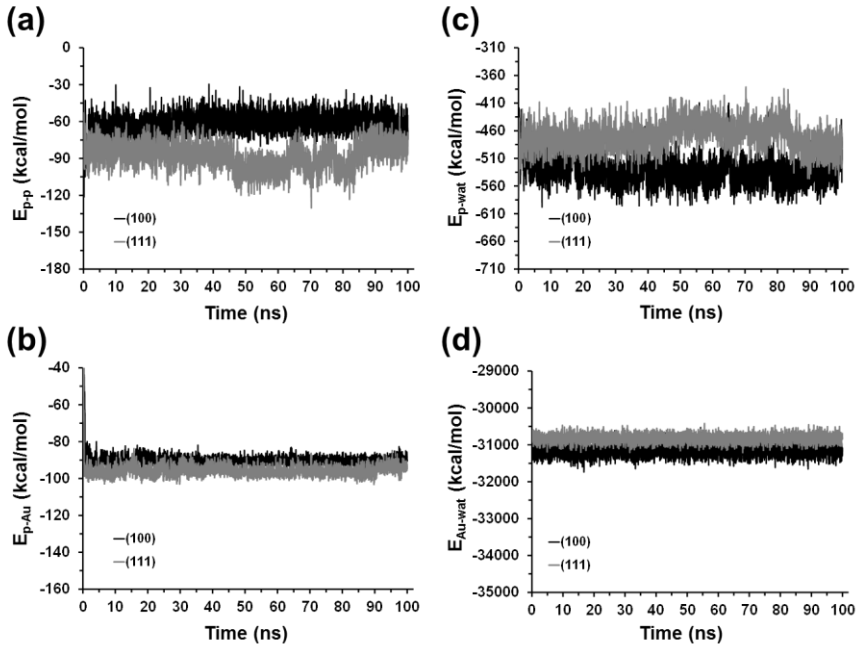


Figure 4.1. Temporal evolution of the different energy contributions extracted from MD_{100} and MD_{111} trajectories. Energy contributions arise from (a) intramolecular peptide interactions, $E_{p\text{-}p}$; (b) interactions between the peptide and the gold surface, $E_{p\text{-}Au}$; (c) interactions between the peptide and the solvent molecules, $E_{p\text{-}wat}$; and (d) interactions between the gold surface and solvent molecules, $E_{Au\text{-}wat}$.

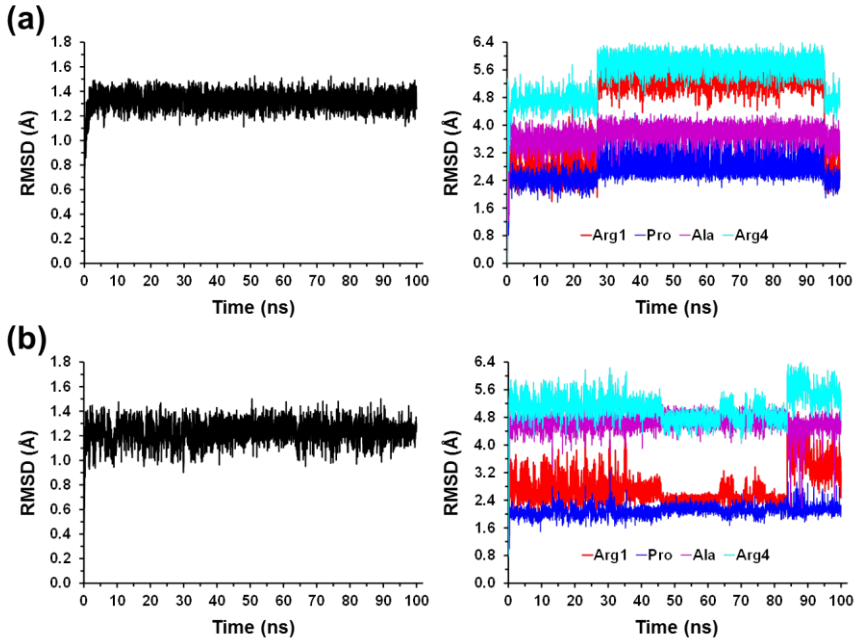


Figure 4.2. For (a) MD_{100} and (b) MD_{111}: Backbone RMSD of the whole peptide calculated considering N, C^{α} and C(=O) atoms of the four residues (left) and RMSD for each individual residue calculated considering all non-hydrogen atoms (right).

The temporal evolution of the root mean square deviation (RMSD) of RPAR, which was calculated with respect to the initial conformation and considering the N, C^{α} and C(=O) backbone atoms of each residue, is displayed in Figure 4.2a-b for MD_{100} and MD_{111}, respectively. As it can been, after adsorption of the peptide onto the gold surface, the RMSD exhibits very small fluctuations. Thus, the average RMSD is 1.32±0.07 and 1.23±0.09 Å for MD_{100} and MD_{111}, respectively, indicating that the backbone remains structurally stable during the whole simulations. However, the RMSD calculated for the four individual residues considering all

non-hydrogen atoms of both the backbone and the side chains, which are included Figure 4.2, reflects the conformational flexibility of the two Arg side groups. This is particularly remarkable for the Arg1 residue in the MD_{100} simulation, where the value averaged during the first 27 ns, 2.75±0.39 Å, increases to 5.15±0.68 Å for the interval comprised between 30 and 100 ns. In spite of this, Figure 4.2 suggests that the conformational flexibility of RPAR is quite restricted when adsorbed onto gold surfaces, fluctuations in the calculated RMSDs being essentially attributed to local re-arrangements involving a reduced number of atoms.

The root mean square fluctuations (RMSF) of all non-hydrogen RPAR atoms with respect to the averaged conformation are displayed in Figure 4.3 for both MD_{100} and MD_{111}. The higher values correspond to the side chain atoms of the Arg residues, especially to Arg1. These results corroborate the remarkable conformational variability of the side chain of Arg1, and to a lesser extent that of Arg4, while the backbone atoms remain stable.

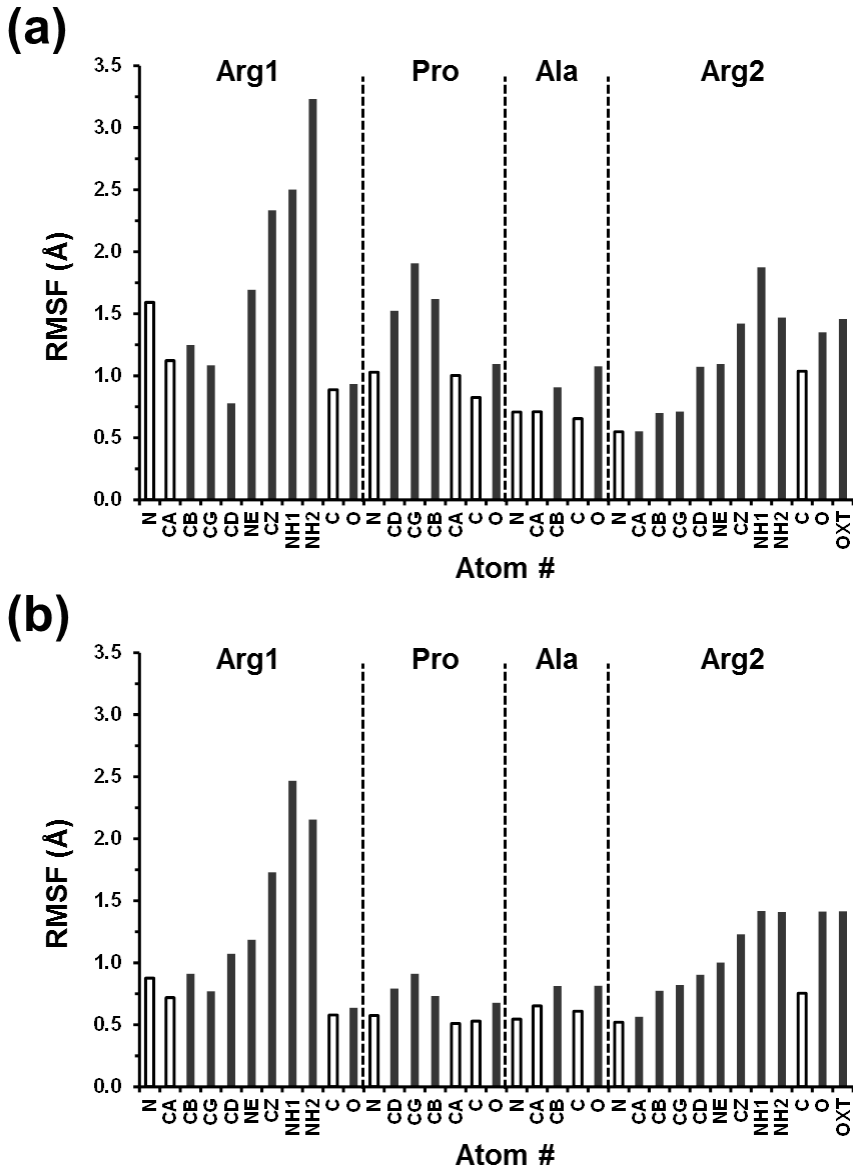


Figure 4.3. For (a) MD_{100} and (b) MD_{111}: RMSF of non-hydrogen atoms with respect to the averaged conformation. Backbone atoms (N, C^{α}, named CA, and C) are displayed by empty bars while the rest of atoms are represented by filled bars. Atoms have labelled using the nomenclature for atom types of the AMBER force-field.[21]

4.4.2 Clustering analyses

In order to rigorously evaluate the conformational preferences and flexibility of RPAR when adsorbed onto (100) and (111) facets of gold, the similarity between the 5000 peptide structures saved during each MD trajectory has been analyzed through clustering analyses. Initially, the structures stored during MD_{100} and MD_{111} simulations were classified in list of "*different structures*". Two structures were considered different when they differ in at least one of their backbone or side chain dihedral angles by more than a value defined as Δ (in degrees). Figures 4.4a and 4.4b represents the number of different structures against Δ, which has varied between 1º and 90º in steps of 1º, for MD_{100} and MD_{111}, respectively. For MD_{100}, a flat profile is observed from Δ= 1º to Δ= 11º. At Δ > 11º, the slope experiences a drastic change, indicating that the effectivity of the clustering analysis starts. Obviously, the precision of the clustering analysis decreases with increasing Δ (*i.e.* growing conformational differences among clusters). When Δ= 25º, the slope of the profile experiences another pronounced change, evidencing that the clustering process can become unmeaning for higher Δ values. Inspection of Figure 4.4b indicates that for MD_{111} the effectivity of the clustering analysis starts at Δ= 14º for MD_{111}, while the loss of meaning initiates at Δ > 30º.

In order to evaluate the influence of the backbone rigidity in the clustering process, the classification process was repeated but considering differences in the backbone dihedral angles, exclusively. Results, which are depicted in Figures 4.4c and 4.4d for MD_{100} and MD_{111}, respectively, indicate that the interval of Δ

values in which the clustering process is effective becomes narrower than that achieved by including differences in the side chain dihedral angles. More specifically, the interval extends from Δ= 4º to 12º for MD_{100} and from Δ= 5 to 13º for MD_{111}. This feature clearly supports that the flexibility of the peptide backbone is very restricted, which is agreement with the RMSD and RMSF analyses displayed in the previous sub-section (Figures 4.2 and 4.3, respectively).

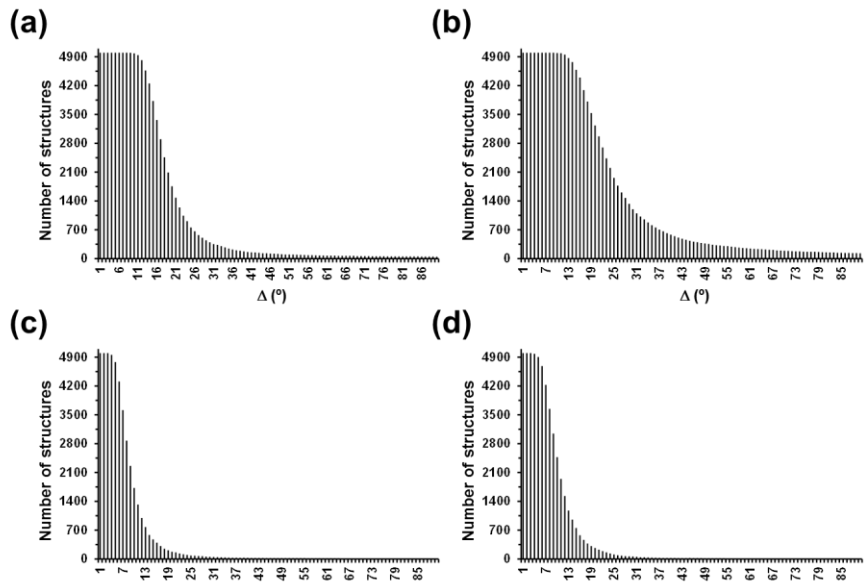


Figure 4.4. Variation of the number of different structures against Δ, which has been increased from 1º to 90º in steps of 1º, considering (a, b) both the backbone and side chain dihedral angles or (c, d) the backbone dihedral angles only for (a, c) MD_{100}; and (b, d) MD_{111}.

Selection of the appropriate Δ value requires the maximum effectivity without loss of significance. Accordingly, the clustering analyses of RPAR adsorbed onto gold surfaces were performed

considering all dihedral angles of the peptide and using $\Delta = 25^{o}$, which resulted in 810 and 1908 different structures, or clusters, for MD_{100} and MD_{111}, respectively. Thus, each cluster is constituted by a set of structures that are not different among them (*i.e.* differences between their backbone or side chain dihedral angles are lower than 25^{o}), the structure considered as head of series in the cluster being the one of lowest energy. It should be remarked that the selected Δ value represents a good compromise between efficacy (*i.e.* the number of structures to analyze decreases by 84% and 62% for MD_{100} and MD_{111}, respectively) and precision (*i.e.* $\Delta = 25^{o}$ is still comprised within the meaning interval of the profiles displayed in Figures 4.4a-b).

The list of clusters was organized considering their head of series structures and ranking them in an increasing order of the energy. Figures 4.5a and 4.5b represent the relative energy, which was calculated with respect to the head of series structure with lowest energy, of the 5000 structures stored during MD_{100} and MD_{111} simulations against their position within the list of clusters. The profiles derived from the two trajectories display a similar behavior with three well-defined regions. The first one, which comprises approximately 15-20 clusters, exhibits a very rapid increment of relative energy (*i.e.* the energy gap between the head of series structure of two consecutive clusters of the list, $\Delta\Delta E$, corresponds to several kcal/mol). After this, the second region, which comprises most of the clusters of the list, exhibits a slower growing of the relative energy (*i.e.* $\Delta\Delta E$ is lower than 1 kcal/mol). Finally, the last clusters of the list show, again, drastic energy variations between

consecutive clusters, resulting in a pronounced slope similar to that of the first region.

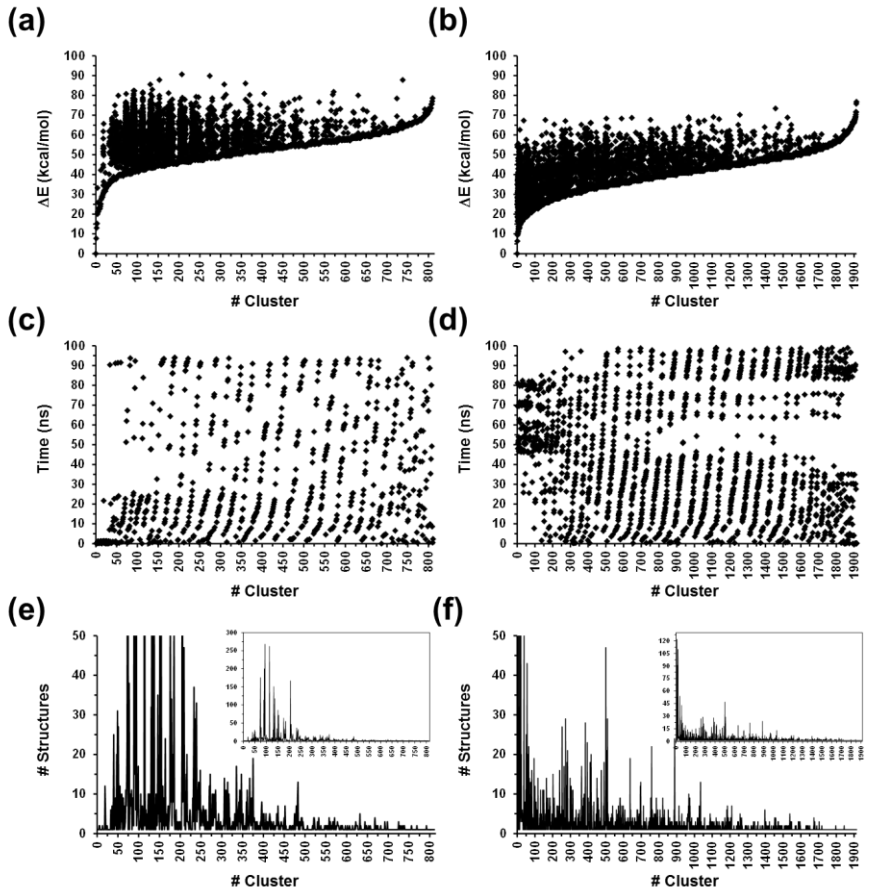


Figure 4.5. For (a, c, e) MD_{100} and (b, d, f) MD_{111}: (a, b) Relative energy (ΔE) of the all saved RPAR structures against the position of the corresponding cluster in the energy ranking list; (c, d) Time at which the head of series structure of the cluster was reached during the MD trajectory against the position of the cluster in the energy ranking list; (e, f) Number of structures contained in the cluster against its position in the energy ranking list.

Figures 4.5c and 4.5d show the point of the MD path in which the head of series structure of each cluster was reached. As it can be seen, there is no correlation between the stability of the cluster (*i.e.* position of the cluster in the ranked list) and the simulation time. This feature suggests that the dynamics of the peptide is not regulated by the formation of specific interactions with the gold surface. Finally, the number of structures per cluster is displayed in Figures 4.5e and 4.5f for MD_{100} and MD_{111}, respectively. The graphic obtained for the clustering of the MD_{100} trajectory can be roughly approximated to a unimodal Gaussian distribution, the most populated clusters being comprised between the positions 74 and 205 of the ranked list. In contrast the population of the rest of the clusters is very small, supporting the fact that differences among them are small (*i.e.* one or more side chain dihedral angles). This Gaussian-like resulting curve is not detected in MD_{111}, which in fact does not exhibit a characteristic shaped distribution. Accordingly, the most populated clusters of MD_{111} are located at the beginning of the ranking list and correspond to those with the more stable head of series.

4.4.3 Conformational analyses

The specific conformations of RPAR can be described considering two virtual dihedral angles, ξ_1 and ξ_2, and two specific dihedral angles of the peptide, ψ_1 and ψ_2. The virtual dihedral angles ξ_1 and ξ_2 have been defined by the C^{α} atoms of the four residues [C^{α}(Arg1)–C^{α}(Pro)–C^{α}(Ala)– C^{α}(Arg2)] and both the C^{η} and the C^{α} atoms of Arg1 and Arg2 [C^{η}(Arg1)–C^{α}(Arg1)–

C^{α}(Arg2)–C^{η}(Arg2)], respectively. The angles ψ_1 and ψ_2 correspond to the ψ angles of Arg1 [N(Arg1)–C^{α}(Arg1)–C(Arg1)–N(Pro)] and Ala [N(Ala)–C^{α}(Ala)–C(Ala)–N(Arg2)], respectively.

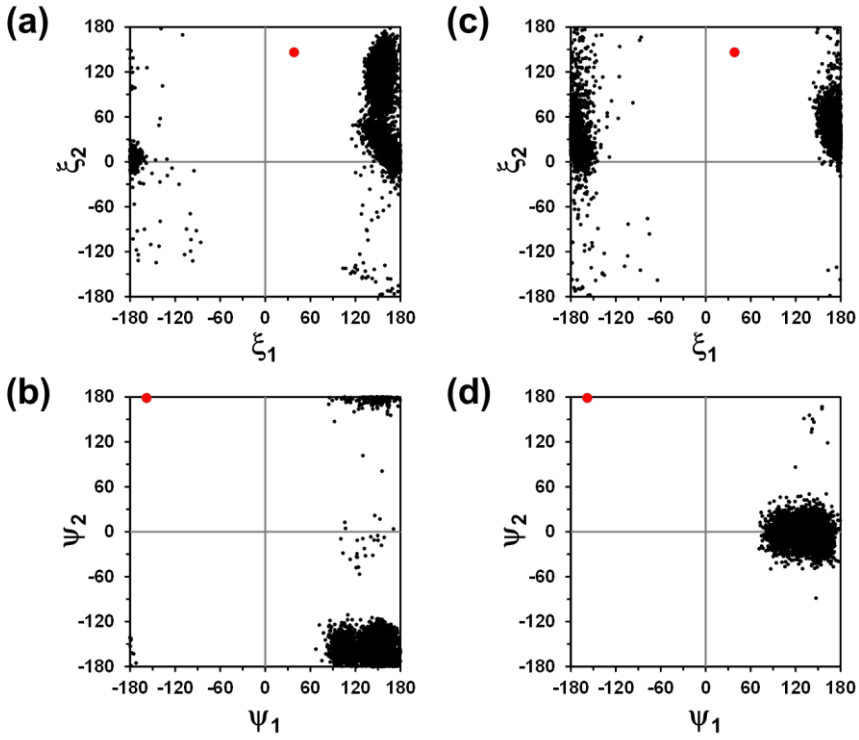


Figure 4.6. For (a, b) MD_{100} and (c, d) MD_{111} trajectories: (a, c) ξ_1–ξ_2 and (b, d) ψ_1–ψ_2 map distribution for RPAR considering all the recorded snapshots. Dihedral angles are defined in the text. The red dot at each map indicates the position of the proposed bioactive conformation.

Figure 4.6 represents the ξ_1–ξ_2 and ψ_1–ψ_2 maps for snapshots recorded during MD_{100} (Figures 4.6a-b) and MD_{111} (Figures 4.6c-d) simulations. The ξ_1–ξ_2 maps are fully consistent with the RMSDs (Figure 4.2) and the results of the clustering analyses (Figure 4.4).

Independently of the two gold surface facets, ξ_1 values are comprised between +150º and -150º, reflecting the low conformational flexibility of the peptide backbone. In contrast, ξ_2 values range from -30º to 180º, demonstrating that the conformational variability of RPAR must be ascribed to the side groups of the Arg residues. Similarly, ψ_1–ψ_2 maps prove the very low flexibility of the peptide backbone, even though the preferred ψ_2 values are influenced by the gold surface facet. More specifically, the ψ_2 visits the region that extends from 180º to -120º during the MD_{100} trajectory, while this dihedral is comprised between 40º and -40º through the whole MD_{111} simulation.

Figures 4.7a and 4.7b show representative snapshots (*i.e.* conformations with the ξ_1–ξ_2 and ψ_1–ψ_2 pairs of dihedrals located at the most populated regions among those displayed in Figure 4.6) extracted from MD_{100} and MD_{111}, respectively. The backbone tends to adopt a pseudo-extended conformation independently of the facet of gold, while charged groups are exposed to the gold surface and/or the solvent. More specifically, the ammonium and carboxylate groups from the N- and C-terminus, respectively, tend to point to the solvent, whereas the guanidinium moieties of the Arg residues are parallel to the gold surface and, therefore, interact with both the metal and the solvent. Indeed, the side group of the two Arg residues also prefers an almost extended conformation. Amazingly, in spite of the high density of positively charged groups, no salt bridge with the end carboxylate group was formed during MD simulations (see below).

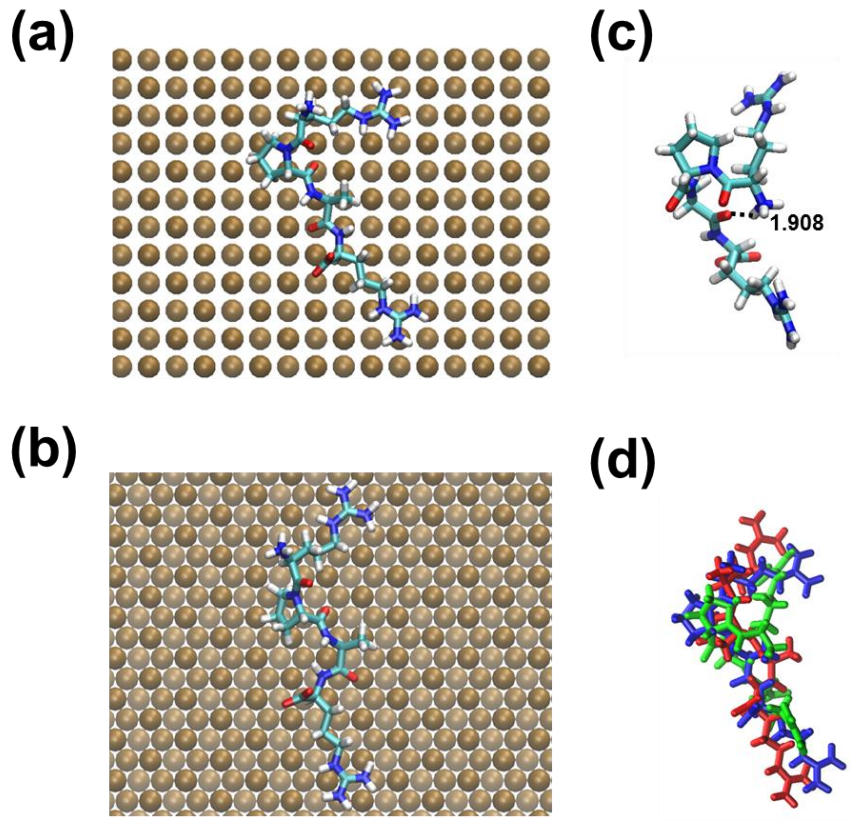


Figure 4.7. RPAR adsorbed onto gold facets: representative snapshots taken from (a) MD_{100} and (b) MD_{111} simulations. Solvent molecules were eliminated for clarity. (c) Bioactive conformation of RPAR (coordinates taken from references 8 and 9). (d) Superposition of the bioactive conformation (green) and representative conformations taken from MD_{100} (blue) and MD_{111} (red).

4.4.4 Comparison with the bioactive conformation

The bioactive conformation of RPAR, which is depicted in Figure 4.7c, was identified in previous work.[8,9] In this conformation the charged side groups of the two Arg residues and the carboxylate of the C-terminus are exposed to the pocket of the NRP-1 receptor (*i.e.* they are not involved in intramolecular interactions). In

addition, the bioactive conformation exhibits a hydrogen bond between the N-terminus and the C=O of Ala.

In the bioactive conformation the values of ξ_1 and ξ_2 are 38.4º and 146.2º, respectively, while ψ_1 and ψ_2 are -158.0º and 178.6º, respectively. These values, which are represented by red dots in Figure 4.6, are not included among the ξ_1–ξ_2 and ψ_1–ψ_2 regions visited during MD_{100} and MD_{111} trajectories. Thus, the gold substrate deeply affects the conformation of RPAR with respect to their preferences in solution and complexed with the NRP-1 receptor.[8,9] Differences between the bioactive and adsorbed RPAR conformation are illustrated in Figure 4.7d, which displays the superposition between the bioactive conformation and representative snapshots extracted from MD_{100} and MD_{111} trajectories.

It is worth noting that the conformational restrictions induced by the substrate are not in detriment of the use of gold nanoparticles as appropriated vehicles for the transport and targeted delivery of the RPAR sequence. On the contrary, the backbone conformational restriction induced by gold surface facets represents a major advantage to improve the cell penetration efficiency. Thus, the elongated conformation found for the peptide adsorbed onto gold is not stabilized by specific intramolecular interactions (*i.e.* salt bridges or hydrogen bonds; see below), which is crucial to facilitate its evolution towards the bioactive conformation. More specifically, once their high affinity for the NRP-1 receptor induces the targeted delivery of peptide molecules from the gold nanoparticles, both the absence of specific intramolecular interactions and the low

conformational variability are expected to favor the successful interaction with the receptor through the bioactive conformation. Thus, the rearrangements required to reach the bioactive conformation are significantly easier (*i.e.* the energetic cost is lower) than for peptide molecules coming from the solution state, in which the RPAR sequence was found to exhibit both high conformational variability and frequent intramolecular interactions.[8,9] Overall, these features are fully consistent with recent experimental observations on gold nanoparticles functionalized with the RGDK CendR peptide,[20] which reflected the efficient bio-effects of these functionalized nanocarriers on prostate cancer cells.

4.4.5 Analysis of interactions

4.4.5.1 Intramolecular peptide–peptide interactions. Salt bridges are among the most important interactions in peptide and protein chemistry. Ion pairs frequently act as binding sites in enzymes, mediate molecular recognition, modulate the allosteric effect and play a role in stabilizing secondary-structural elements. In this sub-section we have analyzed the spurious formation of a salt bridge between the C-terminus carboxylate and the guanidinium moiety of the Arg or the N-terminus ammonium. Figures 4.8a and 4.8b represent the temporal evolution of the three possible $^{-}OOC\cdots N^{+}$ distances for MD_{100} and MD_{111} trajectories, respectively. These distances are very large (*i.e.* > 6.5 Å) in all cases evidencing that salt bridges are not formed at any point in the trajectories.

On the other hand, the presence of intramolecular N–H···O=C and $^{+}$N–H···O=C hydrogen bonds have been examined by analyzing the temporal evolution of the H···O distance for all possible hydrogen bonding pairs. The geometric criteria considered for the identification of such interactions are: $d_{H\cdots O} \leq 3.0$ Å and $\angle$N–H···O $\geq 120^{o}$. Analyses of the trajectories show that intramolecular hydrogen bonds are infrequent, even though they are influenced by the gold surface topology. More specifically, hydrogen bonds were practically inexistent along the MD_{100} simulation, the accumulated lifetime of the three unique observed interactions ranging between 0.04 and 0.25 ns only (not shown). In contrast, analysis of the MD_{111} simulation revealed a very stable intra-residue hydrogen bond involving the first N–H moiety of the Arg1 side group and the backbone C=O moiety of the Pro residue. Figures 4.8c and 4.8d, which represent the temporal evolution of $d_{H\cdots O}$ and $\angle$N–H···O, evidences that the accumulated life time of this intra-residue interaction is 30.2 ns.

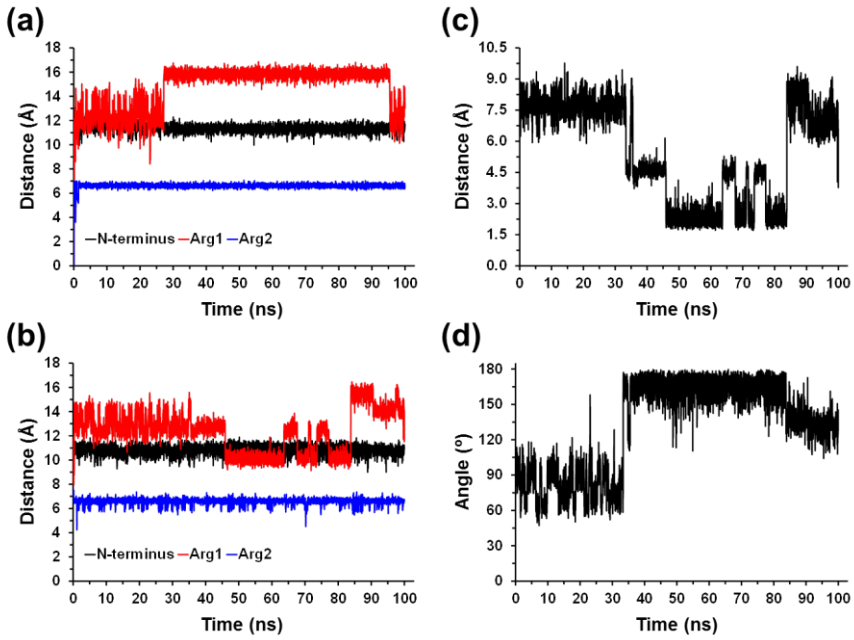


Figure 4.8. For (a) MD_{100} and (b) MD_{111}, temporal evolution of the N···C distance for the three possible $^{-}OOC\cdots N^{+}$ intramolecular salt bridges, where N^{+} corresponds to the ammonium group of the N-terminus or the guanidinium groups of Arg1 and Arg2 (see text). Temporal evolution of (c) the H···O distance and (d) the ∠N–H···O angle in the only intramolecular hydrogen bond observed along MD_{111}, which involves the first N–H moiety of the Arg1 side group and the backbone C=O moiety of the Pro residue.

4.4.5.2 Peptide–solvent interactions. The specific interactions formed between the different moieties of RPAR and the surrounding water molecules have been evaluated by calculating the partial radial distribution functions (RDFs) with respect to the oxygen or hydrogen atom of water for each trajectory. The RDFs obtained considering, on the one hand, the nitrogen of the N-terminus and the oxygen atom of water ($g_{Nt\text{-}O}$), and on the other hand, the carbon atom of C-terminus and the hydrogen atom of

water ($g_{Ct\text{-}H}$), are displayed in Figures 4.9a and 4.9b, respectively. The resulting profiles indicate that the two terminal groups present very similar hydration patterns that, in addition, are independent of the gold facet. Thus, the $g_{Nt\text{-}O}$ profiles display a very sharp and narrow peak at 2.75 Å and a broad peak at 5.05 Å, which clearly correspond to the first and second hydration shells. Similarly, the $g_{Ct\text{-}H}$ profiles show a sharp peak at 2.65 Å and a less intense and broad peak at 4.05 Å. According to these results, the ammonium and carboxylate terminal groups are surrounded by two ordered water shells. The hydration patterns obtained for the charged side groups of Arg1 and Arg2, which are described by the RDFs calculated using the central carbon of the guanidinium group (CZ) and the oxygen atom of water ($g_{CZ\text{-}O}$), exhibit comparable characteristics. Thus, the $g_{CZ\text{-}O}$ curves, which are displayed in Figures 4.9c and 4.9d for Arg1 and Arg2, respectively, shows a sharp peak at 3.85 Å that corresponds to a well-defined cluster of water molecules directly bounded to these charged side groups (*i.e.* first solvation shell). After this, a broad peak appears at ~6.1 Å, coinciding with the less ordered second solvation shell, is observed for the four calculated $g_{CZ\text{-}O}$ profiles.

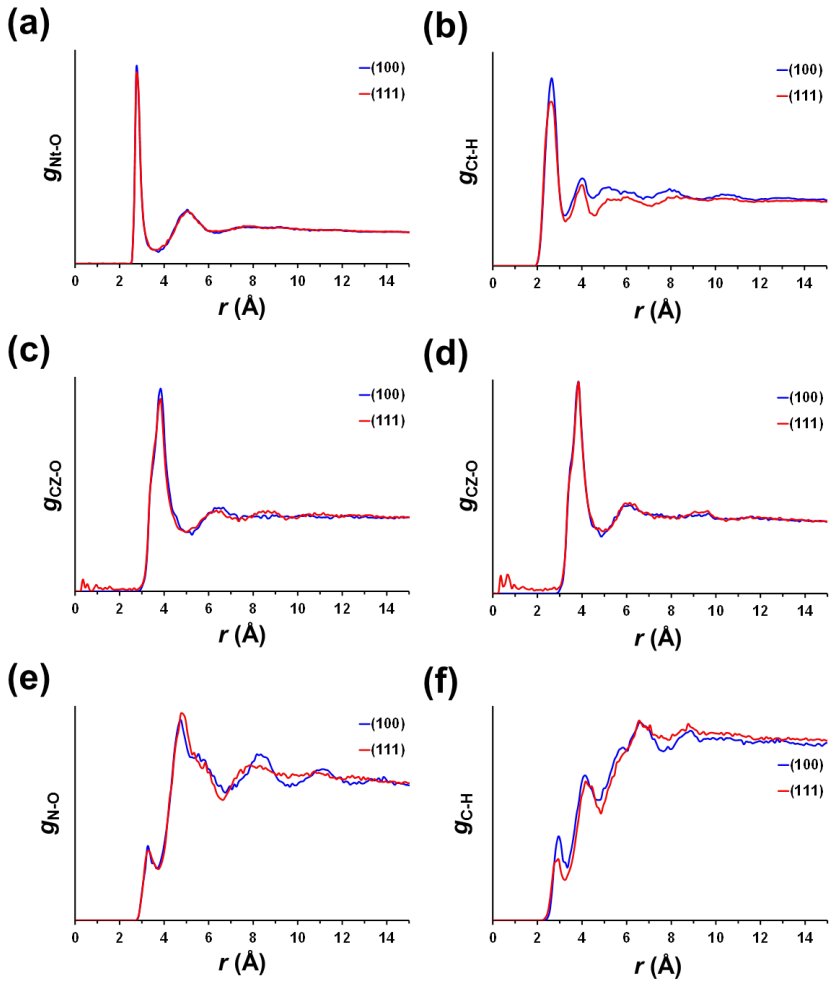


Figure 4.9. For MD_{100} and MD_{111} simulations, analysis of the RPAR···water interactions considering the RDFs for the following pairs: (a) the nitrogen of the N-terminus and the oxygen atom of water (g_{Nt-O}); (b) the carbon atom of C-terminus and the hydrogen atom of water (g_{Ct-H}); (c) the central carbon of the guanidinium side group of Arg1 and the oxygen atom of water (g_{CZ-O}); (d) the central carbon of the guanidinium side group of Arg2 and the oxygen atom of water (g_{CZ-O}); (e) the backbone nitrogen atom of Ala and the oxygen of water (g_{N-O}); and (f) the carbon atom from the backbone C=O of Ala and the hydrogen of water (g_{C-H}).

A different hydration pattern is detected when non-charged polar groups, as for example the N–H and C=O moieties of the backbone amide groups, are considered. This is illustrated in Figures 4.9e and 4.9f, which display the RDFs calculated considering the N–H and C=O moieties of the Ala residue. The intensity of the peak associated to the first hydration shell (r= 3.2 Å) in the profiles obtained using the nitrogen atom of N–H and the oxygen of water ($g_{N\text{-}O}$) is significantly lower than that of flat region associated to disordered water molecules in the bulk liquid. The same feature is observed in the profiles calculated considering the carbon atom of the C=O and the hydrogen of water ($g_{C\text{-}H}$), even though in this case the peak of the second hydration shell (r= 4.1 Å) is also below the bulk water profile. Overall, these results reflect an expected trend: solvent molecules form stronger specific interactions with charged peptide groups than with polar neutral groups of the peptide backbone. Moreover, RPAR⋯water interactions are practically independent of the gold surface used for the peptide adsorption, as is clearly evidenced in Figures 4.9a-f.

4.4.6 Structural parameters

The temporal evolution of the radius of gyration (R_g) and the hydrodynamic radius (R_h) of RPAR along MD_{100} and MD_{111} simulations is displayed in Figure 4.10. R_h is usually defined as the radius of an equivalent hard sphere diffusing at the same rate as the molecule under observation, whereas R_g is the mass weighted average distance from the core of a molecule to each mass element in the molecule. The ratio of R_g and R_h provides shape information

the molecule. The characteristic R_g/R_h value for spherical molecules is ~0.77, this value being extracted from the analysis of globular proteins. This means than R_g is smaller than R_h. However, when molecules deviate from spherical to non-spherical or elongated conformations, then R_g/R_h tend to values upwards of 0.77, as R_g becomes larger.

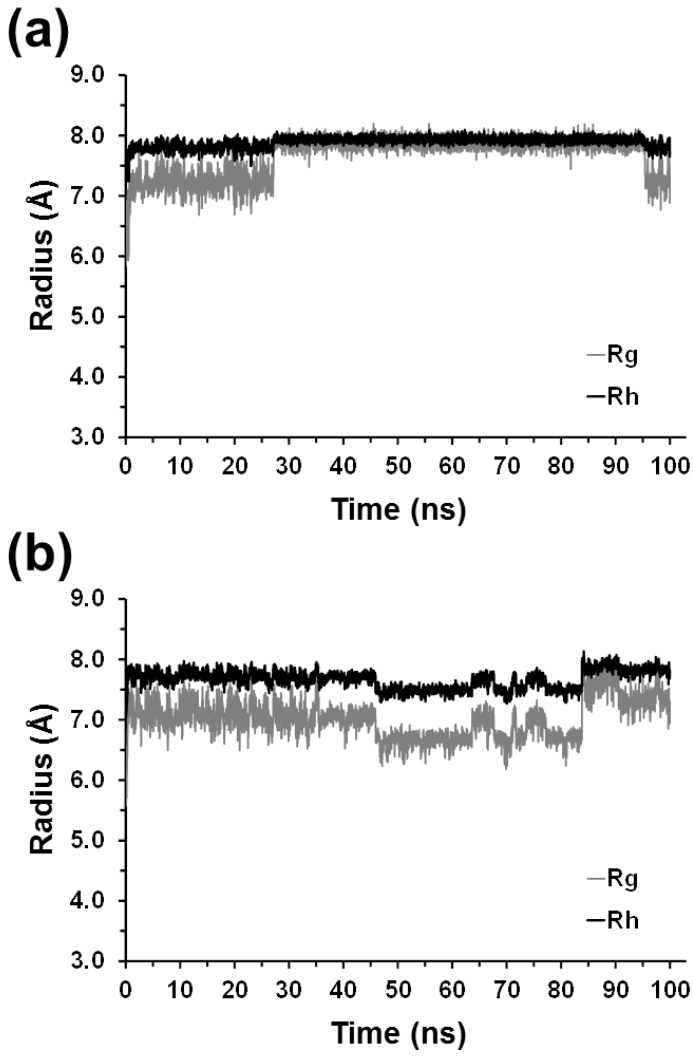


Figure 4.10. For (a) MD_{100} and (b) MD_{111}: Temporal evolution of R_g and R_h.

The average values of R_g and R_h for RPAR adsorbed onto the (100) gold surface are 7.69±0.31 Å and 7.89±0.31 Å, respectively, decreasing to 7.01±0.31 Å and 7.66±0.15 Å, respectively, when the adsorption is onto the (111) surface. Accordingly, R_g/R_h is 0.97 and 0.91 for the peptide adsorbed onto the (100) and (111) surface, respectively. These ratios are consistent with both: *i)* the elongated conformation found for the adsorbed peptide (Figures 4.7a-4.7b), in which Arg1 and Arg2 side groups act as an extension of the backbone enhancing such elongated shape; and *ii)* the first and second water shells detected for the charged groups of both the backbone and side chain (Figures 4.9a-4.9d). The influence of *i)* in the mass distribution is higher than those of both *i)* and *ii)* in the value of R_h and, therefore, R_g/R_h grows to values above 0.77 but below 1.

4.5 CONCLUSIONS

In conclusion, in this work we have demonstrated that the RPAR CendR sequence, which has promising therapeutic uses as internalization trigger in anticancer nanodevices, binds both the (100) and (111) gold surface facets. Accordingly, gold nanoparticles can be considered as suitable vehicles for the transport and targeted delivery of this CendR peptide. The conformation of the adsorbed peptide differ considerably from the bioactive conformation, which was derived in previous work using as a molecular template the crystal structure of NRP-1 bound to tuftsin.[3] However, the conformations adopted by the adsorbed peptide, which display elongated-like shapes, are not stabilized by strong intramolecular

interactions, like salt bridges or specific hydrogen bonds. This feature favors, despite the changes introduced in the conformation, the utilization of gold nanoparticles as carriers for the targeted delivery of the RPAR sequence. Thus, the elongated-like conformation of the peptide adsorbed onto gold nanoparticles can easily change towards the bioactive conformation once the affinity for the NRP-1 receptor induces their targeted delivery. The absence of intramolecular salt bridges and hydrogen bonds reduces considerably the energy penalty associated to the structural rearrangements necessary for such conformational transition.

4.7 REFERENCES

1. Loomis, K., McNeeley, K. & Bellamkonda, R. V. Nanoparticles with targeting, triggered release, and imaging functionality for cancer applications. *Soft Matter* **7,** 839–856 (2011).
2. Chari, R. V. J. Targeted cancer therapy: Conferring specificity to cytotoxic drugs. *Acc. Chem. Res.* **41,** 98–107 (2008).
3. Teesalu, T., Sugahara, K. N., Kotamraju, V. R. & Ruoslahti, E. C-end rule peptides mediate neuropilin-1-dependent cell, vascular, and tissue penetration. *Proc. Natl. Acad. Sci.* **106,** 16157–16162 (2009).
4. Takashima, S. *et al.* Targeting of both mouse neuropilin-1 and neuropilin-2 genes severely impairs developmental yolk sac and embryonic angiogenesis. *Proc. Natl. Acad. Sci.* **99,** 3657–3662 (2002).
5. Staton, C. A., Kumar, I., Reed, M. W. R. & Brown, N. J. Neuropilins in physiological and pathological angiogenesis. *J. Pathol.* **212,** 237–248 (2007).
6. Sugahara, K. N. *et al.* Coadministration of a tumor-penetrating peptide enhances the efficacy of cancer drugs. *Science* **328,** 1031–5 (2010).
7. Vander Kooi, C. W. *et al.* Structural basis for ligand and

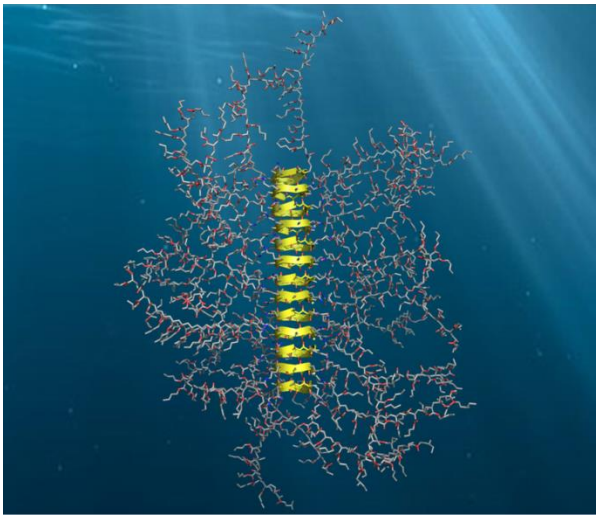
heparin binding to neuropilin B domains. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 6152–7 (2007).

8. Haspel, N. *et al.* Binding of a C-end rule peptide to the neuropilin-1 receptor: A molecular modeling approach. *Biochemistry* **50,** 1755–1762 (2011).
9. Zanuy, D. *et al.* Sequence dependence of C-end rule peptides in binding and activation of neuropilin-1 receptor. *J. Struct. Biol.* **182,** 78–86 (2013).
10. Simberg, D. *et al.* Biomimetic amplification of nanoparticle homing to tumors. *Proc. Natl. Acad. Sci.* **104,** 932–936 (2007).
11. Zanuy, D. *et al.* The energy landscape of a selective tumor-homing pentapeptide. *J. Phys. Chem. B* **112,** 8692–8700 (2008).
12. Zanuy, D., Curcó, D., Nussinov, R. & Alemán, C. Influence of the dye presence on the conformational preferences of CREKA, a tumor homing linear pentapeptide. *Biopolymers* **92,** 83–93 (2009).
13. Agemy, L. *et al.* Nanoparticle-induced vascular blockade in human prostate cancer. *Blood* **116,** 2847–2856 (2010).
14. Curcó, D., Zanuy, D., Nussinov, R. & Alemán, C. A simulation strategy for the atomistic modeling of flexible molecules covalently tethered to rigid surfaces: Application to peptides. *J. Comput. Chem.* **32,** 607–619 (2011).
15. Connor, E. E., Mwamuka, J., Gole, A., Murphy, C. J. & Wyatt, M. D. Gold nanoparticles are taken up by human cells but do not cause acute cytotoxicity. *Small* **1,** 325–327 (2005).
16. Kumar, A. *et al.* Gold nanoparticles functionalized with therapeutic and targeted peptides for cancer treatment. *Biomaterials* **33,** 1180–1189 (2012).
17. Tiwari, P. M. *et al.* Enhanced intracellular translocation and biodistribution of gold nanoparticles functionalized with a cell-penetrating peptide (VG-21) from vesicular stomatitis virus. *Biomaterials* **35,** 9484–9494 (2014).
18. Paciotti, G. F., Kingston, D. G. I. & Tamarkin, L. Colloidal gold nanoparticles: A novel nanoparticle platform for developing multifunctional tumor-targeted drug delivery vectors. *Drug Dev. Res.* **67,** 47–54 (2006).
19. Ghosh, P., Han, G., De, M., Kim, C. K. & Rotello, V. M. Gold nanoparticles in delivery applications. *Adv. Drug Deliv. Rev.* **60,** 1307–1315 (2008).

20. Kumar, A. *et al.* Neuropilin-1-targeted gold nanoparticles enhance therapeutic efficacy of platinum(IV) drug for prostate cancer treatment. *ACS Nano* **8,** 4205–4220 (2014).
21. Duan, Y. *et al.* A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **24,** 1999–2012 (2003).
22. Iori, F., Di Felice, R., Molinari, E. & Corni, S. GolP: An atomistic force-field to describe the interaction of proteins with Au(111) surfaces in water. *J. Comput. Chem.* **30,** 1465–1476 (2009).
23. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926–935 (1983).
24. Barnard, A. S., Lin, X. M. & Curtiss, L. A. Equilibrium morphology of face-centered cubic gold nanoparticles >3 nm and the shape changes induced by temperature. *J. Phys. Chem. B* **109,** 24465–24472 (2005).
25. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).
26. Toukmaji, A., Sagui, C., Board, J. & Darden, T. Efficient particle-mesh Ewald based approach to fixed and induced dipolar interactions. *J. Chem. Phys.* **113,** 10913–10927 (2000).
27. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81,** 3684–3690 (1984).
28. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101,** 4177–4189 (1994).
29. Andersen, H. C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **52,** 24–34 (1983).

- Chapter 5 -

Peptide –Polymer Conjugate Material: Understanding its Conformational Preferences in Front of the Degree of Solvation



5 PEPTIDE –POLYMER CONJUGATE MATERIAL: UNDERSTANDING ITS CONFORMATIONAL PREFERENCES IN FRONT OF THE DEGREE OF SOLVATION[§]

5.1 ABSTRACT

The syndiotactic cyclic peptide cyc[(L-Gln-D-Ala-L-Lys-D-Ala)$_2$] coupled with two poly(*n*-butyl acrylate) blocks is reported to form a self-assembled material by the coherent association of the cyclic peptide to produce a nanotube-like structure with the acrylate chains protruding from the cylindrical scaffold provided by the peptide structure. In previous works, the self -assembly structural details of this conjugated material in several environments were studied in Silico. However, the conformational preferences of the poly(*n*-butyl acrylate) segments when adsorbed on inorganic surfaces did not reproduce the experimental observations and did not provide a sufficient explanation for such unexpected results. While the parent polymer, poly(*n*-butyl acrylate), adopts a fully extended conformation once adsorbed on mica, X-ray diffraction data had showed that when those segments were part of the conjugated material were partially folded. In this work we report an atomistic detailed examination of the deposition process. We have being able to demonstrate that this behavior can only be understood if solvent traces are present while the material was been deposited. Under these conditions, the solvent molecules compete with other species present in the system bulk for forming attractive van der Waals interactions with the acrylate segments before and during its

[1] This work has been has been submitted for publication.

adsorption on mica. This competition interferes with the known affinity of the acrylate segments towards the inorganic surface and conditions their structural preferences. Hence, the acrylate conjugated segments will only behave as the parent polymer if the deposition takes place after removing the solvent molecules. If solvent traces are present, their structural preferences will solely be driven by the amount of solvent molecules that can interact around those acrylate blocks.

5.2 INTRODUCTION

Comprehension of microscopic driving forces very often holds the key for further molecular engineering. Understanding the details of how macromolecules assemble, not only the association geometries but also the path towards a proper assembly, becomes essential for improving those features that might turn to be interesting for its industrial use. In this context, researching on materials that are capable to consistently self-assemble has been the focus of many studies for the las to decades.[1,2] Within this group of materials peptides of natural origin emerge as especially suited constituents.[3] They are abundant in nature and they self-assemble into fibrous structures under mild physical conditions. Conjugating peptides with other macromolecules of synthetic origin has become almost a new technological branch within the macromolecules and material sciences.[4,5] The main drawback of this strategy lies on the conformational flexibility of many of those peptides that quite often drives them to aggregate into amorphous complexes and limits their practical. To that end, several approaches have been developed to

structurally force the peptide to adopt conformations compatible with β-strand arrangements, strategy known as preorganization. Among different preorganization strategies, one of the most popular is to cyclize the main chain of a peptide. This approach is especially efficient when the peptide chirality is alternated within the main chain (known as *D-alt-L*) and the number of residues in the peptide matches an even number. These two features are present in the peptide constituent of the conjugated material studied in this work.

This substance is obtained by the self-assembly of a peptide-polymer conjugate that consists of a cyclic syndiotactic octapeptide, cyc[(L-Gln-D-Ala-L-Lys-D-Ala)$_2$], with a poly(*n*-butyl acrylate) blocks, (P*n*BA), tethered to each Lysine side chains (see a schematic representation in **Figure 5.1a**).[6] From this point, in order to simply the readers task, we would refer to the peptide segment as the *cyclic peptide*, the poly(*n*-butyl acrylate) blocks as the *acrylate side chains*, each conjugated block as the *assembly unit* and the whole assembled system as the *nanotube-like*. Therefore, the stack of assembly units leads to the final material.

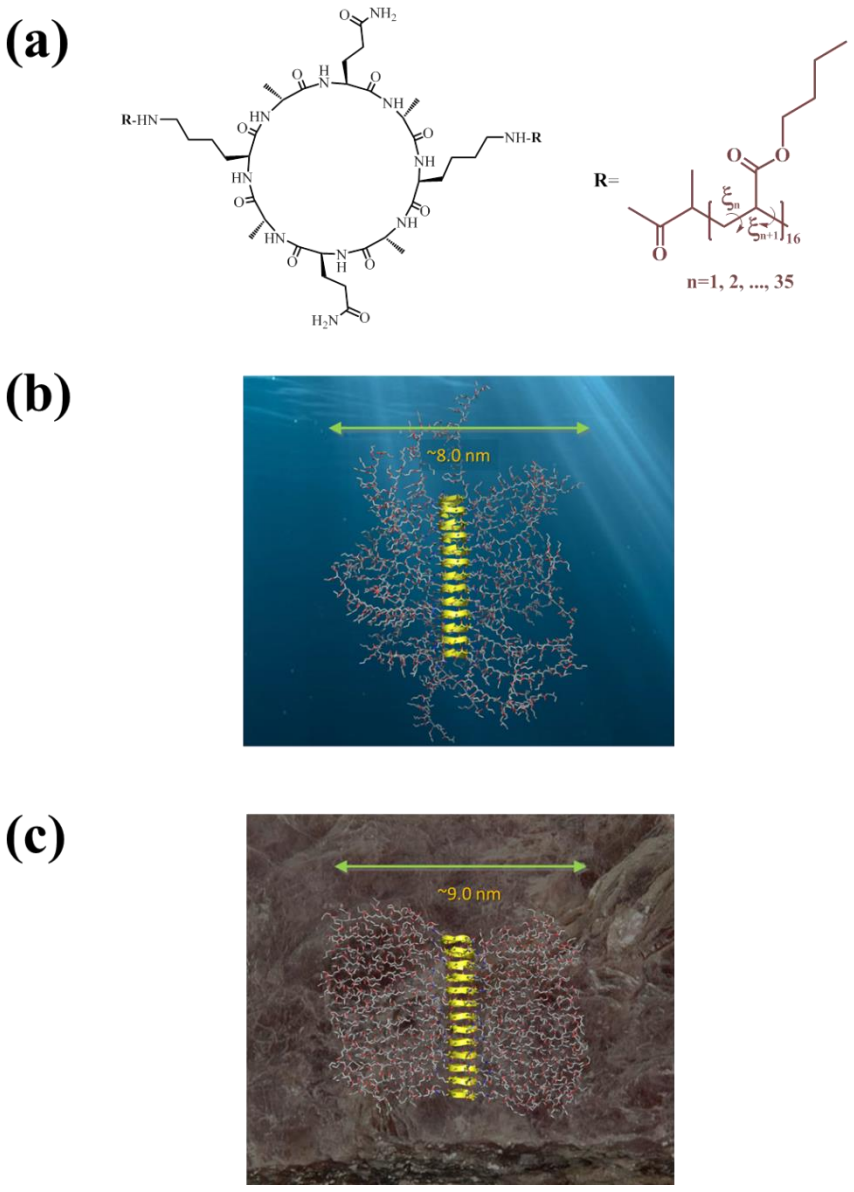


Figure 5.1.- (a) Chemical structure of the self-assembly unit constituted by cyc[(L-Gln-D-Ala-L-Lys-D-Ala)$_2$]-(PnBA)$_2$. Right Chart shows a detail of the free rotation dihedral angles of the acrylate segments. **(b)** Molecular representation of the nantube-like organization and apparent diameter in solution. **(c)** Molecular representation of the nantube-like organization and apparent diameter deposited on mica.

The general structural features of the nanotube-like were studied by a wide collection of experimental techniques that provided a global description of its size and shape in different environments.[6] Yet, the impossibility of getting atomistic details of its ultrastructure made not possible to determine several fundamental details that could condition its large scale production. The combination of the inherent limitations of the spectroscopic techniques such FTIR and the scale limits of microscopy, either classical TEM images or AFM visualizations, let some questions about its structural preferences unanswered. Some of them were partially addressed in a first computational study,[7] in which a thorough investigation explored the majority of features that the resolution of the current microscopic techniques could not address. After studying the relative orientation of the association of the cyclo-peptide unit, the parallel orientation was discarded and the inferred model from FTIR and X-ray scattering experiments[6] could be confirmed. Independently of the environment (gas phase, in organic solvent or deposited over mica) the nanotube-like that was constructed upon stacking peptide rings in antiparallel fashion consistently showed much better stability than the assembled in parallel orientation.

Having set the structural preferences of the peptide domain, the conformational preferences of the acrylate side chains were investigated. Experimental observations showed a substantial increment in the apparent length of these side chains when the nanotube-like was deposited on a mica surface. Dynamic light scattering measurements of the self-assembled nantube in DMF solution showed that the average diameter of the whole nanotube-

like was about 8.0nm, length that is slightly shorter than what should correspond to a fully extended conformation oriented outwards of the peptide scaffold. [6] Molecular Dynamics simulations in both DMF and chloroform showed a clear tendency to present extended conformations for the acrylate side chains, although in any case the distribution of chains was even around the peptide cyclic core. On the contrary, as **Figure 5.1b** shows, most of them tended to remain in close contact with side chains belonging to neighboring assembly units. Moreover, only the finite nature of the model allowed the side chains located at the edges of the nantube-like to freely protrude outbound.[7]

On the other side, electron diffraction experiments performed on selected areas of TEM images of the nanotube-like deposited on mica demonstrated that the arrangement of its side chains did not correspond to a fully extended conformation: the average diameter of the nanotube-like under those conditions only reached values around 5.0 nm.[6] This value represented a contraction of about 38% in the observed diameter of the nanotube-like compared with its organization in solution. Therefore, the conclusion drawn was that the physical attraction between the inorganic surface and the acrylate side chains, which had already been observed in previous work,[8] must had been somehow responsible for the partial contraction in the nanotube-like diameter. However, this interpretation did not account for the fact that the observed conformation changes were limited. The measured diameter remained distant from the maximum extension that would be expected if the contact area was the sole responsible for the

conformational behavior of the acrylate side chains. In that case, the observed length increase should reach diameter lengths much closer to 10.0 nm (which is the approximate diameter of two fully extended linear alkyl chain with 35 carbon atoms tethering from a cyclic peptide of 1.4 nm of diameter)[6] This important discrepancy was rationalized as a consequence of the topological limitations imposed to the side chains by being evenly distributed cyclic peptide frame. In other words, those side chains tethering from sites away of the contact area would need to bend inwards in order to contact with the surface, reducing their apparent length. Molecular dynamics simulations performed in gas phase conditions of the nantube-like deposited on mica apparently contradicted the previous interpretation.[7] The computed average diameter when the nanotube-like lied over mica was close to 10 nm, as should be expected if the interaction between the acrylate side chains and mica surface was the driving force that conditioned the acrylate groups conformation. **Figure 5.1c** shows that even those acrylate segments that needed bending to contact the ceramic surface do not reduce their apparent length to values that matched the observed contraction when the material was deposited on mica.

Therefore, some crucial aspect of the deposition process might have been missed in the MD simulations previously performed. At this point, new scenarios unfold and require to be explored: either the deposition process was not completely solvent free as was initially inferred[7] or the deposition of the material was performed prior to the complete removal of solvent for latter AFM measurements.[6] In this work, we focus on how the deposition

process and the conformational preferences of the acrylate side chains can be affected by how the solvent is eliminated. In order to assess if the presence of residual solvent was the responsible of the unexpected contraction of the acrylate tales, we envisaged a quasi-reversible extraction / substitution of DMF solvent molecules by inert particles that will allow evaluating the effect of low concentration of active solvent over the deposition of the nanotube-like on mica. This work will show that DMF molecules are the main responsible of the unexpected conformation preferences and how their gradual disappearance is the main factor that induces that contraction, which becomes irreversible once the partially folded chains start interacting with the mica surface when the DMF is almost removed.

5.3 METHODS

5.3.1 Force Field details

All parameters were obtained from AMBER03 libraries[9] with exception of non-bonding parameters of the ceramic surface, which were obtained from the specific parametrization done by Heinz et al.[10] For the inert particle (see text below for details), a combination of a big radius with a low hardness was chosen ($r_P = 3.565$ Å, $\varepsilon_P = 0.01$), in order to guarantee a constant volume occupancy and a low energy perturbation. The chosen radius corresponds to the approximate radius that a single DMF molecule would present if it is represented as a spherical particle: the average distance of the center nitrogen atom to any of the carbon atoms is 1.4 Å[9] and the

average radius of a methyl group is 2.165Å,[9] the sum of both being the chosen value.

Periodic boundary conditions were applied using the nearest image convention and atom pair cut-off distance was set at 14.0 Å to compute the van der Waals interactions. In order to avoid discontinuities in the potential energy function, non-bonding energy terms were forced to slowly converge to zero, by applying a smoothing factor from a distance of 12.0 Å. Beyond cut off distance, electrostatic interactions were calculated by using Particle Mesh of Ewald, with a points grid density of the reciprocal space of $1Å^3$.[11]

5.3.2 Molecular Dynamics details

All simulations were performed using the NAMD 2.9 program.[12] Before a new system was simulated, the energy was relaxed by performing 5000 steps of minimization (Newton Raphson method). The numerical integration step for all runs in this work was set at 2 fs and the non-bonded pair list was updated every 1000 steps (2 ps).

The following protocol was applied to all studied conditions: Solvent phase was first equilibrated. 10^5 steps of heating and equilibration at NVT conditions were run to stabilize the solvent phase at 298K. The Berendesen thermostat was used[13] with a relaxation time of 1 ps. Then, the solvent density was equalized to its optimum value using $2.5 \cdot 10^5$ steps of NPT simulation at 298 K. The Nose–Hoover piston[14] combined with the piston fluctuation control of temperature implemented for Langevin Dynamics[15] was used in the late cycle. Pressure was kept at 1.01325 bars, the

oscillation period was set at 1 ps while the piston decay time was set at 0.001 ps. The piston temperature was set at the same value as the thermostat control, 298K, which used a damping coefficient of 2 ps. During these couple of cycles, all non-solvent atoms were kept frozen (mica surface and nanotube-like system).

Once the simulation box dimensions were stabilized, the final equilibration mini-cycles would begin. After unfreezing the rest of atoms (with exception of the mica atoms[7]), $2 \cdot 10^5$ steps of NVT simulation were performed to thermally equilibrate the whole system. For the latter run the Langevin method[16] was used to maintain the system temperature constant with a damping coefficient of 2 ps. Finally, $1.0 \cdot 10^6$ steps of anisotropic pressure control (NP*z*T conditions) were run to reach the required simulation conditions. In the NP*z*T ensemble, only the box length in the *z*-direction is allowed to change, where the *z*-component of the pressure tensor is equal to the external pressure.[17] The last snapshot of the anisotropic equilibration was the starting point of the first 10 ns of production time.

The production runs were all performed under NVT conditions, using Langevin method[16] to keep the temperature constant with a damping coefficient of 2 ps. This approach was chosen in order to avoid the formation of undesired bubbles of free space that would cause discontinuities in the energy function and the consequent crash of the simulation. However, right after replacement cycle, $3.0 \cdot 10^5$ steps of anisotropic pressure control (NP*z*T conditions) were run in order to equilibrate the box density to the new solvent

composition. Each production time, per each solvent mixture composition, was of 10 ns and a snapshot was saved every 10 ps.

5.3.3 The molecular models

The mica K $[Si_3Al_1O_8][Al_2O_2(OH)_2]$ surface was represented using the force-field parameters reported by Heinz and co-workers,[9] which reproduce the unit cell properties of this material within the AMBER force field framework.[10] The mica super cell model was adapted to the required dimensions of the simulation box: a= 207.675 Å and b= 216.367 Å.[7] The thickness (i.e. extension in the *c*-direction) of this sheet model was 6.7 Å. In order to avoid the bending of the mica sheet during the MD simulations, in all studied cases the position of the ions at the surface was fixed at the equilibrium position.

The simulation box was obtained from a previously equilibrated DMF simulation box of initial dimensions a= 207.675 Å, b= 216.367 Å and c=205 Å, that contained up to 69898 molecules of solvent. This initial box was equilibrated following the protocol described above, in which after an energy minimization, the temperature first, and then, density were equilibrated until reaching the experimental density of DMF at 298K (0.994 $g \cdot cm^{-3}$). Then the nanotube-like, constituted by 13 assembly units, and the mica surface were placed inside the simulation box, eliminating those solvent molecules that overlapped with any atom of the newly introduced solutes. The final simulation system was constituted by total of 799812 explicit atoms, containing 62420 molecules of DMF. The density of this initial model was again equilibrated at

298K and 1 atm of pressure, using the same protocol described above.

5.3.4 The quasi reversible desolvation

In order to reproduce a slow desolvation, which should provide us a better perspective over the cohesion forces that affect the conformational preferences of the acrylate side chains, the DMF solvent molecules were replaced by inert particles of similar size than DMF molecules at consecutive time periods, generating a biphasic solvent in which the newly added particles act as an inert atmosphere. Every 10 ns the solvent composition was modified in order to decrease the molar fraction of DMF one ten percent. Therefore, after 100 ns all the solvent molecules would have been replaced by inert particles, without losing the space occupancy or the molecular crowding. In order to avoid density discontinuities within the solvent bulk, once the density of the initial solvent bulk was equilibrated ($\chi_{DMF} = 1$), all the subsequent production runs were performed under NVT conditions. This procedure was selected ahead of any NPT approach to avoid the generation of free space bubbles once the non-bonding interactions of a cohesive solvent are modified. NVT conditions ensure the consistency of the simulation each time the solvent composition was modified. The composition of the solvent was systematically impoverished of DMF every 10 ns by a 10%, until all the remaining solvent is solely composed by inert particles.

5.4 RESULTS AND DISCUSSION

One of the main conclusions f previous investigations was that the interactions between the inorganic surface and the acrylate side chains are responsible for the conformational behavior of the latter segments.[6,7] Previous studies of branched acrylates[8] showed that these molecular species, when not conjugated with peptides, present extended arrangements while lying on mica surfaces. MD simulations in the gas phase (**Figure 5.1c**)[7] always led to highly extended organization for the acrylate segments, demonstrating that the peptide scaffold did not affect the acrylate structural preferences in absence of external elements, such a the mica substrate, or solvent. Yet, TEM images of the acrylate side chains showed that were always partially contracted.[6] Thus, the desolvation process followed to achieve the dried supported samples either alters the conformational preferences of the acrylate segments or does not completely eliminate the solvent prior to the nanotube-like deposition. In order to investigate the possible sources of this unexplained behavior we designed a quasi–reversible experiment in which the solvent used in previous experiments, DMF, would be systematically removed and substituted by inert particles with similar size to that of the solvent molecules. This way, any conformational change of the acrylate side chains due to the loss of interactions with DMF could be monitored. At the same time, this strategy would allow detecting any variation of the organization of those segments when the nanotube-like, as a whole, approaches to the surface. Previous works in gas phase were performed at low temperature and starting with the nanotube-like positioned

sufficiently close to the surface. This starting geometry allowed the edges of some of the acrylate side chains to almost contact mica right at the beginning of the simulation, presumably conditioning the further organization of the acrylate segments.

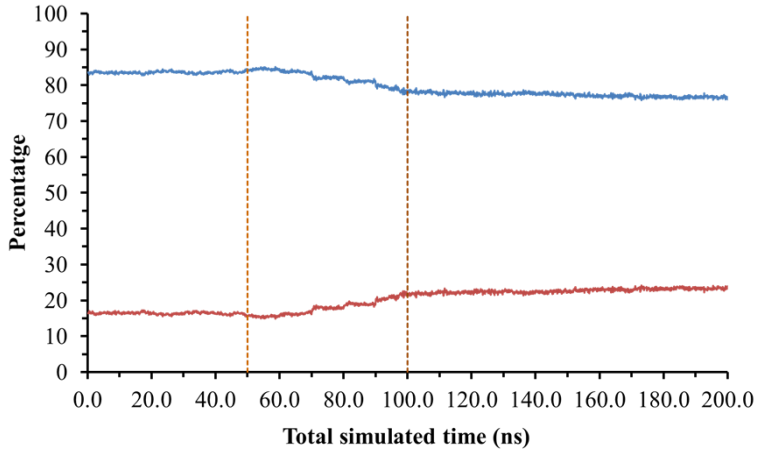


Figure 5.2.- Percentage of torsions of the acrylate segments that present *trans* conformation (blue continuous line) and *other* conformations (red continuous line) as function of the simulated time. Two vertical dashed lines indicate when the molar mass of DMF reaches 50% (left line) and when DMF is completely removed (right line).

As mentioned in previous sections, the first 110 ns of trajectory were divided in windows of 10 ns, in which the system was allowed to freely evolve and react to the solvent composition changes. Initially, 62420 molecules of solvent are present in the simulation box (χ_{DMF}= 1) and every 10 ns the amount of DMF molecules is reduced by a 10%. After 100 ns of simulation the molar fraction of DMF in the solution in which the nanotube-like is immersed is only 10% (χ_{DMF}= 0.1). In order to assess if the desolvation of the

material has an effect on the conformational arrangement of the acrylate side chains, a simple graphical monitoring was used. First, every dihedral angle with free rotation for each acrylate segment present in the studied model was computed for each saved frame. This amounts to 936 computed angles per studied frame: 13 assembly units, each of them with 2 acrylate segments, and 36 dihedral angles with free rotation per segment (Figure 5.1a). Second, all computed dihedral angles values were rank classified as belonging to a *trans* conformer (180°± 30º) or to any other arrangements (*i.e.* any value different than 180°± 30º will be classified as *other*). Third and last, for each frame the percentage of all torsions that were in *trans* were computed and graphically plotted versus time in **Figure 5.2**. This approach might have not produced any significant information, due to the huge amount of data incorporated in each frame. Apparently though, the conformational change is so severe than even this crude approach shows structural changes as the DMF fraction decreases. Figure 5.2 clearly demonstrates that the amount of segments that adopt *trans* conformations is directly correlated with the amount of DMF present in solvent the mixture. When the proportion of DMF molecules reaches 50%, the acrylate segments tend to arrange into contracted conformations. Moreover, this phenomenon is not correlated with the deposition onto mica, because the contraction occurs earlier than the nanotube-like gets in contact with the ceramic surface.

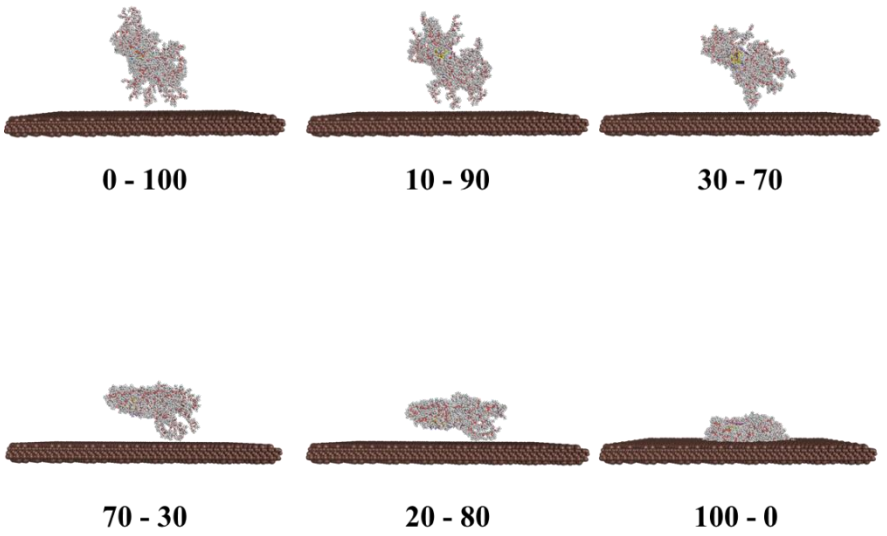


Figure 5.3.- Selected snapshots of representative geometrical arrangements of the nanotube-like respect to the mica surface depending on the DMF relative concentration in the solvent mixture. Legends at the bottom of each frame indicate the relative proportion (in %) of DMF versus inert particle).

Figure 5.3, which depicts representative snapshots of the desolvation process, shows the nanotube-like close to the surface but not in contact whereas the side chains have already started to fold (chart tagged as *20 - 80*). It is worth noting that while the acrylate segments remain unfolded, the conformational preferences are very restricted by the environment. Our data show that not only the *trans* arrangement is overwhelmingly preferred by the vast majority of the rotamers but also the fluctuations around the reference value 180° are very small, providing a narrow cluster of averaged dihedral angle values (for numerical details see **Figures 5.4-5.16** of *Supporting Information*).

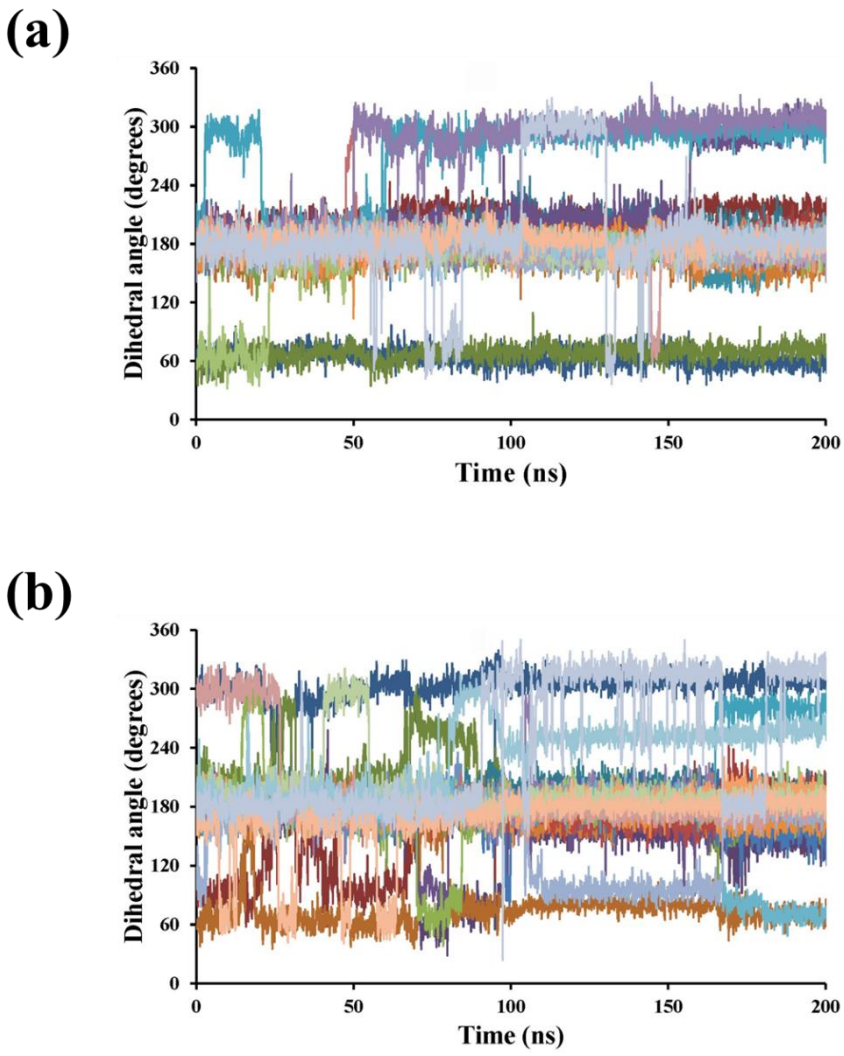


Figure 5.4.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **first** unit of the 13 used to simulate the nanotube-like.

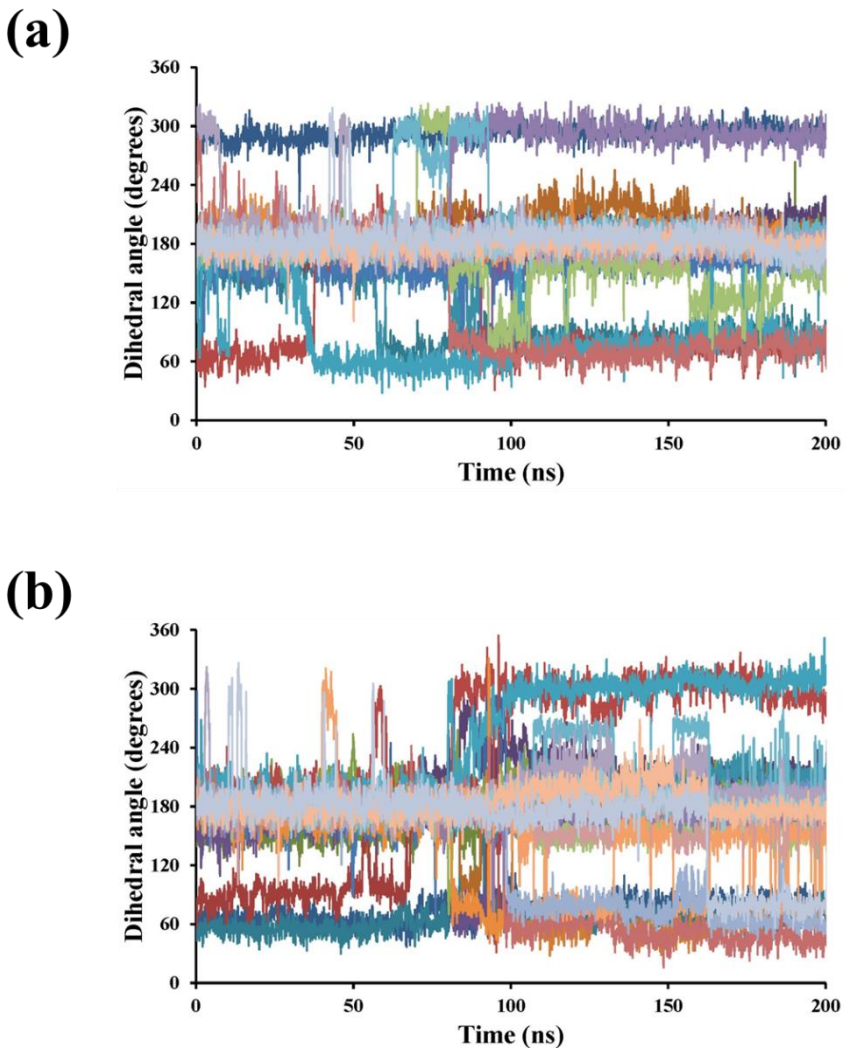


Figure 5.5.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **second** unit of the 13 used to simulate the nanotube-like.

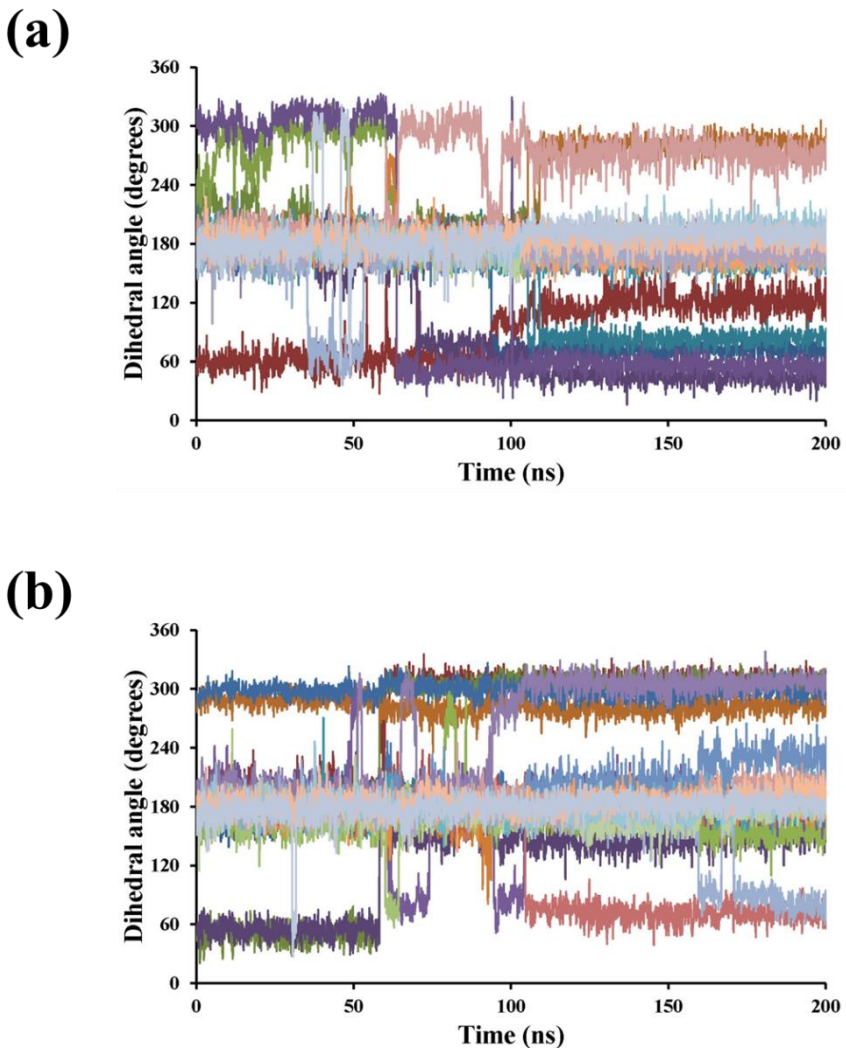


Figure 5.6.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **third** unit of the 13 used to simulate the nanotube-like.

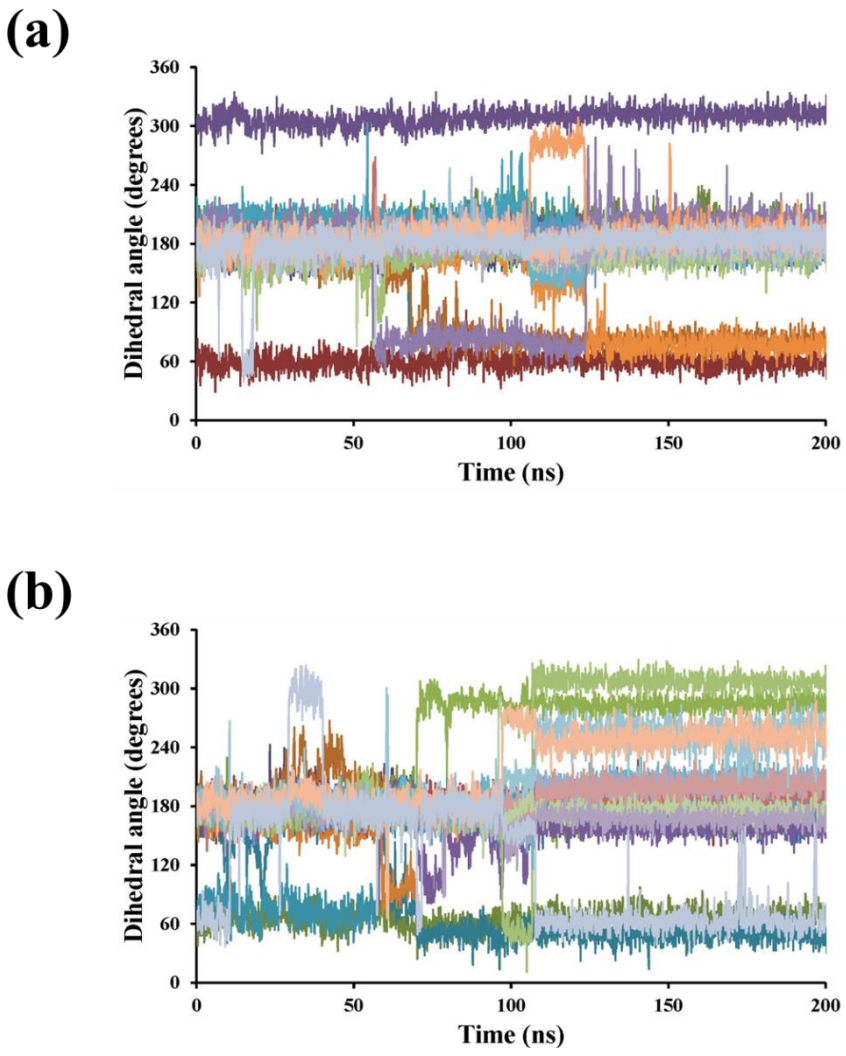


Figure 5.7.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **forth** unit of the 13 used to simulate the nanotube-like.

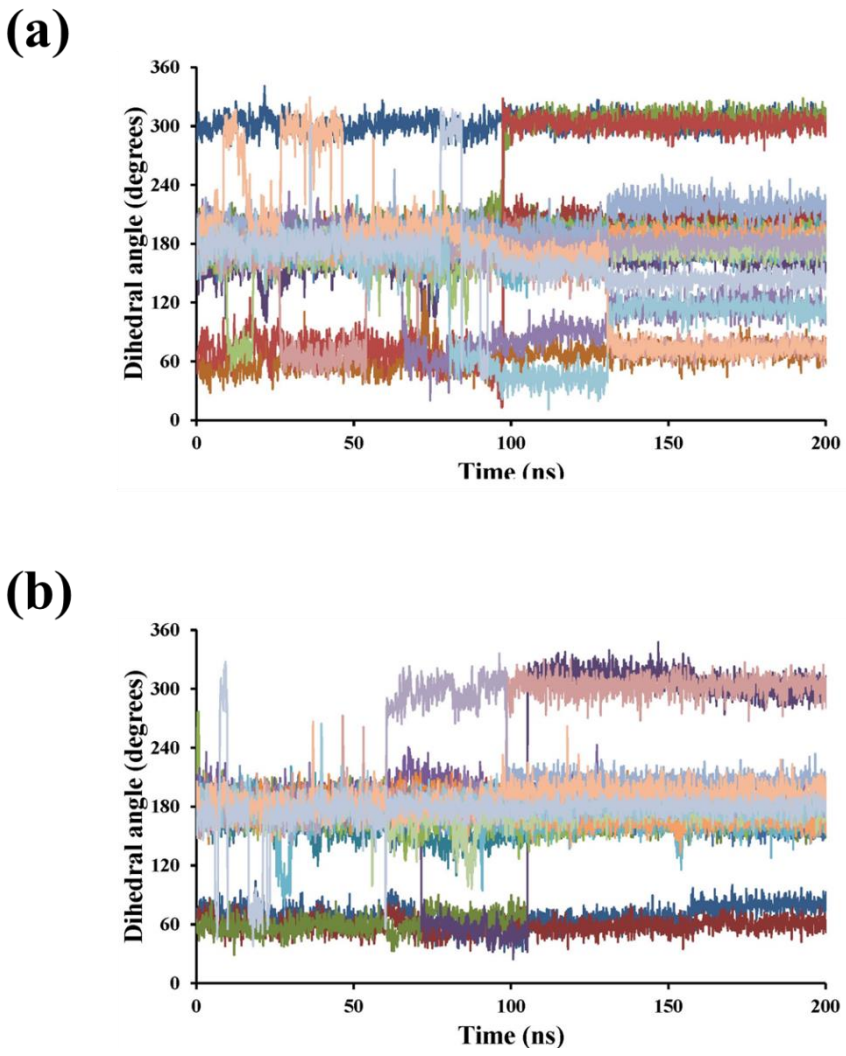


Figure 5.8.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **fifth** unit of the 13 used to simulate the nanotube-like.

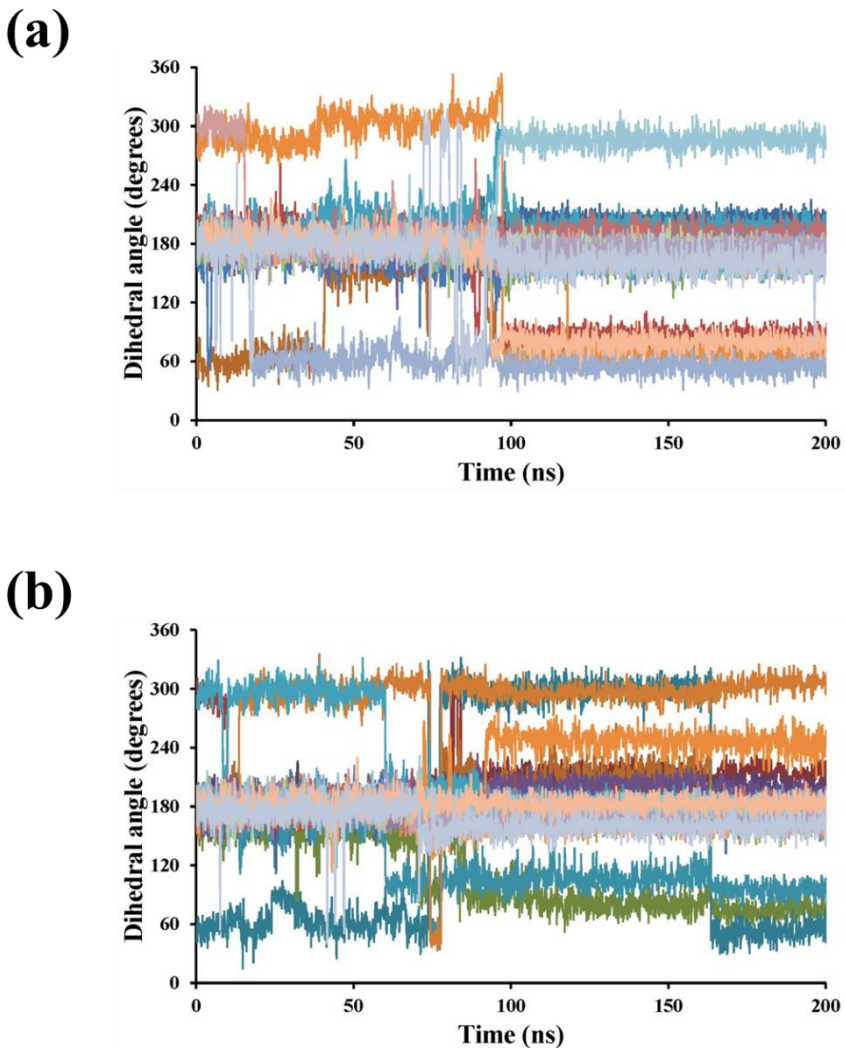


Figure 5.9.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **sixth** unit of the 13 used to simulate the nanotube-like.

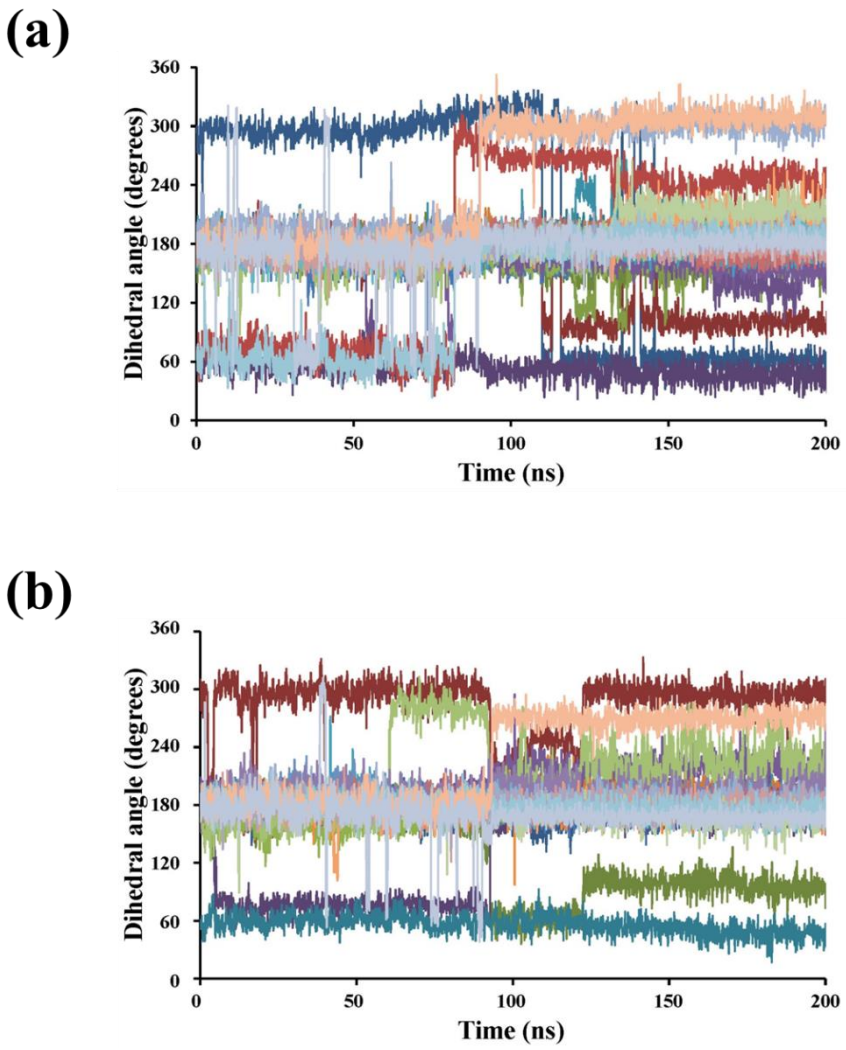


Figure 5.10.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **seventh** unit of the 13 used to simulate the nanotube-like.

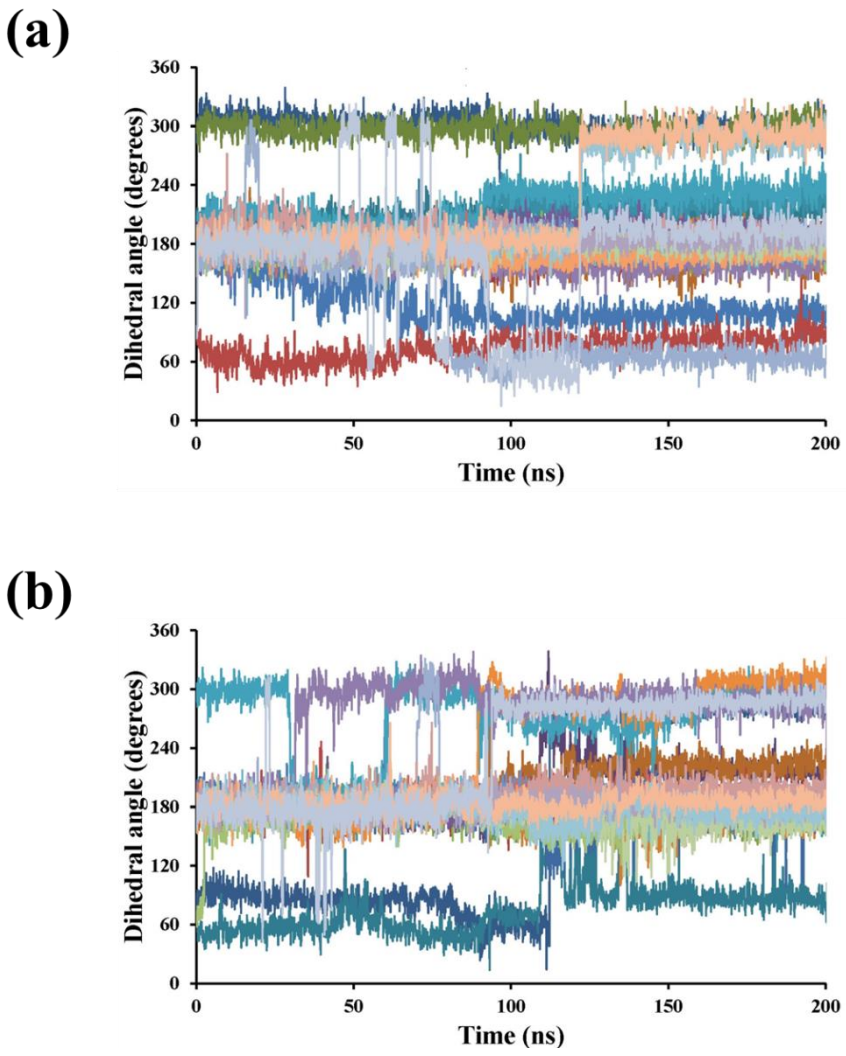


Figure 5.11.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **eighth** unit of the 13 used to simulate the nanotube-like.

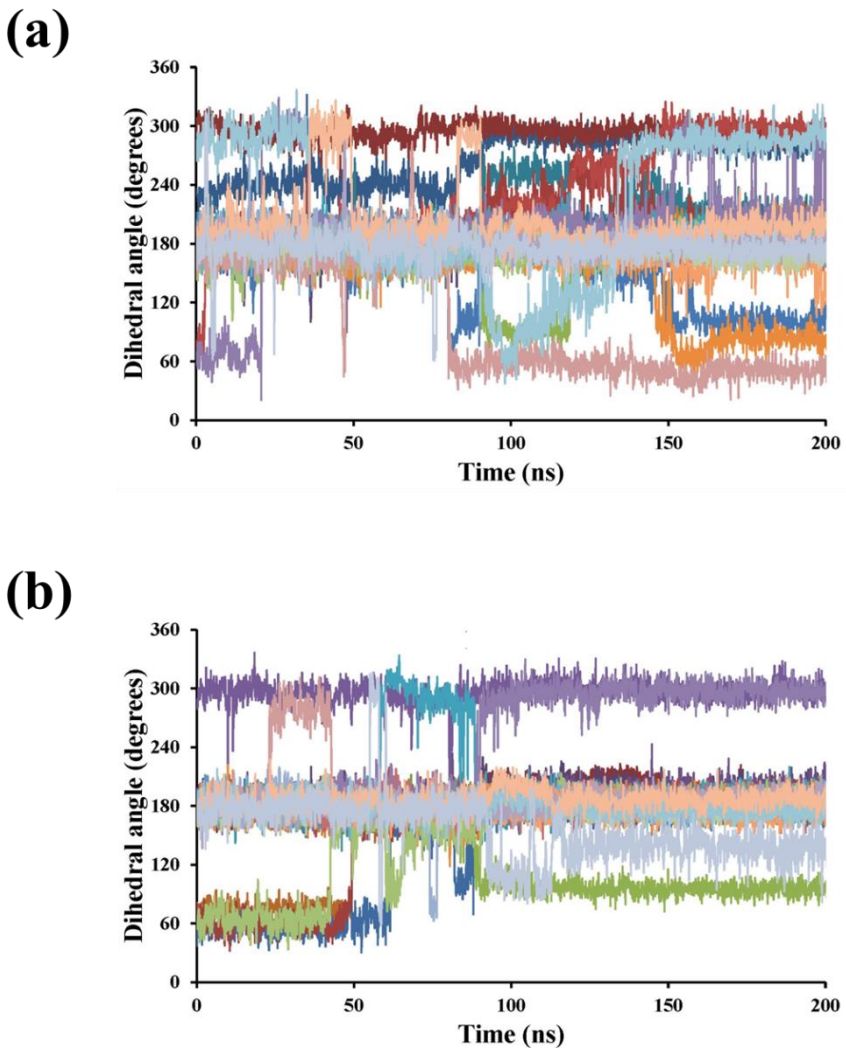


Figure 5.12.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **nineth** unit of the 13 used to simulate the nanotube-like.

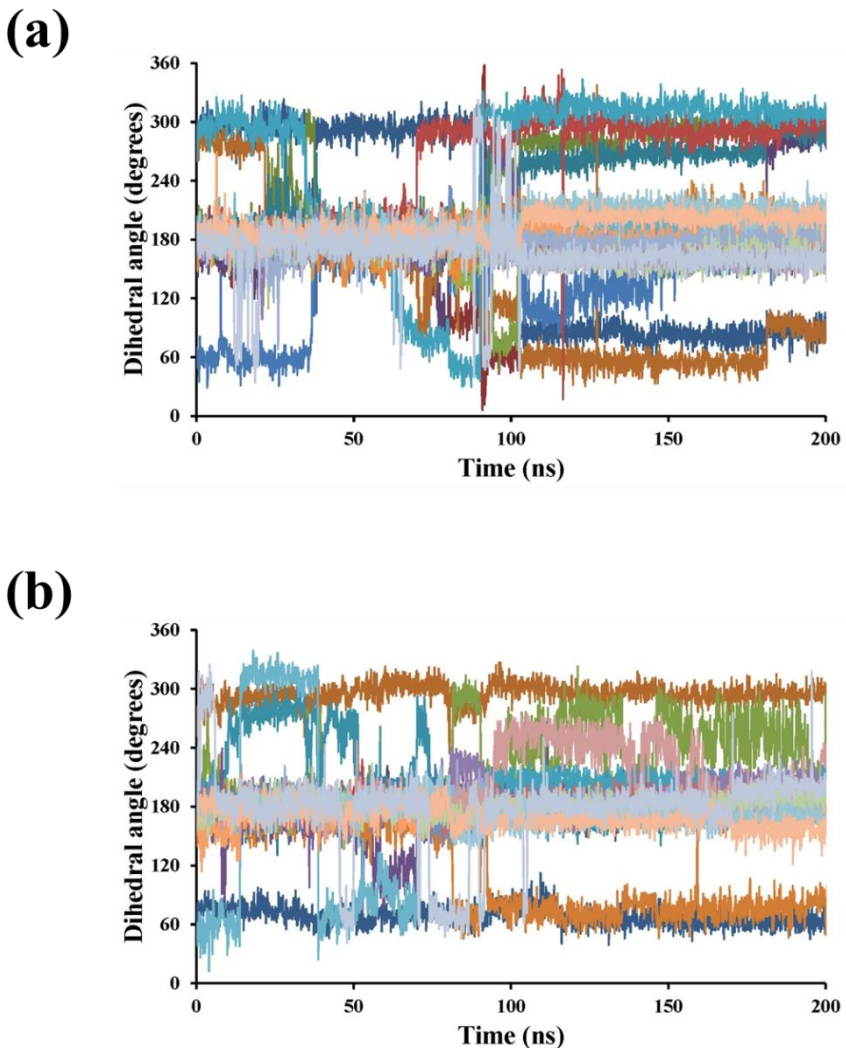


Figure 5.13.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **tenth** unit of the 13 used to simulate the nanotube-like.

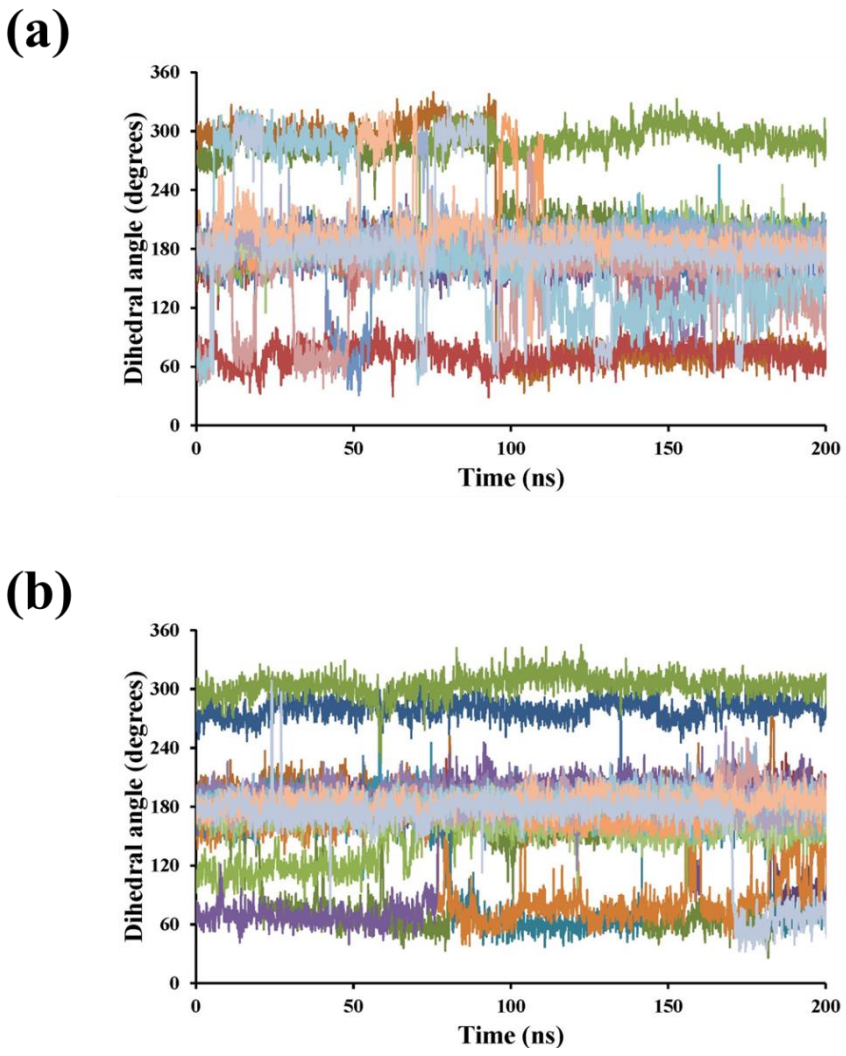


Figure 5.14.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **eleventh** unit of the 13 used to simulate the nanotube-like.

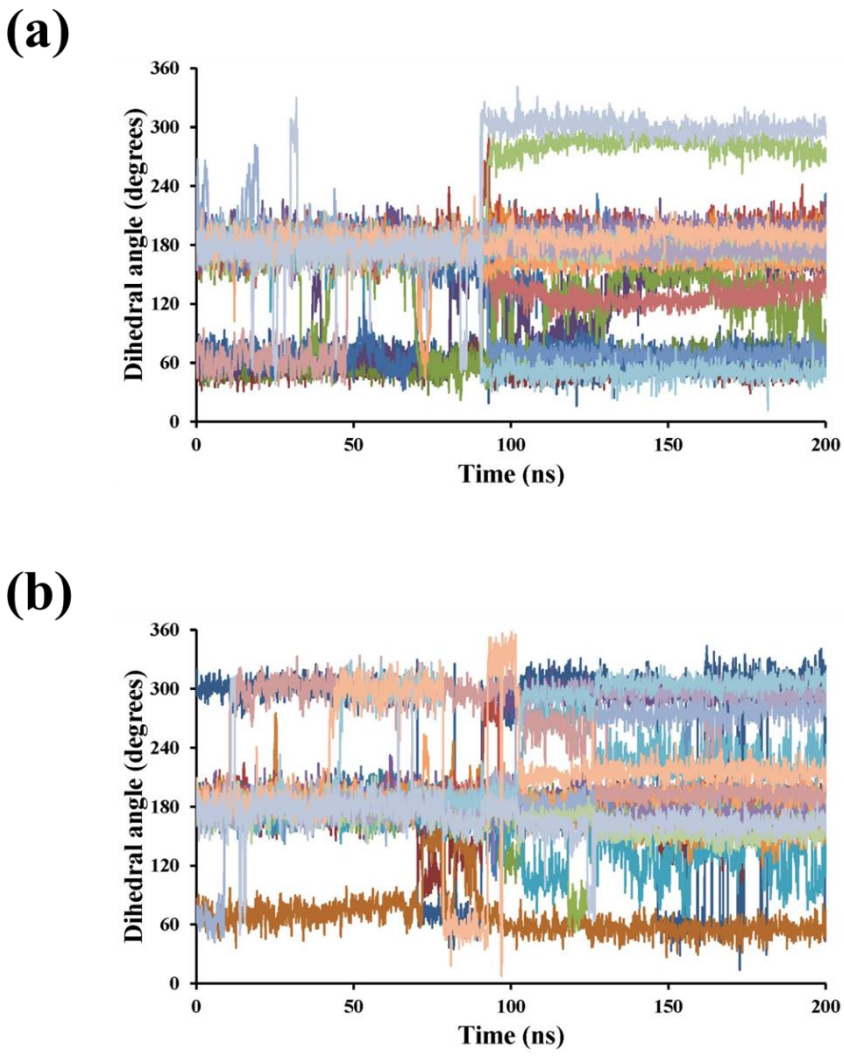


Figure 5.15.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **twelfth** unit of the 13 used to simulate the nanotube-like.

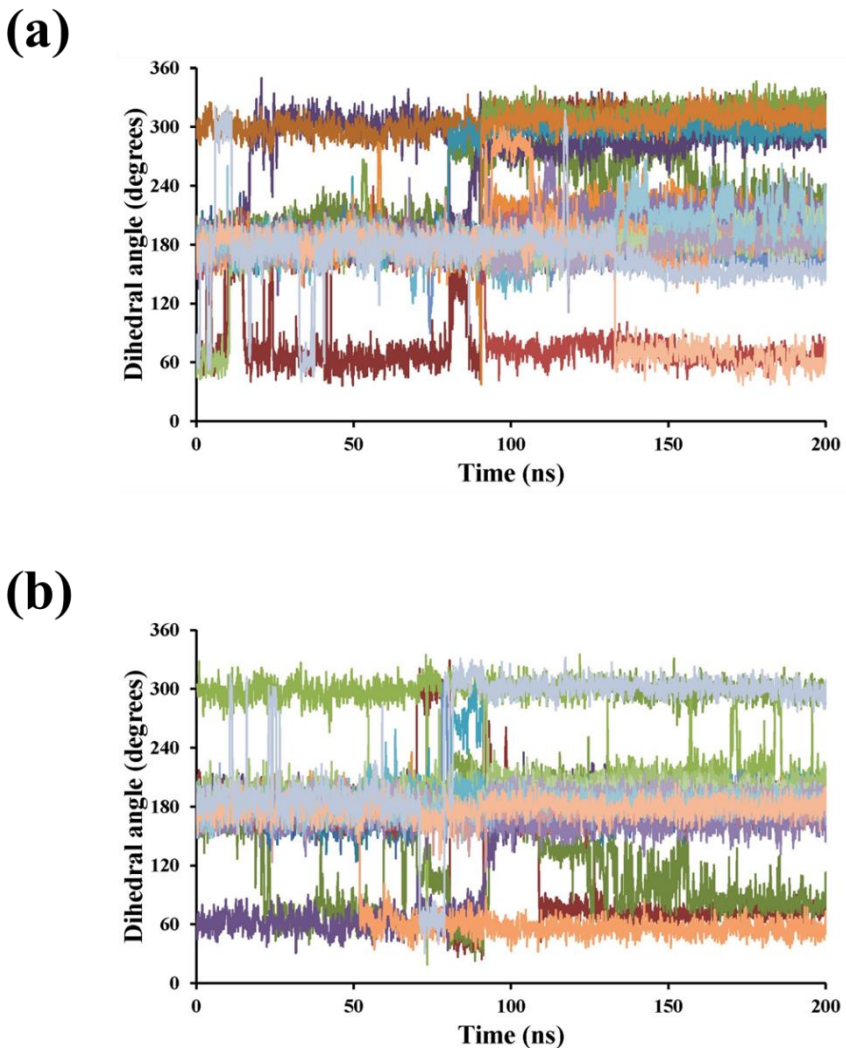


Figure 5.16.- Time evolution of each dihedral associated with the poly(n-butyl acrylate) segments conjugated to the cyclic peptide. Each association unit presents two acrylate side chains that are presented in **(a)** and **(b)** plots. This Figure corresponds to the **thirteenth** unit of the 13 used to simulate the nanotube-like.

These new results, combined with the previous work, point towards an apparent contradiction: acrylate side chains extended arrangements are only favored either when enough DMF molecules are present or when the whole system is totally dry prior to the deposition. In the latter conditions, the deposition was paired with the acrylate segments totally extended. The events here described indicate that the desolvation process itself might be the responsible to the unexplained behavior.

Another remarkable point is that while solvent molecules are majority in the solution mixture, the nanotube-like remains barely in contact with the ceramic surface. **Figure 5.3** shows the nanotube-like arranging the side chains in the same way that was described when the complex was simulated inside a solvent box, independently of the solvent kind. Despite the nanotube-like tends to get close to the surface, the acrylate side chains remain out of contact (approximately 10 to 12 Angstroms apart). Only when DMF proportion decreases to 30% or less ($\chi_{DMF} \leq 0.3$) a very small portion of the nanotube-like side chains start interacting with mica. Even under such conditions, the deposition is not completed. Once the solvent is totally removed and the solution mixture is solely constituted by inert particles, the nanotube-like rests on the mica surface, with the maximum amount of acrylate side chains in contact with it. Again, the presence of solvent molecules in different proportions appears to be the key to understand the experimentally observed conformational preferences of the acrylate segments. At high concentrations of DMF chains are extended. A slow and

progressive substitution of those by inert particles induces contraction. Despite this contraction, which does not favor the interaction with the surface, the acrylate side chains continue to be contracted until the end of the simulation. Once the nanotube is deposited on mica, the acrylate segments do not regain their initial conformation. In other words, DMF favors the maximum extension of the acrylate side chains at high concentration whereas it hampers this arrangement when its molar fraction within the solvent mixture makes it a minority component. At last, DMF molecules seem to bear the whole responsibility of hampering the acrylate chains to adopt the proper orientation required to fully interact with mica.

In order to assess the role of the solvent on the conformational preference of the acrylate side chains, several temporary dependent geometrical properties were computed. First, the DMF geometrical distribution around the acrylate segments was studied as function of the molar fraction. Also, the contacts between acrylate chains belonging to different Lys residues were followed for each studied solvent content. **Figure 5.17a** shows the partial radial distribution function of all atom pair distances between DMF mass centers and any aliphatic carbon atom belonging to the acrylate side chains. **Figure 5.17b** displays the radial distribution function of all atom pair distances between all carbon atoms present in the acrylate segments. The inset plot of **Figure 5.17b** depicts the remaining profile once the intra–residue pair distances and bond related distances have been purged.

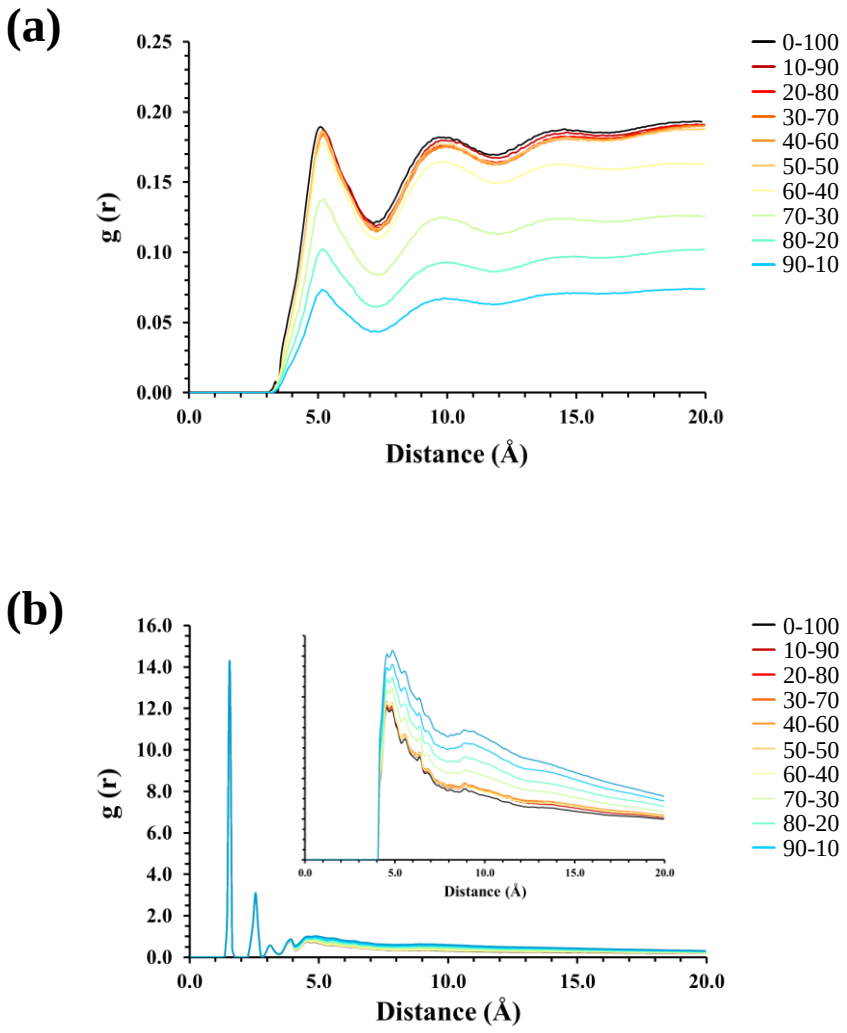


Figure 5.17-. (a) Radial distribution function of distances between DMF mass center and any aliphatic carbon atoms belonging to acrylate segments **(b)** Radial distribution function of distances between aliphatic carbon atom belonging to acrylate side chains. In set plot at the right side shows the profile obtained after purging distances corresponding to bond geometry and those of atoms belonging to the same side chain. Legend colors indicate the DMF proportion in the solution mixture.

The initial image provided by **Figure 5.17a** demonstrates that DMF molecules are systematically and coherently arranged around the acrylate segments, independently of the DMF proportion in the solvent mixture. While the solvent is present, there is a neat peak around 5.5Å, indicating the presence of a solvation shell surrounding the acrylate segments. DMF molecules favorably interact with the side chains, conditioning those segments to adopt arrangements that let a maximum surface contact with DMF molecules. Under such conditions, the conformational preferences of the acrylate segments are totally driven by the interaction with the solvent, being this effect the responsible of imposing an all *trans* arrangement to those dihedral angles with free rotation. A reflection of this intense attraction can be found on the fact that, even when the molar fraction of DMF drops to 10%, the peak at 5.5Å remains present. Thus, the same solvent molecules that favored *trans* conformers seem to hamper the adsorption of the nanotube-like onto the mica substrate once the solution mixture becomes impoverished in DMF molecules.

Previous works that explored the nanotube-like deposition in the gas phase showed that acrylate side chains did not contract in absence of solvent.[7] Hence, there must be some correlation between the depletion of solvent content and the conformational changes undergone by those segments. Detailed analysis of contacts between heavy atoms of acrylate side chains tethered to different Lys residues hints the answer to that question. The inset plot in **Figure 5.17b** shows a steady increase of the contacts between atoms of the investigated segments whilst the solvent mixture loses DMF

content. The arrangement of the acrylate side chains evolves from low contacts between atoms belonging to different nanotube-like chains towards a remarkable increase of those contacts when DMF content drops down to 50%. The loss of solvation is answered by the acrylate side chains with a severe increase in the interactions between neighboring acrylate segments.

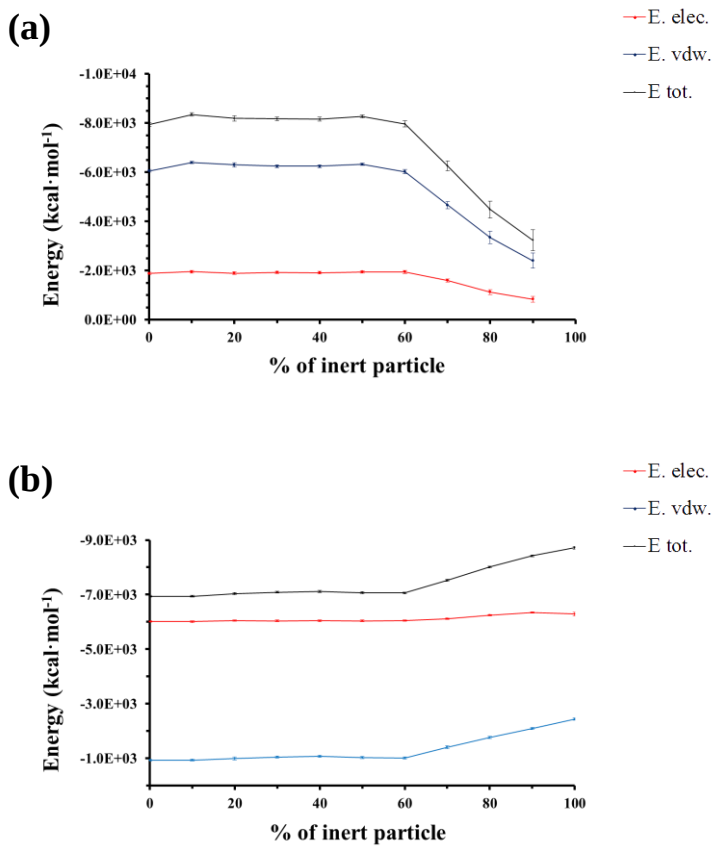


Figure 5.18.- (a) Interaction energy of the nanotube-like with DMF solvent molecules as function of the solvent mixture composition. **(b)** Interaction energy between acrylate side chains belonging to different assembly units. In both cases

total interaction energy, electrostatic component and van der Waals components are separately plotted.

In order to confirm these indications a final analysis was performed. The interaction energy between the different components of the studied system was computed and studied as function of the solvent mixture composition. **Figure 5.18a** shows the evolution of the interaction energy between the nanotube-like complex with the DMF solvent, while and **Figure 5.18b** represents the interaction energy between the acrylate side chains. The picture obtained in this analysis confirms the indications that we already hinted. While the amount of solvent molecules is large enough, the solvation of the nanotube-like is extremely favored as both electrostatic and van der Waals energy terms show. Independently of the DMF content in the solvent mixture, the interaction energy is very negative, with very low fluctuations through each simulated time period (the error bars indicate the standard deviation for each mini-cycle of 10 ns). Once the amount of DMF does not suffice the solvation of the nanotube-like, especially of the acrylate segments (i.e. when $\chi_{DMF} < 0.5$), the increment of the total potential energy of the nanotube-like must be compensated by new interactions.

Figure 5.18b confirms that the shift on the interaction preferences is directly linked to an increase in the contacts between neighboring chains that to this point were fully solvated. The intra-nanotube-like interaction energy suddenly displays a drastic improvement when the interaction energy between the acrylate segments and the DMF molecules drops. This interpretation is especially supported by the comparison of both energy components

of the intra- nanotube-like interaction energy. Independently of the composition of the solvent mixture, the electrostatic component remains steady, without showing any meaningful change for the completion of the simulation. This behavior is expected because this component is dominated by the interactions between assembly units. The main stabilizing interactions, responsible to hold the cyclic peptides bounded together, are the hydrogen bonds that provide the stability to the structural scaffold of the material. As long as the peptide assemblies does not fall apart, the energetic stability provided by them remains steady as observed. However, the van der Waals component, mainly dominated by hydrophobic interactions, abruptly improves in the moment the solvent gets impoverished in DMF. Hence, the loss of a complete solvation of the acrylate segments induces them to contract in order to compensate the destabilization produced by the desolvation by increasing the interaction with other neighboring acrylate side chains. Thus, our study shows that the acrylate segments must have retained some degree of solvation prior to the deposition, otherwise their contraction would have never taken place.

5.5 CONCLUSIONS

MD simulations performed on conjugated peptide-polymer species that spontaneously self-assemble into nanotube-like structures have aided to rationalize the conformational behavior exhibited by the polymeric blocks. The assembled unit, a molecular specie that consists of the syndiotactic cyclic peptide cyc[(L-Gln-D-Ala-L-Lys-D-Ala)$_2$] coupled with two poly(n-butyl acrylate) blocks,

were already characterized as stable nanotube structures consisting on antiparallel stacks of the cyclic peptide units stabilized by inter-strand hydrogen bonds between main chain amide polar groups.[6] Acrylate segments, though, exhibited unexpected conformational behavior, hardly depending on either being constituent part of a conjugated material or a polymeric material on its own. Poly(n-butyl acrylate) deposited on mica always tended to adopt geometrical arrangements that increased the exposed surface of its chemical groups, in order to optimize the formation of stabilizing interactions between the ceramic surface and the acrylate polymer.[8] Oppositely, when those segments were part of conjugated materials, the acrylate chains tended to be partially contracted when lying on mica.

Using a theoretical approach that simulates a selective and progressive desolvation of the nanotube-like assembly, which mimics the presumed conditions in which the nanotube-like was deposited, we have demonstrated that the solvent presence during deposition is the main responsible for the unexplained conformational preferences of the acrylate blocks. DMF molecules and the acrylate segments always show very attractive interactions, with intensively stabilizing interaction energies. This work has shown that those side chains are principally extended as long as the solvent mixture presented enough solvent molecules to cover the acrylate blocks. Our simulation experiment demonstrated as well that when the proportion of DMF molecules within the solvent drops to a critical point, determined to be between 50 and 40%, the loss of many attractive interactions DMF-acrylate segments solely

can partially be compensated by setting new stabilizing interactions between acrylate blocks tethered to neighboring assembly units. In order to achieve a new energetically stable state a significant contraction of the acrylate side chains, which is reflected at the same time by a meaningful increase in the number of torsions that are arranged in conformations different from the *trans*.

This chain of concerted events eventually affects the mode in which the nanotube-like deposits on mica. Despite the attraction between the nanotube-like material and the ceramic surface, the loss of stability by desolvation prior to completing the deposition, hampers the adoption of necessary geometrical rearrangements that the acrylate chains require to expose the highest possible interaction area. Only a sudden desolvation, prior to being deposited on mica, would facilitate the acrylate side chains to optimally interact with the ceramic surface. Our previous work had already demonstrated that when the deposition took place in complete absence of solvent the acrylate blocks were extended when laying on mica.[7] The experimental procedure to desolvate the nanotube-like was coupled with an accelerated deposition, allowing the interaction of the studied material with the solvent molecules on its way down to mica. The solvent presence while the sample was submitted to a spin–coating procedure was large enough to condition the acrylate side chains conformational preferences prior to its deposition.

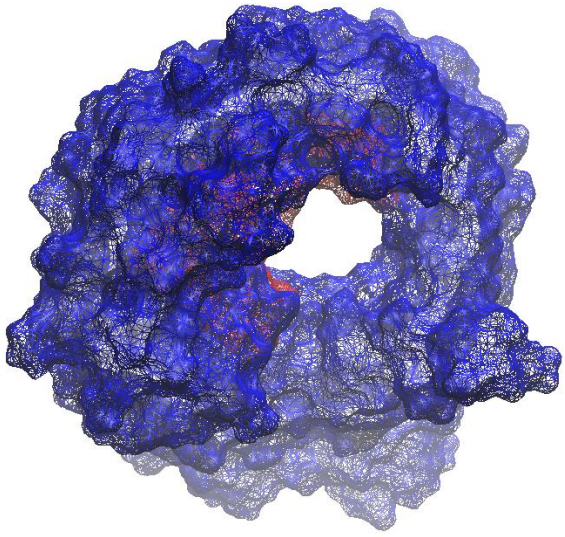
5.6 REFERENCES

1. Aida, T., Meijer, E. W. & Stupp, S. I. Functional supramolecular polymers. *Science (80-.).* **335,** 813–817 (2012).
2. Wang, A., Shi, W., Huang, J. & Yan, Y. Adaptive soft molecular self-assemblies. *Soft Matter* **12,** 337–357 (2015).
3. Webber, M. J., Appel, E. A., Meijer, E. W. & Langer, R. Supramolecular biomaterials. *Nat. Mater.* **15,** 13–26 (2015).
4. Busseron, E., Ruff, Y., Moulin, E. & Giuseppone, N. Supramolecular self-assemblies as functional nanomaterials. *Nanoscale* **5,** 7098 (2013).
5. Edwards-Gayle, C. J. C. & Hamley, I. W. Self-assembly of bioactive peptides, peptide conjugates, and peptide mimetic materials. *Org. Biomol. Chem.* **15,** 5867–5876 (2017).
6. Ten Cate, M. G. J., Severin, N. & Börner, H. G. Self-assembling peptide-polymer conjugates comprising (D-alt-L)- cyclopeptides as aggregator domains. *Macromolecules* **39,** 7831–7838 (2006).
7. Bertran, O., Curcó, D., Zanuy, D. & Alemán, C. Atomistic organization and characterization of tube-like assemblies comprising peptide-polymer conjugates: Computer simulation studies. *Faraday Discuss.* **166,** 59–82 (2013).
8. Börner, H. G., Beers, K., Matyjaszewski, K., Sheiko, S. S. & Möller, M. Synthesis of molecular brushes with block copolymer side chains using atom transfer radical polymerization. *Macromolecules* **34,** 4375–4383 (2001).
9. Heinz, H., Koerner, H., Anderson, K. L., Vaia, R. A. & Farmer, B. L. Force field for mica-type silicates and dynamics of octadecylammonium chains grafted to montmorillonite. *Chem. Mater.* **17,** 5658–5669 (2005).
10. Duan, Y. *et al.* A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **24,** 1999–2012 (2003).
11. Toukmaji, A., Sagui, C., Board, J. & Darden, T. Efficient particle-mesh Ewald based approach to fixed and induced dipolar interactions. *J. Chem. Phys.* **113,** 10913–10927 (2000).
12. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).

13. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81,** 3684–3690 (1984).
14. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101,** 4177–4189 (1994).
15. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **103,** 4613–4621 (1995).
16. Toxvaerd, S. Molecular dynamics calculation of the equation of state of alkanes. *J. Chem. Phys.* **93,** 4290–4295 (1990).
17. Wennberg, C. L., Murtola, T., Hess, B. & Lindahl, E. Lennard-Jones lattice summation in bilayer simulations has critical effects on surface tension and lipid properties. *J. Chem. Theory Comput.* **9,** 3527–3537 (2013).

- Chapter 6 -

Influence of the surrounding environment in re-naturalized β-barrel membrane proteins



6 INFLUENCE OF THE SURROUNDING ENVIRONMENT IN RE-NATURALIZED β-BARREL MEMBRANE PROTEINS[§]

6.1 ABSTRACT

Outer-membrane porins are currently being used to prepare bioinspired nanomembranes for selective ion transport by immobilizing them into polymeric matrices. However, the fabrication of these protein-integrated devices has been found to be strongly influenced by the instability of the β-barrel porin structure, which depends on surrounding environment. In this work, molecular dynamics simulations have been used to investigate the structural stability of a representative porin, OmpF, in three different environments: (i) aqueous solution at pH= 7; (ii) a solution of neutral detergent in a concentration similar to the critical micelle concentration; and (iii) the protein embedded into a neutral detergent bilayer. The results indicate that the surrounding environment not only alters the stability of the β-barrel but affects the internal loop responsible of the ions transport, as well as the tendency of the porin proteins to aggregate into trimers. The detergent bilayer preserves the structure of OmpF protein as is found bacteria membranes, while pure aqueous solution induces a strong destabilization of the protein. An intermediate situation occurs for detergent solution. Our results have been rationalized in terms of protein···water and protein···detergent interactions, which

[1] This work has been has been published as Triguero, J. *et al.* Influence of the surrounding environment in re-naturalized β-barrel membrane proteins. *Biophys. Chem.* **234,** 6–15 (2018).

makes them extremely useful for the future design of new generation of bioinspired protein-integrated devices.

6.2 INTRODUCTION

Outer-membrane porins (OMPs) of Gram-negative bacteria are transmembrane proteins allowing the bacterial cells to interact with their environment through passive diffusion across their outer membranes of water, ions, or small hydrophilic molecules (<600 Da) across their outer membranes.[1–3] OMPs are β-barrel structures usually forming homotrimeric water-filled pores. The porin channel is partially blocked by a loop (L3) folded inside the β-barrel, which consequently affects the size of the solutes that can traverse the channel. Moreover, porins exhibit ion selectivity, *e.g. Escherichia coli* OmpF and *Pseudomonas aeruginosa* OpP438 are cation- and anion-selective porins, respectively.[4,5] Many efforts have been carried out in order to understand the dynamics of these porins in their natural environment. Yet, the inner complexity of amphiphilic systems showed that atomistic details are far from being understood. Dynamic studies under non-physiological ionic strength showed that both the variable dielectric behavior in the lipid – water interface and the trimeric nature of the functional porin influences the ionization states of the acid residues constituting the L3 loop. Nonetheless structural differences were smaller than expected when D121 was tested under both potential states, whereas all comparisons were made on a crystal structure obtained at 77 K.[6]

In recent studies we have explored the very remarkable technological potential of several devices based on the coupling of

OMPs to polymeric nanomembranes.[7,8] Firstly, we functionalized a supporting matrix made of poly(N-methylpyrrole) (PNMPy) with a β-barrel OMP, Omp2a from *Brucella melitensis*, that forms channels and pores.[7] The OMP was successfully immobilized onto the PNMPy surface forming a hydrophilic, electroactive and biocompatible bio-interface promoting the passive transport of ions. More recently, we have fabricated bioinspired free-standing nanomembranes (FsNMs) for selective ion transport by immobilizing the Omp2a β-barrel protein inside nanoperforations created in flexible poly(lactic acid) (PLA) nanomembranes.[8] The functionalization of the nanoperforations caused effects similar to those observed in biological nanomembranes.[8]

The complex nature of OMPs, their poor solubility in aqueous solutions, and their instability in presence of charged surfactants severely complicated the fabrication of the above mentioned protein-integrated devices.[7,8] As a result, the preservation of the Omp2a native β-barrel structure was crucial since the structure and function of OMPs are drastically affected by environmental conditions.[7–12] The problem was partially solved using a recently reported methodology,[13] which promoted Omp2a refolding using neutral detergents, in particular *n*-dodecyl β-D-maltoside (DDM). As a major step to improve the production and efficiency of bioinspired OMP-based devices, understanding the effects induced by the environmental conditions in the structural stability and characteristic nanofeatures of these proteins is essential.

In this work we examine the impact of solvation medium in the structure of a representative OMP by using atomistic molecular

dynamics (MD) simulations. The crystal structure of Omp2a is unfortunately not yet available, and a similarity criterion has been used to select a comparable OMP. Three different situations have considered: (i) the OMP in aqueous solution at pH= 7; (ii) the OMP in aqueous solution with a number of DDM molecules corresponding to the critical micelle concentration (CMC) typically found for neutral detergents; and (iii) the OMP protein embedded in a DDM bilayer, which mimics the lipid environment typically found in cellular membranes.

6.3 MATERIALS AND METHODS

All MD simulations were carried out using the NAMD 2.9 program[14] and the Amber potential energy function.[15] The numerical parameters for the selected OMP protein (next section) were taken from Amber99 force-field.[16] Water molecules were represented using the TIP3P model,[17] while parameters for DDM compatible with the Amber force-field were taken from the literature.[18] No algorithm was used to force the bond lengths at their equilibrium distances, whereas atom pair distance cut-offs were applied at 14.0 Å to compute the van der Waals interactions. To avoid discontinuities in this energy component, the van der Waals energy term was forced to slowly converge to zero by applying a smoothing factor from a distance of 12.0 Å. Beside electrostatic interactions were extensively computed by means of Ewald summations. The real space term was defined by the van der Waals cut-off (14.0 Å), while the reciprocal space was computed by

interpolation of the effective charge into a charge mesh with a grid thickness of 10 points per volume unit.[19]

Due to its fast convergence, the weak coupling method[20] was used to heat the system and to rapidly equilibrate its pressure and temperature around 1 bar and 298 K, respectively. The relaxation times used for the coupling were 1 and 10 ps for temperature and pressure, respectively. For final relaxation and for all production runs, both temperature and pressure were controlled by the Nose–Hoover piston[21] combined with the piston fluctuation control of temperature implemented for Langevin dynamics.[22] Pressure was kept at 1.01325 bars, the oscillation period was set at 400 fs while the piston decay time was set at 100 fs. The piston temperature was set at the same value as the thermostat control, 298 K, which used a damping coefficient of 2 ps. The integration step was 2 fs in all simulations.

Before the relaxation and production steps described in the Results section, models were thermalized and equilibrated using a model-dependent four-step process. Such processes, which applied keeping the protein frozen, can be summarized as follows:

- Model for the OMP in aqueous solution at pH= 7: 1) 2×10^5 steps of energy minimization; 2) 1 ns of NVT MD at 500 K; 3) 0.4 ns of NVT MD at 298 K; and 4) 0. 5 ns of NPT MD at 298 K and 1.01325 bars.
- Model for the OMP in aqueous solution with a number of DDM molecules corresponding to the CMC: 1) 1.5×10^4 steps of energy minimization; 2) 1 ns of NVT MD at 800 K; 3) 0.6 ns of NVT MD at 298 K freezing the positions of the

detergent molecules; and 4) 0. 5 ns of NPT MD at 298 K and 1.01325 bars keeping fixed the positions of the detergent molecules.

- Model for the OMP embedded in a DDM bilayer: 1) 1×10^4 steps of energy minimization; 2) 0.6 ns of NVT MD at 800 K; 3) 1 ns of $NP_{zz}T$ (*i.e.* only the box length in the *z*-direction was allowing to such that the P_{zz} component of the pressure tensor was equal to the imposed pressure) at 500 K and 1.01325 bars freezing the positions of the detergent molecules; and 4) 0.75 ns of $NP_{zz}T$ at 298 K and 1.01325 bars enabling the movement of the detergent molecules.

In order to ensure our results reproducibility, additional production MD simulations were performed considering as starting points models obtained using the same procedures described above but introducing small changes at the equilibration step. The length of these additional production trajectories ranged from 30 to 40 ns, depending on the model. As results were very similar to those described in the Results and Discussion section, these production trajectories were not enlarged.

The *gmx sasa* program contained within the GROMAC package was used to compute the SASA[23] for all production trajectories.

6.4 RESULTS AND DISCUSSION

6.4.1 Selection, construction and equilibration of the models

As mentioned in the Introduction section, the 3D structure of Omp2a protein, which was used in our previous experiments,[7,8,13]

remains unknown. Consequently, the first step of the modeling study was to select an alternative OMP in the Protein Data Bank (PBD), which collects the experimentally-determined structures of proteins. Table 6.1 summarizes important characteristics of OMPs from Gram-negative bacteria that present greater similarity with Omp2a. From this list, OmpF with PDB code of 3K19[24] has been selected as the most appropriate due to its likeness with Omp2a in terms of function, structural and chemical constitution, and membrane location. The OmpF transmembrane protein, which is active as a trimer for ion transport and contains a 16-stranded β-barrel, share with Omp2a the ability to refold by addition of ~1.5 mM DDM[13,25]. Furthermore, Omp2a is constituted by 367 amino acids, which is close to the 340 amino acids of OmpF. Although the ion permeability and ion selectivity through the OmpF channel have been investigated using MD simulations,[26–29] the impact of the environment on its structural stability has not been studied yet.

Once the OmpF was selected as representative OMP, the following three models were constructed:

(i) The OmpF immersed in a box with 56711 explicit water molecules, 116 Na^+ and 104 Cl^- explicit ions (*i.e.* 100 mM NaCl concentration), which represents an aqueous solution at pH= 7.

(ii) The OmpF in a solution made of 50660 explicit water molecules, 116 Na^+ and 104 Cl^- explicit ions, plus 62 DDM molecules.

(iii) The OmpF embedded in a bilayer formed by 884 DDM molecules, which in turn is immersed in a bath of 45940 explicit water molecules.

Models described in (i), (ii) and (ii), are hereafter denoted $(OmpF)_{aq}$, $(OmpF)_{DDM}$ and $(OmpF)_{bl}$, respectively, and involve 170353, 157222 and 209644 explicit particles. The $(OmpF)_{aq}$ and $(OmpF)_{DDM}$ models were constructed by placing one protein molecule in its crystal structure, 3K19[24], at the center of a $120\times120\times120$ $Å^3$ simulation box fulfilled with previously equilibrated (1 atm and 298 K) explicit water molecules. The charge of all potential titratable residues was fixed to values corresponding to neutral pH (*i.e.* all Asp and Glu side chains were negatively charged, while Lys and Arg side groups were represented in their positively charged forms). Then, the ions required to neutralize charged protein side chains, as well as to reach the desired anionic strength (*i.e.* 116 Na^+ and 104 Cl^-) were added to each simulation box. For $(OmpF)_{DDM}$, 62 DDM molecules were randomly added, avoiding any overlap with the protein atoms. Finally, water molecules in steric conflict were deleted. The DDM bilayer included in the $(OmpF)_{bl}$ model was prepared using the PACKMOL program.[30] The rest of the construction process was identical to that described for the other models, except the simulation box dimensions set to $125\times125\times150$ $Å^3$.

	Omp28	Maltoporin	OmpG	PagP	Omp32	OmpA	OmpF	Omp2a
PDB code	2R2C	1MAL	2IWW	1THQ	2FGR	2KOL	3K19	-
Number of amino acids	183	421	281	170	332	216	340	367
Number of β-strands	10	18	14	10	16	8	16	16
Function	Receptor	Porin	Porin	Enzyme	Porin	Structural	Porin	Porin
Functional form	Monomer	Trimer	Monomer	Monomer	Trimer	Dime	Trimer	Trimer

Table 6.1. Major characteristics of the OMPs found in the PDB for comparison with non-crystallized Omp2a.

6.4.2 Relaxed models

MD simulations on $(OmpF)_{aq}$, $(OmpF)_{DDM}$ and $(OmpF)_{bl}$ were performed using the Amber99 force-field.[16] After thermal and structural equilibration using the careful protocols described in the Methods section, the structure of each system was relaxed by applying 5.5 ns of NPT MD. After this short MD simulation, the root mean square deviation (RMSD) of the atomic positions with respect to the crystal structure was obtained for the three models. The RMSD, which was obtained using the C^{α} atoms, was 3.58, 1.82 and 1.26 Å for $(OmpF)_{aq}$, $(OmpF)_{DDM}$ and $(OmpF)_{bl}$, respectively, reflecting that the environment can significantly affect the main features of the crystallized β-barrel architecture, even during the relaxation step.

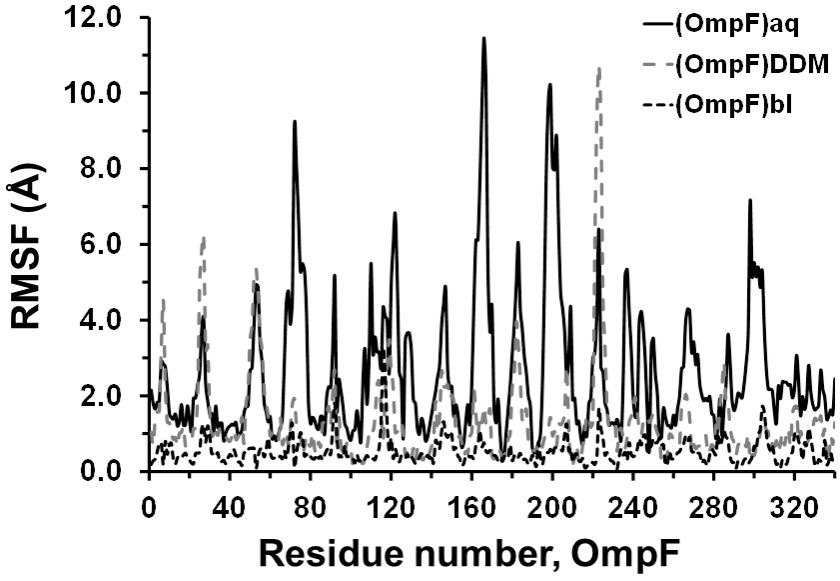


Figure 6.1. Residue-based C^{α}-RMSFs for OmpF after relaxation in several environments compared to the crystal structure.

To assess the location of the major distortions along the backbone of the constituting residues, the root mean square fluctuations (RMSF) of the individual C^{α} atoms from the protein relaxed in the different environments was measured with respect to their positions in the crystal (Figure 6.1). As it can be seen, the largest fluctuations occur at the same positions in all cases, albeit $(OmpF)_{aq}$ RMSFs are significantly higher than their $(OmpF)_{DDM}$ and, especially, $(OmpF)_{bl}$ counterparts. In general, the C^{α}-RMSF in the $(OmpF)_{DDM}$ and $(OmpF)_{bl}$ structures display low fluctuation for all residues, with a maximum of 3 Å in the most distorted regions, which are mainly located at loops. Meanwhile, $(OmpF)_{aq}$ exhibits several regions with C^{α}-RMSFs > 4 Å, with a maximum of 11.4 Å at the center of the helical tract comprising residues 159-171, evidencing important local distortions. These results are fully consistent with the previously presented RMSD values. The next subsection provides details about the structural evolution of the different regions of the protein during the MD production runs, for which the relaxed structures of the three models were used as starting points of 120 ns trajectories for $(OmpF)_{aq}$ and $(OmpF)_{DDM}$ models and of 100 ns for $(OmpF)_{bl}$. The first 10 ns of trajectory was omitted for statistical analyses.

6.4.3 OmpF in aqueous solution at pH= 7: $(OmpF)_{aq}$ model

Figure 6.2a displays the time evolution of the C^{α}-RMSD in the $(OmpF)_{aq}$ model, calculated with respect to the crystal structure. The RMSD increases from 3.58 to 3.86 Å after a 120 ns trajectory, the mean value averaged over the whole production run being 3.75±0.23 Å. Whilst this relatively small variation of C^{α}-RMSD (< 0.3 Å) would suggest that the protein does not undergo important distortions along the production run, comparison between the C^{α}-RMSFs at the beginning and at the end of the production run reflects the opposite. More specifically, the protein experiences a significant re-structuration with respect to the initial structure during the production trajectory. As illustrated by Figure 6.3a, the first part of the protein undergoes a distortion with respect to the relaxed structure. Thus, the level of alteration increases in the regions that were already deformed during the relaxation process. In contrast, the second half of the protein stabilizes with respect to the crystal and the degree of distortion decreases for many secondary motifs. The combination of such distinctive behaviors explains the relatively small increment in C^{α}-RMSD during the production trajectory (Figure 6.2a).

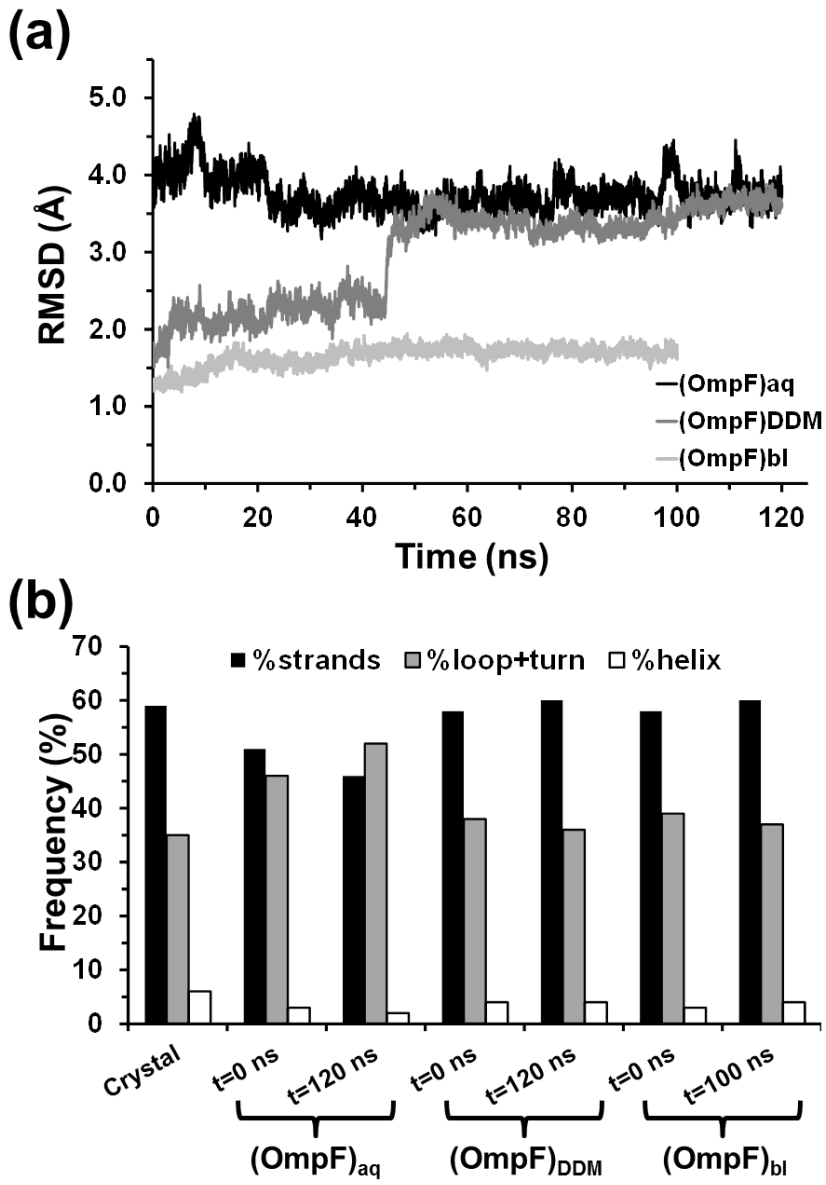


Figure 6.2. (a) Evolution during the production trajectory of the C^{α}-RMSD of the three simulated models compared to the crystal structure. (b) Frequency of the secondary structure motifs for the OmpF protein in the crystal structure in the snapshot used as starting point of the production run (relaxed model corresponding to t= 0 ns) and in the last snapshot of the production run (t= 100 or 120 ns) for the three simulated models.

By using the DSSP program,[31] the secondary structure of the protein has been characterized in terms of β-strand (βS), loops and turns (L+T), and helices (H). The results found for the first and last snapshot of the production simulation are compared with the crystal in Figure 6.2b. The progressive reduction in weight of βSs and Hs (*i.e.* 13% and 4%, respectively, after 120 ns) and the corresponding increment of L+Ts (17%), which becomes the most populated motif with respect to the crystal structure, is fully consistent with the destabilization of the first half of the protein identified by C^{α}-RMSF. However, a detailed inspection of the snapshots indicates that the extension of such structural destabilization is greater than that indicated in the secondary structure analysis. Indeed, some loops of the crystallized proteins transform into regular secondary motives during the production trajectory. This is illustrated by Figure 6.3b where the protein structures at the beginning and at the end of the production MD are displayed along the crystal structure. For example, the region comprised between 9-12 residues forms a disorganized loop in the crystal, but adopts a well-defined helical conformation (light green dashed circle in Figure 6.3b) that re-orients and grows during the simulation. Another source of destabilization corresponds to the loops, which exhibit spatial orientation completely different from that of the crystal, at both the beginning and end of the simulation. This is also illustrated by few punctual results depicted in Figure 6.3b (light blue dashed rectangles).

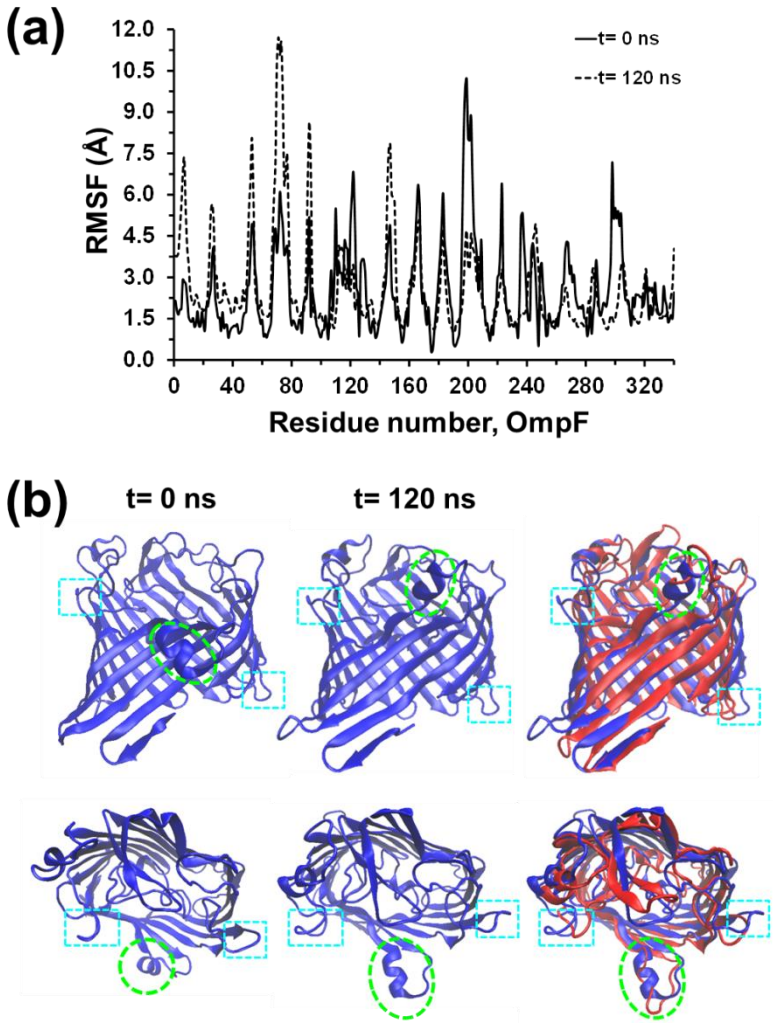


Figure 6.3. For the $(OmpF)_{aq}$ production trajectory: (a) Residue-based C^{α}-RMSFs for OmpF protein at the beginning (relaxed structure at *t*= 0 ns) and at the end (*t*= 120 ns) of the simulation; and (b) Axial and equatorial (top and bottom, respectively) views of the protein at the beginning (left) and at the end (middle) of the simulation. The superposition of the protein structures at the end of the simulation (blue) and in the crystal (red) are displayed in the right column. The light green dashed circle shows that residues 72-78, which form a loop in the crystal, adopt a helical conformation in the model. The light blue dashed rectangles display some loops whose spatial orientations in the model differ from those in the crystal.

6.4.4 OmpF in a detergent solution: $(OmpF)_{DDM}$

The C^{α}-RMSD for the $(OmpF)_{DDM}$ simulation increases slowly and gradually from 1.82 to 2.30 Å during 44 ns (Figure 6.2a). After this, it grows to 3.18 Å in ~1 ns, and at the end of the 120 ns trajectory, the C^{α}-RMSD is 3.63 Å, with an average value of 3.44±0.16 Å in the 45-to-120 ns time interval. This C^{α}-RMSD trends indicate that, during the first part of the production trajectory (*i.e.* 44 ns), the protein remains in a conformation similar to that in the crystal, before undergoing a very fast destabilization. Figure 6.4a compares the C^{α}-RMSFs at the beginning and the end of the production run. It shows that the zones delimited by residues 68-76 on the one hand and 142-152 on the other hand almost exclusively concentrate the observed destabilization, while the rest of protein exhibits moderate or even low distortions with respect to the relaxed structure used as starting point.

The comparison of populations associated to βS, L+T and H secondary motifs highlights the remarkable similarity between the simulated and the crystal structures (Figure 6.2b). Indeed, in this case the most destabilized regions fundamentally correspond to poorly organized L+T patterns. In clear, L+T preserve their disordered state, though their structure differs from that found in the crystal. These features are illustrated in Figure 6.4b, which reflects the large discrepancies between the loops in the crystal and those formed at the end of the production run (*e.g.* the light blue and the green circles correspond to the loops contained in the 68-76 and 142-152 segments, respectively). Essentially, these results reveal that detergent molecules do not prevent changes in the spatial

orientation of the intrinsically disordered loops, whereas they clearly exert some protective effect in βS and H regular motifs, as previously proven in the parent porin Omp2a.[13]

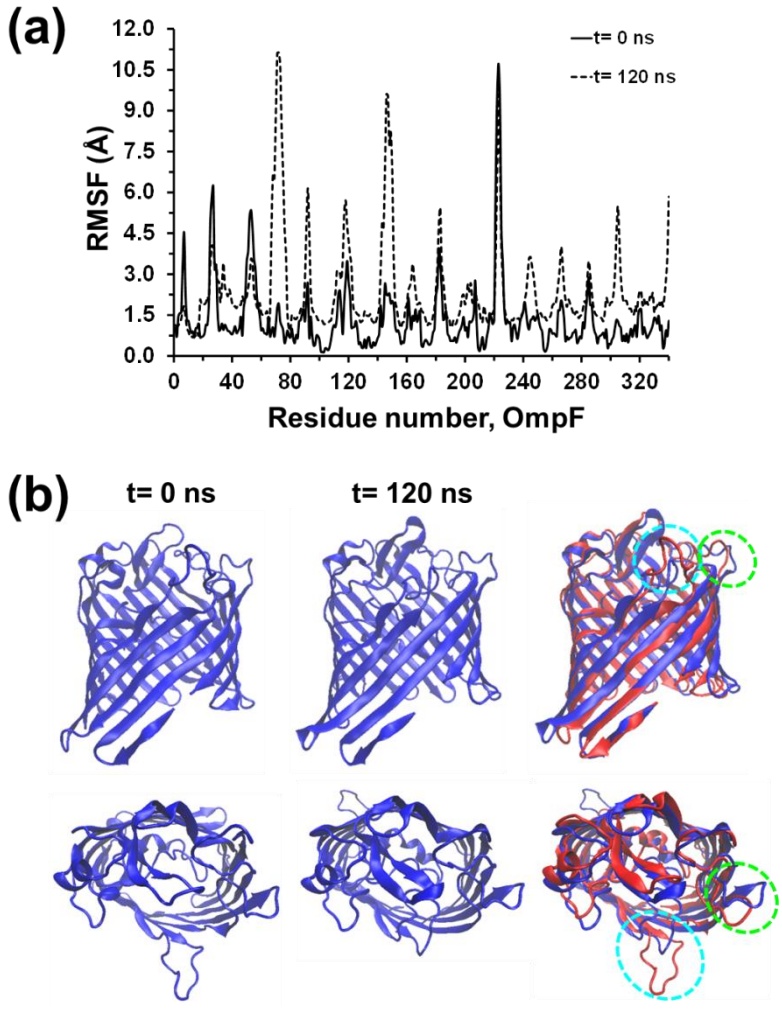


Figure 6.4. For the $(OmpF)_{DDM}$ production trajectory: (a) Residue-based C^{α}-RMSFs for OmpF protein at the beginning (relaxed structure at *t*= 0 ns) and at the end (*t*= 120 ns) of the simulation; and (b) Axial and equatorial (top and bottom, respectively) views of the protein at the beginning (left) and at the end (middle) of the simulation. The superposition of the protein structures at the end of the

simulation (blue) and in the crystal (red) are displayed in the right hand column. Light blue and the green circles, which correspond to the loops contained in the 68-76 and 142-152 segments, respectively, reflect the discrepancy between the simulated and crystal structures.

6.4.5 OmpF embedded in a bilayer: $(OmpF)_{bl}$

The average C^{α}-RMSD in $(OmpF)_{bl}$ is 1.65±0.13 Å after 100 ns (Figure 6.2a), which represents an increment of less than 0.4 Å with respect to the relaxed structure. This suggests that the detergent bilayer protects the protein, thus preserving the structure observed in the crystal. This stability is confirmed by Figure 6.5a, which plots the C^{α}-RMSF for the initial structure and the last snapshots with respect to the crystal structure. Only six residues (92, 116-118, 223 and 304), which are located at turn or bend-regions in the crystal structure, present a RMSF higher than 3.5 Å. Indeed, differences between the simulated and crystal structures only involve a few L+T motifs (Figure 6.5b). Accordingly, the bilayer provides global stability not only to the βS and H motifs but also to the structurally disordered L+T (Figure 6.2b).

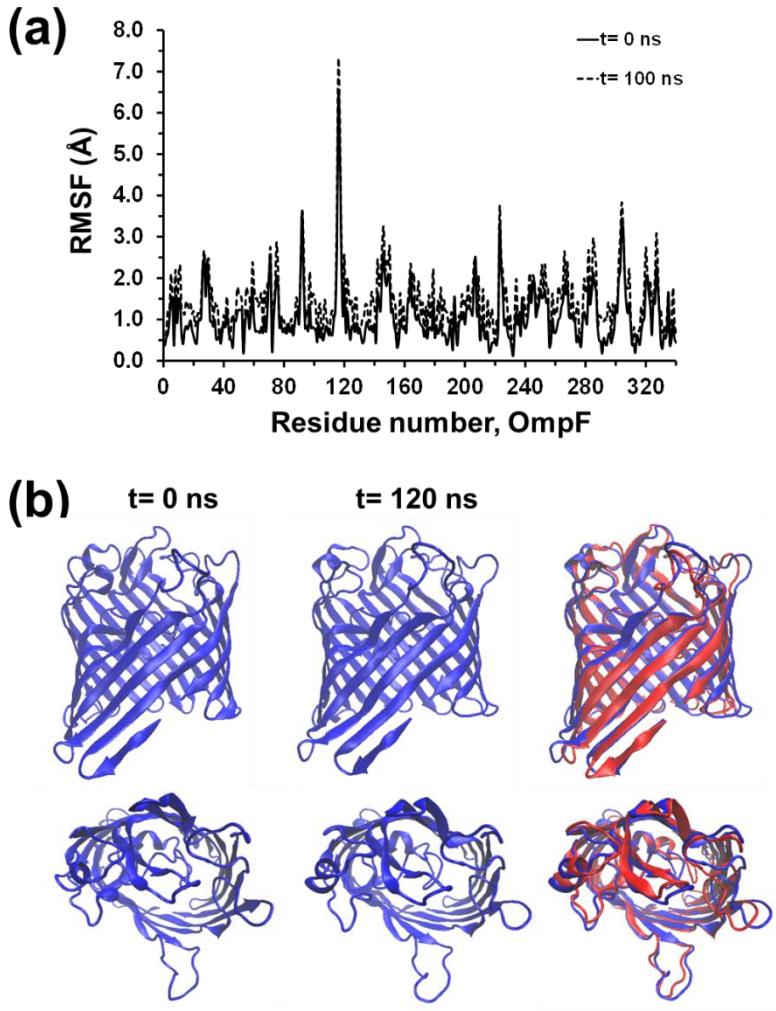


Figure 6.5. For the $(OmpF)_{bl}$ production trajectory: (a) Residue-based C^{α}-RMSFs for OmpF protein at the beginning (relaxed structure at *t*= 0 ns) and end (*t*= 120 ns) of the simulation; and (b) Axial and equatorial (top and bottom, respectively) views of the protein at the beginning (left) and at the end (middle) of the simulation. The superposition of the protein structures at the end of the simulation (blue) and in the crystal (red) are displayed in the right hand column.

6.4.6 Analysis of the protein···water and protein···detergent interactions

Figure 6.6a compares the evolution with time of the solvent accessible surface areas (SASA) of the protein for $(OmpF)_{aq}$, $(OmpF)_{DDM}$ and $(OmpF)_{bl}$ models. In aqueous solution the SASA rapidly stabilizes around an average value of 182±3 nm^2. This behavior, which is fully consistent with the RMSD profile discussed above (Figure 6.2a), indicates a 10% increment with respect to the SASA of the crystal structure (164 nm^2). Consequently, one can draw the hypothesis that the distortions reflected by both the C^{α}-RMSD and C^{α}-RMSF profiles (Figures 6.2a and 6.3a) are related to the attractive protein···water interactions, resulting in a higher exposition of the residues.

The average SASA value obtained from the $(OmpF)_{DDM}$ trajectory is 178±2 nm^2 for the first 44 ns, then decreasing to 174±2 nm^2 for the interval between 44 and 120 ns. Comparison of these variations with the results observed for $(OmpF)_{aq}$ indicates that during the first 44 ns the protein is more protected in the detergent solution than in pure water. After such period of time, the SASA reduction to reach a value closer to the crystal structure might apparently look anomalous since simultaneously both the C^{α}-RMSD and C^{α}-RMSF increase. However, complementary analyses in which the SASA of protein together with DDM molecules, as well as of the DMM molecules alone (OmpF+DDM and DDM, respectively, in Figure 6.6b) highlight such unexpected response. Indeed, the exposition to water of both OmpF+DMM and DDM decreases after 44 ns, indicating that the structure of the protein

changes to achieve the maximum protein···detergent contact. Thus, detergent molecules induce localized distortions at the L+T regions maximizing the hydrophobic contacts, whereas βS and H motifs, which are stabilized by intramolecular protein···protein interactions, preserve their stable structures.

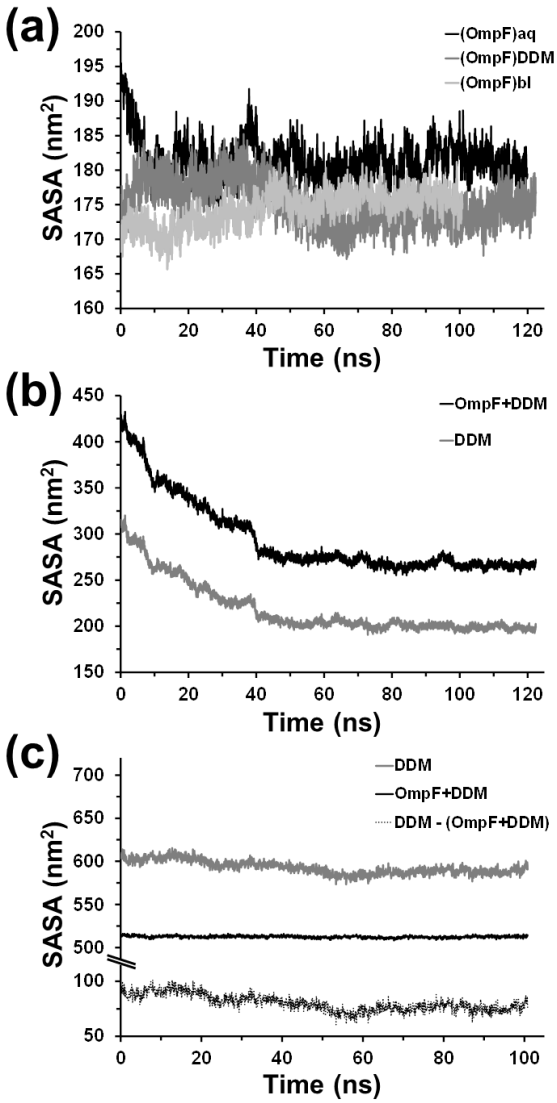


Figure 6.6. (a) Evolution during the production trajectories of the SASA for the three OmpF examined models. For the (b) $(OmpF)_{DDM}$ and (c) $(OmpF)_{bl}$ models, the SASA evolution with time of the protein together with DDM molecules and of the DMM molecules alone (OmpF+DDM and DDM, respectively) are also displayed. In (c) the protein···detergent contact surface has been estimated as the difference between the OmpF+DDM SASA and the DDM SASA.

Finally, the SASA of the protein calculated using the $(OmpF)_{bl}$ trajectory fluctuates around its average value, 172±2 nm^2 (Figure 6.6a), which deviates by less than 5% from the value obtained for the crystallographic structure. This behavior is fully consistent with the small C^{α}-RMSD (Figure 6.2a) and C^{α}-RMSF (Figure 6.5a) found for the protein inside the detergent bilayer. Figure 6.6c, which represents the evolution of the SASA values for OmpF+DDM, reflects very small fluctuations around the average value, 513±1 nm^2, evidencing the structural stability of the bilayer. Moreover, the protein···detergent surface contact, which was estimated as the difference between the SASA values of OmpF+DDM and DDM alone (*i.e.* removing the protein) for each snapshot (Figure 6.6c), exhibits fluctuations similar to those observed for the C^{α}-RMSD.

The interactions between the OmpF protein and water molecules have been analyzed by calculating the pair distribution function between non-hydrogen OmpF atoms and water oxygen atom, $g_{OmpF\text{-}wat}$ (Figure 6.7a). For all models, three peaks have been detected corresponding to directly interacting ($R \approx$ 2.8 Å) and ordered ($R \approx$ 3.8 and 4.9 Å) water molecules. The first peak is mainly associated to hydrogen bonded water molecules, while the second and third peaks correspond to the second and third hydration shells (*i.e.* solvent molecules hydrogen bonded to the water of first and solvation shell, respectively). On the top of that, the number of water molecules interacting with the protein decreases as follows: $(OmpF)_{aq}$ > $(OmpF)_{DDM}$ >> $(OmpF)_{bl}$. Obviously, this can be attributed to the fact that in the $(OmpF)_{aq}$ model the protein is completely immersed

in aqueous solution, while in the $(OmpF)_{DDM}$ model some detergent molecules tend to be located in positions close to the protein (see below). Finally, in the $(OmpF)_{bl}$ model, the OmpF is embedded in a detergent bilayer and the only protein regions accessible to water molecules correspond to the edges of the β-barrel.

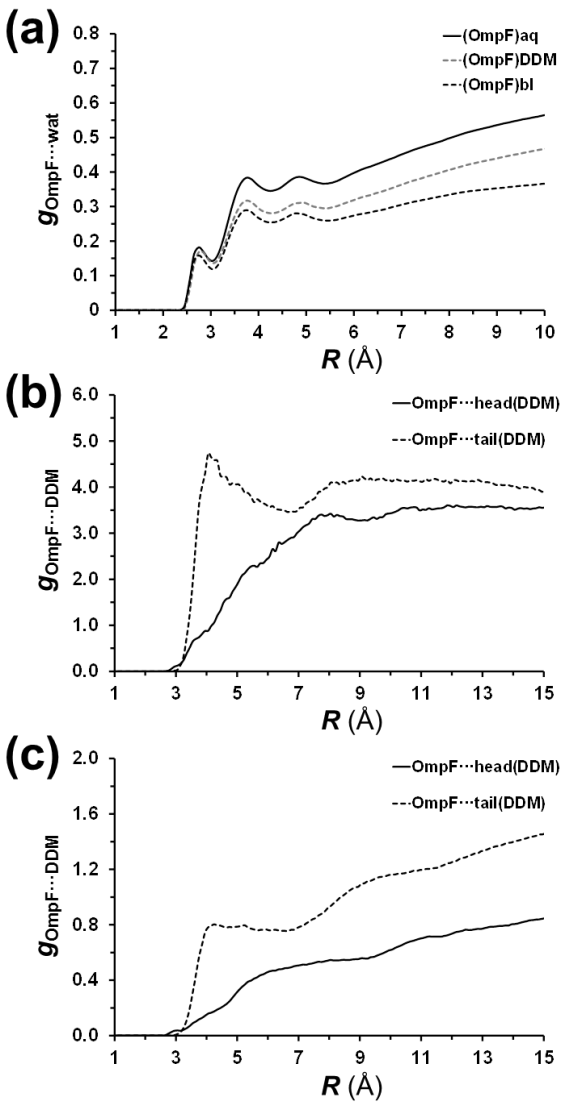


Figure 6.7. Distribution functions for the following pairs of atoms: (a) non-hydrogen atoms from OmpF and oxygen atoms from water for $(OmpF)_{aq}$, $(OmpF)_{DDM}$ and $(OmpF)_{bl}$ models; (b) non-hydrogen atoms from OmpF and non-hydrogen atoms of maltosyl (OmpF···head) or last atom of the dodecyl tail (OmpF···tail) for the $(OmpF)_{DDM}$ model; and (c) non-hydrogen atoms from OmpF and non-hydrogen atoms of maltosyl (OmpF···head) or last atom of the dodecyl tail (OmpF···tail) for the $(OmpF)_{bl}$ model.

On the other hand, OmpF···detergent interactions have been examined considering the head and the tail of the DDM molecules. More specifically, the pair distribution functions between non-hydrogen OmpF atoms and the oxygen atom connecting the two rings of the maltosyl group ($g_{\text{OmpF-head}}$) or the last carbon atom of the dodecyl tail ($g_{\text{OmpF-tail}}$) are represented in Figures 6.7b and 6.7c for $(\text{OmpF})_{\text{DDM}}$ and $(\text{OmpF})_{\text{bl}}$ models, respectively. The lack of peaks in $g_{\text{OmpF-head}}$ profiles reflects the absence of preferential interactions between the protein and the maltosyl group, which is consistent with the fact that heads of the DDM molecules are mainly oriented towards the aqueous environment. On the contrary the $g_{\text{OmpF-tail}}$ profiles display a peak centered at 4.05 Å both for $(\text{OmpF})_{\text{DDM}}$ and $(\text{OmpF})_{\text{bl}}$. That is consistent with the formation of stabilizing lipophilic interactions between the tails of the DDM molecules and the protein. The formation of such attractive interactions is fully coherent with the SASA results discussed above (Figure 6.6), and explains the stabilizing effects experimentally observed when membrane proteins are combined with detergents.[13]

6.4.7 Analysis of the L3 and L2 loops

OmpF forms aqueous voltage-gated channels that span the outer membrane, allowing the diffusion of small polar molecules. Its crystal structure reveals that eight long irregular loops named L1-L8 connect the antiparallel β-strands.[24] L3 is particularly long (33 residues) and folds inside the protein channel defined by the β-barrel, thus restricting its accessibility. This was reported to

constrain the pore dimension from a theoretical 23 Å to 7×11 $Å^2$, [24] allowing the passage of molecules < 600 Da, with slight preferences for those bearing a positive charge. In short, L3 controls the pore diameter, modulating the solute transport (Figure 6.8a).

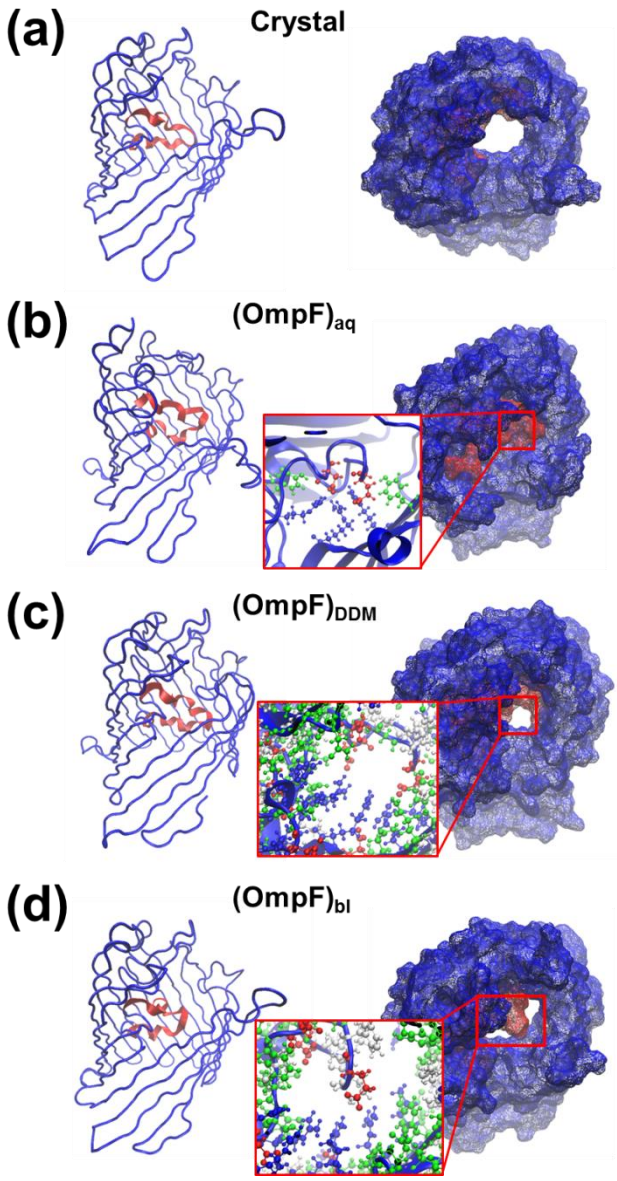


Figure 6.8. Views of the disposition L3 loop, represented in red, in the protein model (left) and representation of the pore (right) in the (a) OmpF crystal and in the last snapshot of the (b) $(OmpF)_{aq}$, (c) $(OmpF)_{DDM}$ and (d) $(OmpF)_{bl}$ production trajectories.

In spite of the L3 loop internal location, analysis of the MD productions trajectories revealed that its conformation depends on the external environment. This is reflected in Figures 6.8 and 6.9, which represent both the structure of the loop and the pore for each model and the temporal evolution of the C^{α}-RMSD considering 105-127 residues (*i.e.* the central and most relevant residues of L3), respectively. The C^{α}-RMSD averaged along the whole $(OmpF)_{aq}$, $(OmpF)_{DDM}$ and $(OmpF)_{bl}$ trajectories is 4.67±0.23, 2.70±0.37 and 3.05±0.70 Å, respectively. Early mutation studies proved that conformational changes in L3 do not affect β-barrel structure or the surface exposed loops.[32,33] Although simulations results on $(OmpF)_{DDM}$ and, especially, $(OmpF)_{bl}$ models corroborate this fact, the reverse is apparently derived from $(OmpF)_{aq}$ results. Thus, the average C^{α}-RMSD obtained for the $(OmpF)_{aq}$, which is the highest one, is probably due to the poor stability of the β-barrel in water, suggesting that structural alterations in the β-barrel induce very important conformational changes in the loop. Overall, these variations close the pore (Figure 6.8b), preventing the passage of any chemical specie. Indeed, the network of hydrogen bonding interactions and salt bridges formed among the guanidinium group of three Arg residues (42, 82 and 132), the hydroxyl moiety of two Tyr residues (40 and 106), and the carboxylate of Glu and Asp (117 and 121, respectively) is responsible for the complete closure of the pore (Figure 6.8b, inset).

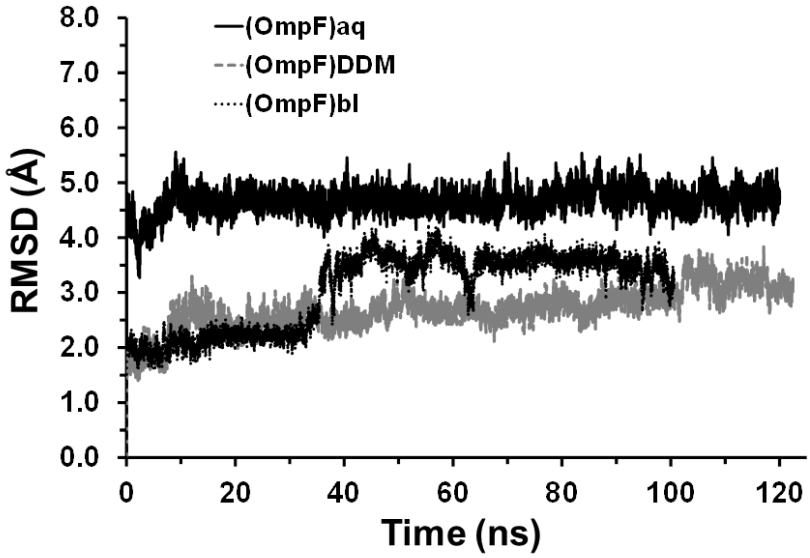


Figure 6.9. Evolution during the production trajectory of the C^{α}-RMDF for the L3 loop (residues 105-127) in the three models compared to the crystal structure.

On the other hand, the L3 motif in $(OmpF)_{DDM}$ and $(OmpF)_{bl}$ models also exhibits differences with respect to the crystal (Figure 6.9). More specifically, the pore remains partially opened in $(OmpF)_{DDM}$ and $(OmpF)_{bl}$ models (Figures 6.8c and 6.8d, respectively). In these cases the very high stability of the β-barrel partially restricts the amount of hydrogen bonds and salt bridges formed inside the cavity, even though the dimension of the pore is strongly affected by the re-arrangements of the L3 loop. In the $(OmpF)_{DDM}$ model, the L3 motif surrounds the cavity, reducing considerably the dimension of the pore to 5×8 $Å^2$ in (Figure 6.8c). In the bilayer, L3 crosses the cavity and forms several salt bridges between the side chains of Arg, Asp and Glu residues (Figure 6.8d, inset). This splits the original pore into two smaller ones with approximate dimensions 3×3 and 3×6 $Å^2$.

The L2 loop (residues 66-80) plays a crucial role in the stability of the OmpF. In bacteria membranes the OmpF porin is associated to homotrimers, each subunit including a hollow β-barrel. These homotrimers are unusually stable due mostly to the hydrophobic interactions between β-barrels. Furthermore, L2 connects one subunit to its neighbor by latching into its channel. More specifically, residue E71 on L2 is integrated into an ionic network and forms salt bridges as well as hydrogen bonds with R100 and R132 on the channel wall of the adjacent subunit. Although in this work production trajectories involve a single β-barrel subunit, analysis of the L2 flexibility or rigidity provides important information about the influence of the environment in the trimer-stabilization mechanism.[34]

As it can be seen in Figure 6.10a, in the crystal L2 projects sideways to dips into the pore of the adjacent β-barrel. However, this disposition is completely different from those achieved at the end of the production trajectories. The C^{α}-RMSD profile calculated for the $(OmpF)_{aq}$ model using residues 66-80 (Figure 6.10b) indicates that in aqueous solution the loop exhibits large fluctuations, even though it remains at the external side of the β-barrel (Figure 6.10a). Thus, the C^{α}-RMSD shows a large variation within the range 3-9 Å, with the maximum and minimum values of 14.5 and 2.1 Å, respectively. Accordingly, the aqueous environment not only affects the stability of the β-barrel but also enhances the already intrinsic flexibility of L2, the latter favoring the dissociation of the characteristic homotrimer.

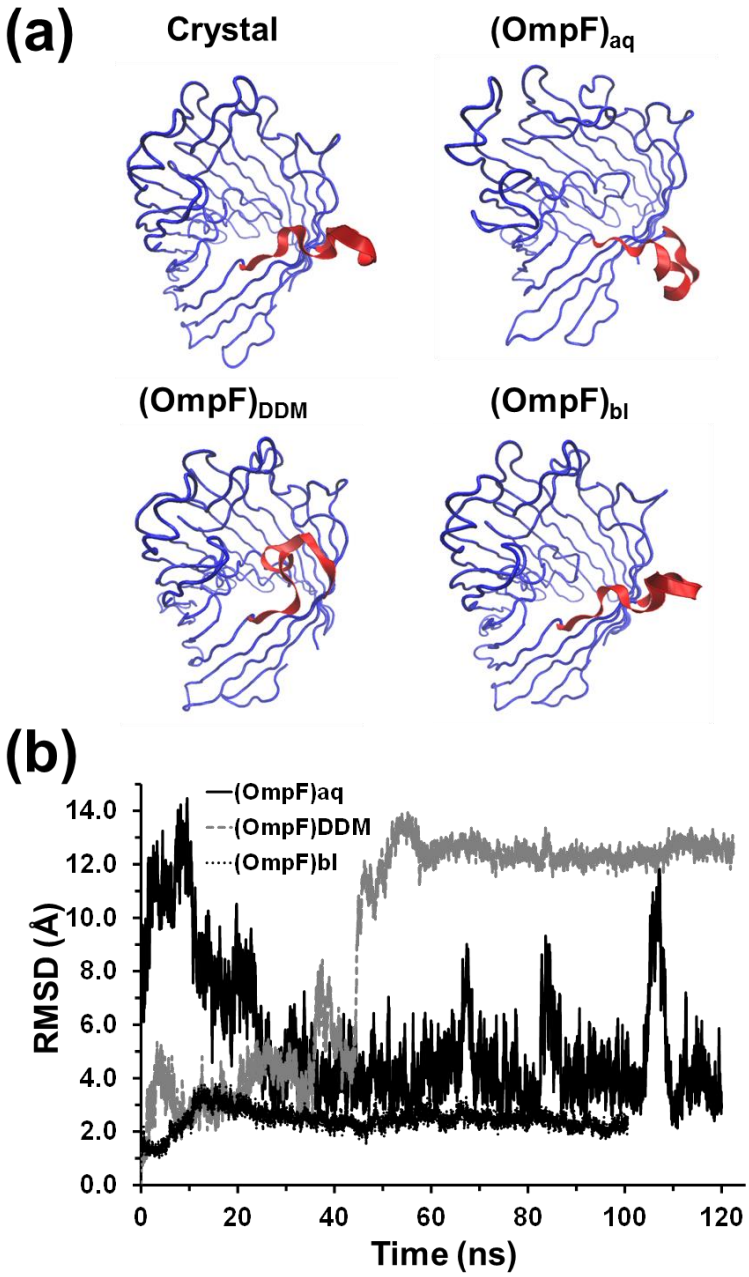


Figure 6.10. (a) Views of the L2 loop (in red) in the OmpF crystal and in the last snapshot of the $(OmpF)_{aq}$, $(OmpF)_{DDM}$ and $(OmpF)_{bl}$ production trajectories. (b) Evolution during the production trajectory of the C^{α}-RMSD for the L2 loop (residues 66-80) of the three models compared to the crystal structure.

A completely different patterns appear in the $(OmpF)_{DDM}$ model, in which the L2 loop folds over itself partially entering inside the β-barrel (Figure 6.10a). This drastic rearrangement of the loop occurs after ~44 ns (Figure 6.10b), simultaneously to the global re-structuring of the protein. Once the re-arrangement happened, the L2 loop needs ~13 ns to accommodate into the new position. After this, it remains stable, as shown by the small C^{α}-RMSD fluctuations along the time interval between 57 and 120 ns. Overall, the analysis of the evolution of L2 in the detergent solution indicates that, after the initial protection effect exerted by DDM molecules, this environment promotes structural changes much more pronounced than the fluctuations observed in aqueous solution. These results are fully consistent with electrophoretic association studies recently reported for the Omp2a porin [8]. Consequently, the homotrimeric form was found to be much less populated than the monomeric one in a detergent solution, and the low self-association capacity in this environment can be attributed to the protein⋯detergent interactions, which induce conformational changes in L2, significantly affecting the homotrimer stability.

Finally, the conformation of L2 is preserved in the detergent bilayer (Figure 6.10a). The C^{α}-RMSD averaged along the whole $(Omp2F)_{bl}$ trajectory is 2.44±0.41 Å only (Figure 6.10b), evidencing that the bilayer protects the loop from undesirable re-arrangements and, in addition, considerably reduces the conformational flexibility of such intrinsically disorganized motif. In summary, the homotrimer is predicted to be preserved in detergent bilayers like those used in this work.

6.5 CONCLUSIONS

We have presented MD simulations for a representative OMP in several surrounding environments. Our results show how the latter may affect not only the stability of the β barrel, but also the internal and surface exposed loops related with the ion transport and with the assembly of protein units in homotrimers, respectively. Water causes a very rapid and practically global destabilization of the protein structure that is accompanied by an increment of the SASA. The new structure is reached through both the spatial re-orientation of disorganized loops and the complete transformation of secondary motifs, as for example the conversion of a disorganized loop into a regular helix. These features seem to indicate that the water···OmpF interactions are apparently stronger than OmpF···OmpF interactions. However, the high dynamical activity observed for such re-orientation and transformation processes suggests a competition between the two types of interactions during the whole production trajectory.

On the other hand, DDM molecules in water induce a partial protective effect on the protein. More specifically, ordered βS and H motifs mostly maintain their structure, while poorly organized L+T motifs undergo severe conformational changes. These changes affect, amongst other things, the association capacity of OmpF, precluding the formation of homotrimers. This explains our recent observation in porin-PLA nanodevices constructed by immobilizing the protein dissolved in a neutral detergent solution onto polymeric nanomembranes.[8] Finally, the protein is perfectly protected by DDM bilayers. This protection refers to ordered motifs but also to

intrinsically disordered loops. Consequently, the OmpF maintains not only a well-defined channel structure but also its ability to form homotrimers through the L2 loop.

These results are extremely relevant for the improvement of bioinspired porin-integrated nanodevices, such as the nanoperforated PLA FsNMs for selective ion transport currently in fabrication. More specifically, we are immobilizing Omp2a, which has been previously protected using detergent bilayers, inside nanoperforations. Indeed, this is expected to facilitate the fabrication process as well as to improve the efficiency of the membrane in terms of transport.

6.6 REFERENCES

1. BENZ, R. & BAUER, K. Permeation of hydrophilic molecules through the outer membrane of gram-negativ bacteria. Review of becterial porins. *Eur. J. Biochem.* **176,** 1–19 (1988).
2. Galdiero, S. *et al.* Microbe-Host interactions: Structure and role of Gram-negative bacterial porins. *Curr. Protein Pept. Sci.* **13,** 843–854 (2012).
3. Nikaido, H. Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol. Mol. Biol. Rev.* **67,** 593–656 (2003).
4. Alcaraz, A., Nestorovich, E. M., Aguilella-Arzo, M., Aguilella, V. M. & Bezrukov, S. M. Salting out the ionic selectivity of a wide channel: The asymmetry of OmpF. *Biophys. J.* **87,** 943–957 (2004).
5. Moraes, T. F., Bains, M., Hancock, R. E. W. & Strynadka, N. C. J. An arginine ladder in OprP mediates phosphate-specific transfer across the outer membrane. *Nat. Struct. Mol. Biol.* **14,** 85–87 (2007).
6. Varma, S., Chiu, S. W. & Jakobsson, E. The influence of amino acid protonation states on molecular dynamics simulations of the bacterial porin OmpF. *Biophys. J.* **90,** 112–

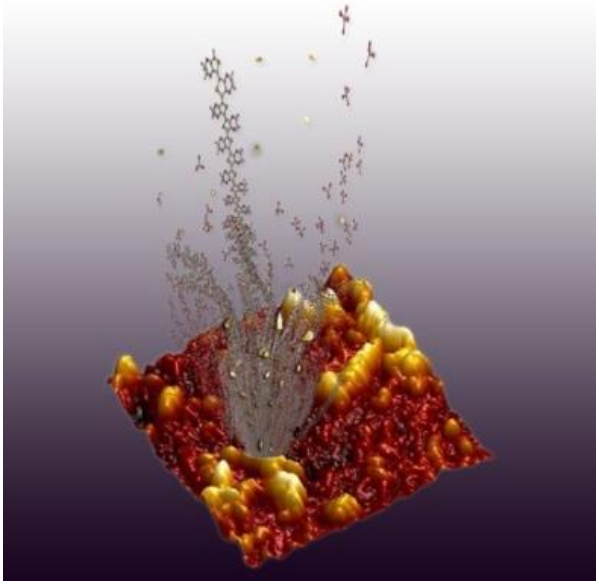
123 (2006).
7. Pérez-Madrigal, M. M. *et al.* Polypyrrole-supported membrane proteins for bio-inspired ion channels. *ACS Appl. Mater. Interfaces* **7,** 1632–1643 (2015).
8. Puiggalí-Jou, A. *et al.* Confinement of a β-barrel protein in nanoperforated free-standing nanomembranes for ion transport. *Nanoscale* **8,** 16922–16935 (2016).
9. Della Pia, E. A. *et al.* A step closer to membrane protein multiplexed nanoarrays using biotin-doped polypyrrole. *ACS Nano* **8,** 1844–1853 (2014).
10. Kumar, M., Habel, J. E. O., Shen, Y. X., Meier, W. P. & Walz, T. High-density reconstitution of functional water channels into vesicular and planar block copolymer membranes. *J. Am. Chem. Soc.* **134,** 18631–18637 (2012).
11. Mecke, A., Dittrich, C. & Meier, W. Biomimetic membranes designed from amphiphilic block copolymers. *Soft Matter* **2,** 751–759 (2006).
12. Valiokas, R., Vaitekonis, Š., Klenkar, G., Trinkunas, G. & Liedberg, B. Selective recruitment of membrane protein complexes onto gold substrates patterned by dip-pen nanolithography. *Langmuir* **22,** 3456–3460 (2006).
13. Roussel, G., Matagne, A., De Bolle, X., Perpète, E. A. & Michaux, C. Purification, refolding and characterization of the trimeric Omp2a outer membrane porin from Brucella melitensis. *Protein Expr. Purif.* **83,** 198–204 (2012).
14. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).
15. Cornell, W. D. *et al.* A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **118,** 2309–2309 (1996).
16. Duan, Y. *et al.* A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **24,** 1999–2012 (2003).
17. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926–935 (1983).
18. Abel, S., Dupradeau, F. Y., Raman, E. P., MacKerell, A. D. & Marchi, M. Molecular simulations of dodecyl-β-maltoside micelles in water: Influence of the headgroup conformation

and force field parameters. *J. Phys. Chem. B* **115,** 487–499 (2011).
19. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089–10092 (1993).
20. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81,** 3684–3690 (1984).
21. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101,** 4177–4189 (1994).
22. Tamerler, C. & Sarikaya, M. Molecular biomimetics: Utilizing nature's molecular ways in practical engineering. *Acta Biomater.* **3,** 289–299 (2007).
23. Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C. & Scharf, M. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem.* **16,** 273–284 (1995).
24. Kefala, G. *et al.* Structures of the OmpF porin crystallized in the presence of foscholine-12. *Protein Sci.* **19,** 1117–1125 (2010).
25. Paquet, J. Y., Vinals, C., Wouters, J., Letesson, J. J. & Depiereux, E. Topology prediction of brucella abortus omp2b and omp2a porins after critical assessment of transmembrane β strands prediction by several secondary structure prediction methods. *J. Biomol. Struct. Dyn.* **17,** 747–757 (2000).
26. Dhakshnamoorthy, B., Ziervogel, B. K., Blachowicz, L. & Roux, B. A structural study of ion permeation in OmpF porin from anomalous X-ray diffraction and molecular dynamics simulations. *J. Am. Chem. Soc.* **135,** 16561–16568 (2013).
27. Faraudo, J., Calero, C. & Aguilella-Arzo, M. Ionic partition and transport in multi-ionic channels: A molecular dynamics simulation study of the OmpF bacterial porin. *Biophys. J.* **99,** 2107–2115 (2010).
28. García-Giménez, E., Alcaraz, A., Aguilella-Arzo, M. & Aguilella, V. M. Selectivity of protein ion channels and the role of buried charges. Analytical solutions, numerical calculations, and MD simulations. *J. Phys. Chem. B* **119,** 8475–8479 (2015).

29. Pothula, K. R., Solano, C. J. F. & Kleinekathï¿½fer, U. Simulations of outer membrane channels and their permeability. *Biochim. Biophys. Acta - Biomembr.* **1858,** 1760–1771 (2016).
30. Martinez, L., Andrade, R., Birgin, E. G. & Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **30,** 2157–2164 (2009).
31. Touw, W. G. *et al.* A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.* **43,** D364–D368 (2015).
32. Bainbridge, G., Armstrong, G. A., Dover, L. G., Whelan, K. F. & Lakey, J. H. Displacement of OmpF loop 3 is not required for the membrane translocation of colicins N and A in vivo. *FEBS Lett.* **432,** 117–122 (1998).
33. Lout, K. L. *et al.* Structural and functional characterization of OmpF porin mutants selected for larger pore size. I. Crystallographic analysis. *J. Biol. Chem.* **271,** 20669–20675 (1996).
34. Phale, P. S. *et al.* Stability of trimeric OmpF porin: The contributions of the latching loop L2. *Biochemistry* **37,** 15663–15670 (1998).

- Chapter 7 -

Detailed Description of the Molecular Organization behind AFM Images of Polymer Coatings: a Molecular Modeling Approach



7 DETAILED DESCRIPTION OF THE MOLECULAR ORGANIZATION BEHIND AFM IMAGES OF POLYMER COATINGS: A MOLECULAR MODELING APPROACH[§]

7.1 ABSTRACT

The current experimental techniques of surfaces characterization provide structural information on a much larger scale than that in which atomistic details can be observed. In this work, this size gap gets reduced by consistently combining different modeling techniques that leads to describe the topographic profiles of thin polymer coatings at the atomistic level. A new modeling protocol, that combines Monte Carlo generation with Molecular Dynamics relaxation, has allowed us reproducing the experimental topography of extremely thin Poly(3,4-ethylenedioxythiophene) coating films using only the generated molecular models. The clue element of this protocol relies on parceling the studied surfaces in independent small plots, of which detailed molecular models are built. Combining a finite number of independent models enables to mimic the molecular organization of a large length of film that is orders of magnitude lager than the commonly used size in molecular models. The reconstructed area reproduces the thickness and roughness of very thin polymer coatings that were explicitly obtained for this study using very short electropolymerization times. This work shows a feasible way of visualizing with atomistic detail coated surfaces with polymeric films.

[1] This work has been has been published as Zanuy, D., Fabregat, G., Triguero, J. & Alemán, C. Detailed Description of the Molecular Organization behind AFM Images of Polymer Coatings: A Molecular Modeling Approach. *J. Phys. Chem. C* **122,** 20261–20270 (2018).

7.2 INTRODUCTION

The atomistic details of polymer based materials are seldom available from microscopic observations. Most of the experimental techniques used to characterize the topographic features roughly encompass scales that reach decimals of micrometers. These technical limitations are generally ignored and most of the available literature about the atomistic organization of polymeric materials shows structures that are fitted to low resolution images based on known systems without much attention to their actual feasibility.[1] Yet, most available structural works are completed without assessing the basic details of those systems, generally by either avoiding the atomistic nature of the polymer ensembles with simplified coarse-grained models,[2] models that do not properly described the system energy,[2] or by totally ignoring the atomic organization approaching the problem to finite elements computations.[3] Beyond these simulation limitations, experimental techniques are not generally tuned for the atomistic description of the polymer bulk, lacking the required structural finesse. Atomic Force Microscopy (AFM) and other more classic electronic microscopy techniques do not provide yet such details on polymers ultrastructure. We believe that these methodologic limitations can be overcome by strategically combining the available mathematical algorithms with Force Fields calculations. In this work, a novel modeling strategy is presented for obtaining a fast and reliable atomistic description of a grafting polymer using molecular models based on classical mechanics. Our results are validated by

comparing experimental topographic magnitudes obtained on the actual polymer films deposited on metal surfaces.

Metal surfaces coated by polymers are commonly found in nowadays literature for micro-engineering applications, because this sort of nanoscale constructs is especially suited for electronic micro-circuitry. Reduced scale circuitry requires the use of low density materials that can efficiently store and release charge. Electroactive Conducting Polymers (ECPs) are an excellent choice for becoming important building blocks on nanocircuitry.[4] ECP are sTable 7.macromolecular scaffolds that show induced and reversible electrochemical activity, which is a direct consequence of allowing several oxidation states. They show as well simplicity and versatility of synthesis (e.g. chemical synthesis, vapor-phase oxidation and electrochemical processes) and satisfactory environmental stability. Poly(3,4-ethylenedioxythiophene), abbreviated PEDOT, features the majority of those properties: excellent conductivity (up to 500 $S \cdot cm^{-1}$), optimal environmental and chemical stability, fast doping–dedoping processes and exceptional biocompatibility.[5,6] Moreover, PEDOT shows excellent electroactivity and electrostability, both explaining its high ability to reversibly exchange charge.[6–9] These features makes PEDOT suited for its inclusion as active parts in both rechargeable a batteries and electrochemical ultracapacitors.[9–12] Symmetric ultracapacitors including very thin PEDOT films (i.e. those with a thickness, ℓ, typically lower than 250 nm) are among the most reliable choices, due to their excellent capacitive properties.[13–15]

This work investigates the ultrastructure of coated metal surfaces that were designed for working as capacitor electrodes.[13–15]

Although PEDOT on steel constructs have extensively been examined by both scanning electron microscopy (SEM) and AFM,[13–15] structural details about its deposition and growth are yet very scarce. Molecular simulation approaches recently hinted few defining features of these nano-constructs never proven before.[16] Our initial attempt to represent the molecular details of the polymer bulk failed to describe the surface properties (e.g. topography and morphology), because we relied on polymer chains of unique lengths, relinquishing to explore the roughness of the polymer coated surface. Unfortunately, surface roughness is one of the ions movement determinant during oxidation and reduction cycles, i.e., explains how topography affects the charge exchange in these systems.[13–15,17]

The results presented in this work represent a significant improvement with respect our initial approach that has allowed us overcoming previous limitations, whereas we kept reducing the gap between the experimental scales in which surfaces are studied and the feasible use of force field based techniques within reasonable computation times. Therefore, we demonstrate the reliability of our approach by exploring with AFM the topography of newly synthesized PEDOT on steel electrodes using extremely short polymerization times (symbolized by the Greek letter θ), $\theta \leq 5$ seconds, and comparing the surface topographies with the roughness computed using our atomistic modeling that combined stochastic generation with molecular dynamics energy relaxation.

7.3 METHODS

7.3.1 Experimental techniques

7.3.1.1 Materials

3,4-Ethylenedioxythiophene (EDOT) and acetonitrile of analytical reagent grade were purchased from Aldrich. Anhydrous LiClO4, analytical reagent grade, from Aldrich was stored in an oven at 80°C before use in the electrochemical trials.

7.3.1.2 Polymerization

Poly(3,4-ethylenedioxythiophene) (PEDOT) was polymerized by chronoamperometry (CA) under a constant potential of 1.40 V. All the anodic electropolymerizations and electrochemical experiments conducted in a three-electrode two-compartment cell under nitrogen atmosphere (99.995% in purity) at 25°C. The anodic compartment was filled with 40 ml of a 10 mM EDOT solution in acetonitrile containing 0.1 M $LiClO_4$ as supporting electrolyte, while the cathodic compartment was filled with 10 ml of the same electrolyte solution. Steel AISI 316 L of 1×1 cm^2 was employed as working and counter electrodes. The reference electrode was an Ag | AgCl electrode containing a KCl saturated aqueous solution (E^o=0.222 V at 25°C), which was connected to the working compartment through a salt bridge containing the electrolyte solution.

7.3.1.3 Atomic force microscopy (AFM)

Topographic AFM images were obtained with an AFM Dimension 3100 microscope using a NanoScope IV controller under ambient conditions, in tapping mode. A silicon TAP 150-G

probe was used (resonant frequency and force constant of 150 kHz and 5 N/m, respectively). AFM measurements were performed on various parts of the films, which produced reproducible images similar to those displayed in this work. The scan window sizes used in this work were 5×5 and 2×2 μm^2.

7.3.1.4 Stylus profilometry

The thickness (ℓ) of the films was determined through profilometry measurements using a profilometer Dektack 6 from Veeco. Imaging of the films was conducted using the following optimized settings: stylus force 1.5 mg and speed 15 nm/s.

7.3.2 Computational methods

7.3.2.1 Generation of Starting Models

All atomistic models were built using an in house program. Our stochastic growth algorithm was recently introduced in reference 18. In this work, 60 identical PEDOT chains (made of different number of EDOT repeat units each one) were used to emulate the polymer deposition on a metallic substrate, following the approaches already described in previous works.[16,18] Before starting the growth cycles, PEDOT chains were arranged fully extended, represented as rigid conformers, oriented parallel to z-axis and were randomly distributed on the iron atoms surface. The only restriction imposed to their distribution over the steel surface (i.e. their equatorial projection on the surface plane) was to minimize the steric hindrance between neighboring chains. Lennard-Jones

potential adapted to AMBER force field was used to achieve this goal.[16,18,19]

In order to represent the asymmetric growth, sets of chains are randomly selected and make them grow by adding a random number of EDOT repeating units. At each simulation cycle, *i* new structures are built and only one will be selected. Every new structure shares with the other new ones the total amount of repeats that will have been added, amount that had been randomly selected before a new cycle started. After selecting how many RUs are going to be added, a number of chains are selected for each new structure to be built. This number of nascent chains is also randomly selected within a previously set margin. In the presented set of simulations this margin was from 1 to 10 (previous trials showed that was a good compromise between acceptance probability and generation speed). Per each chain, the number RUs that will be added is also selected randomly. Although the total number of added RUs is the same for each new built structure, the number of chains that grow and the increment on RUs per chain is totally different. In summary, each cycle generates *i* structures, each of them comes from adding $n=n_1+n_2+...n_m$ repeats to *m* chains. Each structure has a different number *m* growing chains and a different combination of n_m repeats per chain.

Each of these new structures will suffer an energy change that will be computed to ascertain the thermodynamic probability of acceptance for every introduced change. Our program uses a modified Metropolis criterion, analogous to that used in the ConRot method,[20] in which the acceptance probability is not directly

assessed by the energy difference of the starting and final point but by computing the probability of each individual change: if the chain *m* grows n_m RUs, per each new n_m repeat units placed in chain *m*, a thermodynamic probability is computed as:

$$W_{n_m} = \frac{e^{-\beta \varepsilon_k}}{\sum_{k=1}^{n_m} e^{-\beta \varepsilon_k}}$$

Where n_m is the number added new RUs, W_{n_m} is the probability weight of each combination of added RUs and β is $(T\kappa)^{-1}$, where T is Temperature in K and κ the Boltzmann constant.

The final probability of accepting a structure will be the product of all the $m \times n$ individual probabilities.

$$W_{built\ model} = \prod_m W_i$$

Each cycle finishes when one of the new model fulfills the modified Monte Carlo criterion:[18]

$$P_{acc} = \min\left(1, \frac{P_i W_{i+1}}{P_{i+1} W_i}\right)$$

It is very important not to misunderstand the meaning of this generation method: our procedure does not try to reproduce the growth mechanism of PEDOT but to build reliable atomistic models to represent the surface morphology once all the modeling stages are finished. Counterions (dopant agents) are placed along the new EDOT units in positions that reproduce the organization previously observed.[18] An example of a finished model can be seen at **Figure 7.2**, left panel.

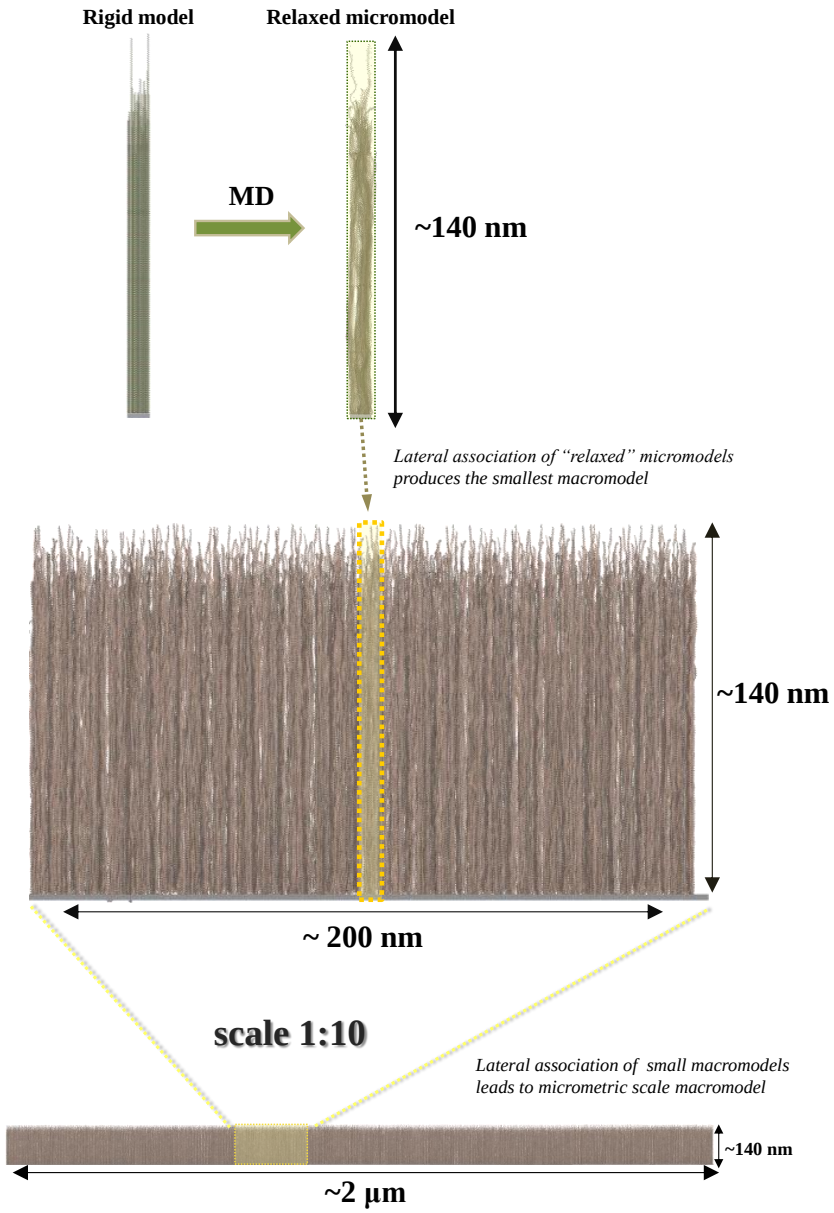


Figure 7.2.- Flow chart of the methodological approach presented in this work. ***Top panel****: from left to right*, an independent rigid micromodel (generated using stochastic methods) and its relaxed form (after a short trajectory of Molecular Dynamics). ***Middle panel:*** lateral association of 26 independent micromodels into a 200 nm surface segment (see text why 26 segments are used). ***Bottom panel*****:** lateral association of 260 micromodels

into a 2μm surface segment. Each change of order of magnitude is remarked by framing the smallest atomistic model that would lead to a larger system by merging several of those small partitions.

7.3.2.2 Energy Relaxation

Models derived from the deposition and growth algorithms are rigid and do not take into account the effect of temperature and pressure. Thus, before characterizing the topology and topography, energy optimization and MD trajectories were run for all generated models. Each model is contained in an orthorhombic cell of 7.665nm × 7.665nm × "*c*", in which the space not occupied by the substrate and polymer is filled with solvent molecules (acetonitrile in this case[16]). Depending on the total amount of RUs present in each generated model (later referred as *micromodel*) different amount of acetonitrile molecules was required to ensure the correct periodical conditions under the used cutoff (see below). Hence the chemical composition of each set of micromodels was equalized both in number of RUs and solvent composition, ranging from 7358 molecules of acetonitrile for 17200 RUs models to 10746 molecules in those that presented 28000 RUs.

7.3.2.3 Force Field details for MD

All parameters were extrapolated from AMBER03 libraries[21] with exception of partial charges of the EDOT repeating unit, which had previously been parametrized,[19,22] and ClO_4^- parameters that were obtained from Baaden et al., 2000.[23] Bond and length distances of PEDOT repeat units were approximately kept at their equilibrium distances using the same strategy employed at reference

22. Steel surface was represented as an iron crystal organized following an FCC lattice, as it was discussed in reference 16. Periodic boundary conditions were applied using the nearest image convention and atom pair cut-off distance was set at 14.0 Å to compute the van der Waals interactions. In order to avoid discontinuities in the potential energy function, non-bonding energy terms were forced to slowly converge to zero, by applying a smoothing factor from a distance of 12.0 Å. Beyond cut off distance, electrostatic interactions were calculated by using Particle Mesh of Ewald, with a points grid density of the reciprocal space of $1Å^3$.[24]

7.3.2.4 Molecular Dynamics: Simulation details

All simulations were performed using the NAMD 2.9 program.[25] Each system was submitted to 5000 steps of energy minimization (Newton Raphson method) and this was the starting point of several mini-cycles of optimization and equilibration. The numerical integration step for all runs was set at 1 fs and the non-bonded pair list was updated every 1000 steps (1 ps).

The following protocol was applied to all 1170 studied models. Solvent phase was first equilibrated. 10^5 steps of heating and equilibration at NVT conditions were run to stabilize the solvent phase at 298K. The Berendesen thermostat was used[26] with a relaxation time of 1 ps. Then, the solvent density was equalized to its optimum value using $2.5 \cdot 10^5$ steps of NPT simulation at 298 K. The Nose–Hoover[27] piston combined with the piston fluctuation control of temperature implemented for Langevin Dynamics[28] was

used in the late cycle. Pressure was kept at 1.01325 bars, the oscillation period was set at 1 ps while the piston decay time was set at 0.001 ps. The piston temperature was set at the same value as the thermostat control, 298K, which used a damping coefficient of 2 ps. During these couple of cycles, all solid phase atoms were kept frozen (iron, polymer and counterions).

Once the simulation box dimensions were stabilized, the final equilibration cycles would begin. After unfreezing the rest of atoms (solid phase with exception of the iron atoms[16]), $2 \cdot 10^5$ steps of NVT simulation were performed to thermally equilibrate the whole biphasic system. For the latter run the Langevin method[29] was used to maintain the system temperature constant with a damping coefficient of 2 ps. Finally, 1.0·106 steps of anisotropic pressure control (NPzT conditions) were run to reach the required simulation conditions. In the NPzT ensemble, only the box length in the z-direction is allowed to change, where the z-component of the pressure tensor is equal to the external pressure.[30] The last snapshot of the anisotropic equilibration was the starting point of 2 ns of production time. The production runs were all performed under the same conditions previously mentioned for the NPzT equilibration.

7.3.2.5 Computations of the topographic magnitudes

7.3.2.5.1 Thickness (ℓ)

Thickness was computed per each of the generated micromodels (see text) as the absolute average of each present chain deposited over the metallic surface.

7.3.2.5.2 Roughness (R_a and R_q)

The arithmetic average height (R_a) and the root mean square roughness (R_q) where computed for each of the generated micromodels using the following expressions:[31]

$$R_a = \frac{1}{n}\sum_{i=1}^{n}|y_i| \quad (1)$$

$$R_q = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {y_i}^2} \quad (2)$$

Where y_i is the distance from the top of a peak to the average line (as the average height of the 60 explicit chains) or the distance from the bottom of a valley to the same average line, and n represent the number of studied peaks or ridges.

7.4 RESULTS AND DISCUSSION

7.4.1 Topography of the polymer coated surface

In order to reduce the size of the structure to be modeled, we used very short polymerization times, which led to an unknown scenario, because of the uncomplete cover of the metallic surface. Under such conditions, the previously observed homogeneity of PEDOT polymer films is lost. **Figure 7.1** shows selected AFM images for several samples of PEDOT polymerized at both θ= 1s and θ= 3s. Inspection to **Figure 7.1a** confirms that the polymer coating is not completed after such short time periods, which is further manifested by higher dispersion of polymer thickness (see

below). Evidently, **Figure 7.1** also manifests that the larger is the polymerization time the more extensive is the morphologic homogeneity, showing much regular dispersion when θ is 3s than when is 1s (**Figure 7.1b** compared with the former). Nonetheless, a more detailed examination of those regions that have already been covered by polymer at θ= 1s show several features yet observed in cases in which polymerization was completed:[32] PEDOT chains tend to organize in a dense distribution of sharp peaks forming two differentiated levels. On the top level, very small number of high and compact clusters forms from aggregation of peaks, whereas the bottom level involves individual peaks and low clusters of reduced dimensions, which can be associated to most recent formed polymer chains that accumulate at the bottom of the valleys. A closer inspection to these regions reveals considerably thicker polymer coatings than could be expected. Moreover, increasing θ apparently only improves the coating uniformity of the metallic surface but not its thickness.

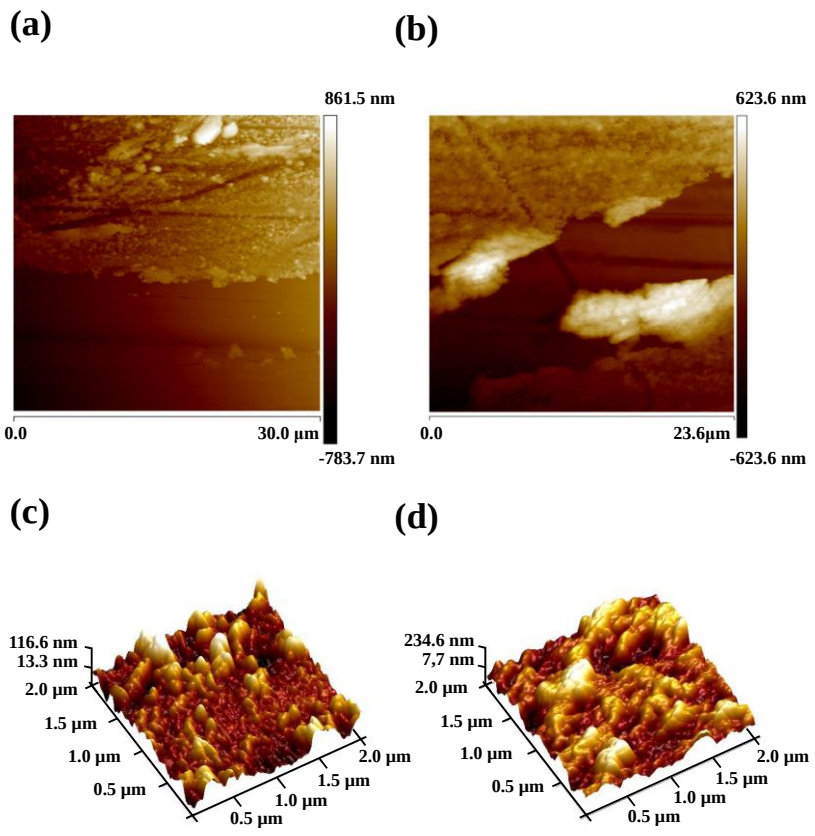


Figure 7.1.- 2D Height AFM images of: **(a)** PEDOT at θ= 1s and **(b)** at θ= 3s. 3D topographic AFM images of: **(c)** PEDOT at θ= 1s and **(d)** at θ= 3s.

A summary of the measured thickness (ℓ) and roughness (R_q and R_a)[31] per each studied case is presented in **Table 7.1**. A detailed collection of both thickness and roughness measurements is presented in the *supplementary information*. The averaged ℓ values observed in the studied PEDOT films ranged from 173±49 to 256±80 nm, respectively after θ = 1 and 3 s of polymerization (**Table 7.1**). These values are in general agreement with previous works in which

polymerization times were long enough to completely cover the metal surface.[32] As mentioned earlier, there are important differences. First of all, greater data dispersion is observed, in agreement with an incomplete deposition of polymer mass. Hence, the standard deviations of the measured thicknesses after $\theta = 1$ and 3s represent 28% and 31% respectively, of the actual measured magnitude. If such dispersion is combined with the values reported in previous works,[12–15,18] a new scenario emerges: there is always a rapid polymerization, in less than a second, that leads to the formation of significantly long polymer chains. At the same time, those early growing chains quickly and irregularly settle over the metallic surface. By contrast, at longer θ part of the deposition must be related to refilling of the aforementioned low clusters, those initially created cavities in between high peaks. This idea is strengthened by how thickness varies as function of polymerization time. A threefold increment in the polymerization time does not lead to a thickness enlargement of that order but only of 1.5 times. An increase of 10 times in polymerization times only provides two times thickness enlargement.

Polymerization time (θ)	ℓ[a]	Roughness[b]	
		R_q[c]	R_a[c]
1	173±49	45±14	37±13
3	256±80	61±6	48±5
10*	350±10	117±3	94 ±7

Table 7.1. Thickness (ℓ, in nm), roughness (R_a and R_q, in nm) measured by AFM for PEDOT films deposited on steel substrate using different polymerization times (θ, in s). Values are obtained over 5 samples per each θ.

[a] Measured by ASH profilometry. [b] Measured by AFM. [c]RMS and average roughness (R_q and R_a, respectively). *Data extracted from reference 32, for context and comparison with smaller θ shown in this work.

Roughness, determined as both root mean square (R_q) and mean average (R_a) of the deviation of the sample high from the mean line over the sampling length, shows very similar trends. None of both magnitudes follow a linear correlation with respect to θ increments, being the differences in roughness with respect to polymerization time even smaller than those observed for thickness. Visual inspection to the organization of PEDOT films shows polymer layers that feature a profoundly rugged surface, with deep valleys surrounded by significantly high peaks, as can be observed in **Figure 7.1b** that compares the 3D roughness profiles for samples obtained at both studied θ's. Again, the reduction on the surface roughness when polymerization time increases is noticeable at first sight, clearly pointing towards the same idea introduced above: first

chains of minimal length get deposited on the surface leaving pronounced valleys, partially unoccupied spaces, which at larger θ's are refilled by newly polymerized chains. Thus, the thickness increment is much higher than the increment in roughness from θ= 1s to θ= 3s.

The averaged value of roughness reveals that multiplying by 3 and by 10 the polymerization time, only implies a roughness enhancement of factors of 1.3 and 2.6, respectively (**Table 7.1**). Moreover, even the data dispersion of roughness measurements decreases when the polymerization period increases. While the standard deviation of thickness increased 63% when the θ is triplicated, roughness dispersion dramatically drops to values that are 130% lower when going from θ =1s to θ =3s. This correlation between smoother topography and larger polymerization times was already characterized in PEDOT films but using much longer polymerization times (i.e. from 10 to 300 s).[12] This effect is significantly more pronounced in the PEDOT films studied in this work, which are much thinner due to shorter θ, and again points towards the aforementioned new mechanistic interpretation. If larger θ implies smaller roughness dispersion, it necessary means that at certain time frame, between θ=1s and θ= 10s, the new polymerized segments preferably fill the gaps left in early generation steps rather than contribute to enlarge existing polymer chains. Thus, the longer is θ the

larger the polymer baseline will become (i.e. both observed thickness and roughness).

7.4.2 Modeling Polymer Roughness: Reducing dimensions of system to increase feasibility

Modeling the ultrastructure of a macromolecular assembly and doing it in a size scale comparable to that of the experimental methods is still a challenge of complex solution. Commonly, computational chemists rely on coarse graining models to face this complexity and sizes.[33,34] Yet, the information provided by these approaches is very dependent on the molecular templates used to parametrize not only the magnitudes that define the inter-molecular relationships but also the mathematic expressions used to compute them. This limitation affects the structural and physical features that can be characterized at length scales between the μm and the nm. Within this context, we decided to face these limitations trying to achieve the nm scale by reducing the scale of at least two of the three space dimensions, which made possible to obtain polymer assemblies of chain lengths closer to the μm scale than to the nanometer scale. In order to achieve this goal we adapted a previously develop strategy[16] to simulate the outcome of polymer growth.[18] It is important to clarify that our methodology does not pretend to model the polymerization reaction, neither its kinetics nor its reaction mechanism. We developed a simple and effective strategy that generates

polymer assemblies with consistent intermolecular organization.

Trying to comprise the mesoscopic organization on the surface shown in **Figure 7.1b** at once is still an impossible mission by all means. However, the observed structural disposition allows some degree of simplification. Most of the topographic features of these coated surfaces are the consequence of an asymmetric disposition of the polymeric material. There is a consistent scale misalignment between the extension of the metal surface covered by polymer chains (and their lateral organization) and their vertical growth. In other words, the topographic exploration is performed over the μm scale (pictures shown in **Figure 7.1b** represent surfaces of 4 μm^2) whereas the vertical scale is within an acceptable range for atomistic models, especially in cases corresponding to polymerization times of 1s. This quick polymerization leads to thickness and roughness in the order of hundreds nm, which is right on the feasibility of atomistic constructs. Thus, the classic strategy of divide and conquer can be applied to our problem.

We reduce most of the system complexity to a small set of variables to be optimized onto the growing direction, which is coincident with the vertical one, as we previously showed.[16,35] Hence, the lateral organization of polymer chains and their vertical growth are initially modelled disregarding the macro-organization over the metallic surface and we focus our efforts on reproducing both thickness and roughness. In order to achieve this goal the equatorial superficial projection for our

molecular model is reduced to a minimal expression (7.665 × 7.665 nm^2) whereas the vertical extend of polymer chains is only limited by the total numbers of repeating units (*RUs*) to be added onto a fixed set of nascent chains, number previously fixed to maintain the amount of polymer mass per unit of area. The only variable that will be modified in order to keep constant the density of the polymer bulk is the vertical "addition" of EDOT RUs. Before moving forward on our modelling results, we summarize the main approaches on which our in house code was based.

In order to expedite the generation of atomistic models, rigid geometry was assumed for the polymer molecules. All molecules present their molecular geometries at the equilibrium values during the polymer coating generation. The energy of the system only includes non-bonding terms, with Van der Waals energy represented by a Lennard – Jonnes potential and the electrostatic energy by the untruncated Coulomb potential. The averaged degree of oxidation per EDOT RUs had already been determined to be +0.5 charge units, under the experimental conditions used in PEDOT electropolymerization.[36] A counter-ion (perchlorate molecular anion) is added per each pair of EDOT units included in the molecular model. Thus, when our computational algorithm starts making the nascent chains grow the acceptance criterion guarantees generating assemblies without neither high steric hindrance nor electrostatic repulsions.

Before this work, we could only demonstrated that PEDOT chains laterally associated with the perchlorate anions acting as cement via electrostatic attraction between oppositely charged centers, whereas they shielded the inherent repulsion between positively charged EDOT units.[16] However, we could not correlate topography with molecular structure, mainly because our modelling algorithm dealt with very short polymer chains (10 nm), and furthermore, all those chains had the same number of repeating units, making models totally flat without roughness. In order to overcome this shortcoming, without saturating the generation of new models, we focused on elongating polymer chains that had been previously distributed in very small areas. Thus, if experimental samples extended over surfaces of 2×2 μm^2, we generated a representative number of independent models, on per each starting distribution of 60 chains over the selected steel area (7.665 × 7.665 nm^2). Details about how to choose a representative number of atomistic models are provided below for each studied topographic feature. Thus, each atomistic model based on the aforementioned surface plotting will be further referred as *micromodel* (see **Figure 7.2**, Top panel, *rigid model*).

The generation of any micromodel is finished when a targeted total amount of RUs is reached, i.e., when the summation of the RUs that each nascent chains initially present plus the number RUs incorporated during the growth procedure reaches that targeted RUs number. In order to reach an atomistic representation of the experimental size (lateral

length of one of the surface directions) a single molecular model can be built by laterally combining all generated micromodels (see below for practical details). All models are uncorrelated by definition and all of them are local energy minima respect the non-bonding potentials. This strategy is faster and more efficient than trying to cover a $4 \cdot 10^6$ nm^2 of alloy with 60000 chains of polymer and then simulate their growth, which includes also adding a perchlorate anion per every two new EDOT RUs. Such simulation would imply computing millions of non-bonding pairs per each generation cycle (addition of new RUs), making the model production totally unfeasible in terms of both computing time and memory demands. Our divide and conquer idea allows to produce in just few hours micromodels with chain lengths in the order of the μm scale (vertical growth). Each micro-model implies the computation of only few hundred thousands of non-bonding pairs per cycle, which is totally affordable with the contemporary computational power.

7.4.3 What conditions the final topography?

Despite the simplicity of our rational there are several questions that are not straightforward. We knew thickness should be the average high of the shortest chains present in the systems whereas roughness should be a reflection of the dispersion of chain lengths. Therefore, we first assessed what was the minimal amount of independent models required for reproducing the experimental thickness. As stated above, this

feature depends on the average high of the shortest chains. Within our molecular production scheme is equivalent to ascertain how many RUs are incorporated to each micromodel once the production run is finished. At this point, we needed to produce enough micromodels to make the computed data representative but we did not require yet to reproduce the whole extend of the experimentally studied surface length. If the simulated surface were within the averaged horizontal length of observed peaks, the thickness information would be statistically demonstrative. The average length of most peaks depicted in **Figure 7.1c** is 0.196 μm, which is approximately 26 times the horizontal length that was set to define the deposition surface of each micromodel.

As it was introduced earlier, each microscopic model presents only 7.665 nm of lateral length in both x and y directions. In order to have an acceptable atomistic representation, we will rely on sets of 26 independent micromodels to compute thickness per each presented case. In other words, each studied case will constituted by 26 independent microscopic distributions of 60 independent nascent chains over a metallic surface of 7.665 × 7.665 nm2 (**Figure 7.2**, *middle panel*). These 26 micromodels will only share two molecular features, the initial number RUs that the nascent chains presented at the beginning of the "growth simulation" and the total amount of EDOT RUs that each micromodels presents once the Monte Carlo procedure has finished. Lateral association of the 26 independent microscopic

models will represent the average length of the actual observed peaks. It is important to remark that the macromodel will be built by single micromodel addition at the time. Every new micromodel laterally added will then be energetically relaxed (repeating this step 25 times in this case and 259 in larger models. See below). Because we had built each micromodel with Periodic Boundary Conditions, the steric hindrance generated in the border zones of the structures placed in contact will be easily relaxed with energy minimizations.

However, there is not clear experimental evidence neither previous simulation that could convey what structural features will be reflected in the numerical value of the computed polymer thickness. Both, the length of the polymer chains that initially were deposited (or what we call in this work the *length of nascent chains*) and the random incorporation of new RUs onto those nascent chains are potential structural descriptors of this nanometric magnitude. In order to address this matter, multiple combinations of those two variables were studied, 7 different targeted final RUs amounts were set, whereas per each of these RUs Figures, 5 different starting chain lengths were set. If each case was to be represented by building 26 micromodels, a total 910 micromodels were built and further relaxed using MD simulations, as described in Methods section (An example of the morphological differences after relaxing can be seen in **Figure 7.2**, Top panel, *relaxed micromodel*).

Before getting into how PEDOT ultrastructure can be built, the actual representability of the generated macromodels was

assessed. **Figure 7.3a** plots the radial distribution function of all inter-ring distances between EDOT units belonging to different chains, represented for each of the previously built macromodels. In the same Figure 7.the main structural information obtained from X-Ray diffraction of PEDOT films deposited on steel is included.[36,37] Despite presenting large amount of disordered segments, the computed models are able to reproduce the most relevant crystallographic features observed in deposited semicrystalline films. Among them, the characteristic reflections of the known possible orthorhombic arrangements, in which the inter-chain distances within the crystalline structure and their multiple reflections, are fairly reproduced. The distances that represent the averaged inter-chain distances along x- and y-axes, at 15.2 Å and 13.6 Å, are present as distinguishable peaks. Furthermore, the characteristic reflection that corresponds to the distance between repeat units from the same chains (around 7Å) is also systematically reproduced in our built structures. All geometrical indications point towards a successful achievement of representative models.

At this point, it is possible to continue exploring the structural features correlated with measurable magnitudes: at molecular level polymer thickness would only depend on the total amount of RUs that are placed within a confined space, as can be inferred from **Table 7.2**, which shows a summary of the structural features derived from all studied cases. It is remarkable that independently of the initial chain length,

thickness converges at very similar values when the same total number of RUs is reached, which demonstrates both the consistency of our stochastic approach for building models that emulate polymer growth and the validity of our divide and conquer approach. This point is verified when the average thickness is calculated over all the models that reached the same total number of RUs. Consistently, in all cases standard deviations are below 0.5% of the averaged magnitude. Finally, this particularity also shows that our generation approach is accurate enough to provide a realistic representation of the polymer chains organization at almost nanometric scale. Once the necessary degrees of freedom had been established for reaching the right thickness, we proceed to investigate what variables conditioned the roughness description.

Total number of RUs	**Initial number of RUs per polymer chain**					
	2	**20**	**40**	**60**	**120**	$\langle \ell \rangle^a$
17200	110.5	110.8	110.9	110.9	110.9	110.8±0.2
18000	116.1	115.9	116.1	116.2	116.2	116.1±0.1
20000	128.8	128.9	129.1	129.0	129.2	129.0±0.2
22000	141.9	141.9	142.1	142.0	142.1	142.0±0.1
24000	154.8	154.7	154.8	154.8	154.9	154.8±0.1
26000	167.9	167.7	167.9	167.9	168.0	167.9±0.1
28000	180.4	180.2	180.4	180.6	180.5	180.4±0.2

Table 7.2. Average computed value of thickness depending on the total number of EDOT RUs used per each set of micromodels (left column) and

the initial length of the nascent chains (length expressed as number of RUs per polymer chain).

[a] averaged thickness over the 5 different starting points per each set of 26 micromodels with the same final amount of RUs. As mentioned in the text each micromodels consisted of 60 polymer chains randomly distributed over 58.75 nm^2. All presented values are presented in nm.

This second question is closely related with the topography of the polymer coated surface. How well roughness can be described using our random growth approach, in which polymer chains grow upwards into the accessible space? One of our clue assumptions was that local lateral organization of polymer chains would not directly interfere with the final roughness of the modeled bulk polymer. If this assumption were to be incorrect, independently of the thickness our generated micromodels, we should not be able to compute an average roughness comparable to the experimental observations. In order to answer these questions, two different possibilities were explored: either only the new RUs were responsible for the observed topographic accidents or the combination of the newly added EDOT units and the preformed chains lengths were the determinants of that observable variables.

Table 7.3 shows the comparison between experimental values obtained at θ= 1s and the computed roughness using different starting points, after reaching the amount of EDOT units required for describing the experimental thickness. As can be clearly observed, at equal amount of final RUs in each micromodel, the key factor for comprehend the surface roughness lies on the initial length of the nascent chains. Thus, independently of the total RUs

incorporated in a confined space, the topographic description of the surface seems to only depend on how large polymer chains were before beginning the deposition itself. It is remarkable that in order to reproduce the topographic profile of the studied bulky surface, each deposited chain must reach a vertical length equivalent to 130 RUs of EDOT. Therefore, the electropolymerization must be fast enough to produce polymer chains of about 60 nm even before starting its deposition over the steel surface. This statement also fits our previous observations, in which triplicating the polymerization time did not increase three times fold the thickness of the polymer coat but mainly diminished the roughness of the studied surface. This observation can be rationalized if the polymer grows fast enough and the surface coating quickly increases its thickness by the deposition of preformed PEDOT chains.

# of initial RUs[a]	**R_a**	**R_q**
1200	26.4±6.3	35.3±1.5
2400	27.4±6.4	36.3±5.6
3600	30.1±5.7	40.4±7.2
7200	35.4±7.5	45.4±1.1
7800	36.7±8.7	47.1±2.3
12480	52.7±4.6	57.0±7.7
Experimental	37±13	45±14

Table 7.3.- Comparison of computed roughness from 6 different sets of micromodels. All values are expressed in nm and per each set the standard deviation has been included (as mentioned above, 26 micromodels per set).

[a] Each set differs in the initial chain lengths but all them incorporated the same amount of residues (28000, value derived from Table 7.2), with the experimentally obtained roughness in deposited PEDOT on stainless steel after one second of polymerization. Those models that are closer to the real polymer organization have been highlighted with a gray shade.

Yet, a final question remains to be answered. We reproduce the topographic features of the studied system using very small pieces of the surface but we did not show if the combination of many small pieces reproduces a whole real profile. In other words, if our divide and conquer approach is reliable, a lateral combination of enough micromodels should provide a good explanation of the experimental topography over a lateral length of 2 μm (which is the length of used on AFM experiments). Thus, in order to have sufficiently large amount of macromodels to compare with experimental perfilometries, 1300 new micromodels were stochastically generated: each model presented 60 nascent chains with 130 RUs and the generation was completed when 12480 RUs were present in each micromodel. This latter amount of microstructures corresponds to 5 different and independent profiles over 2 μm of length per each case.

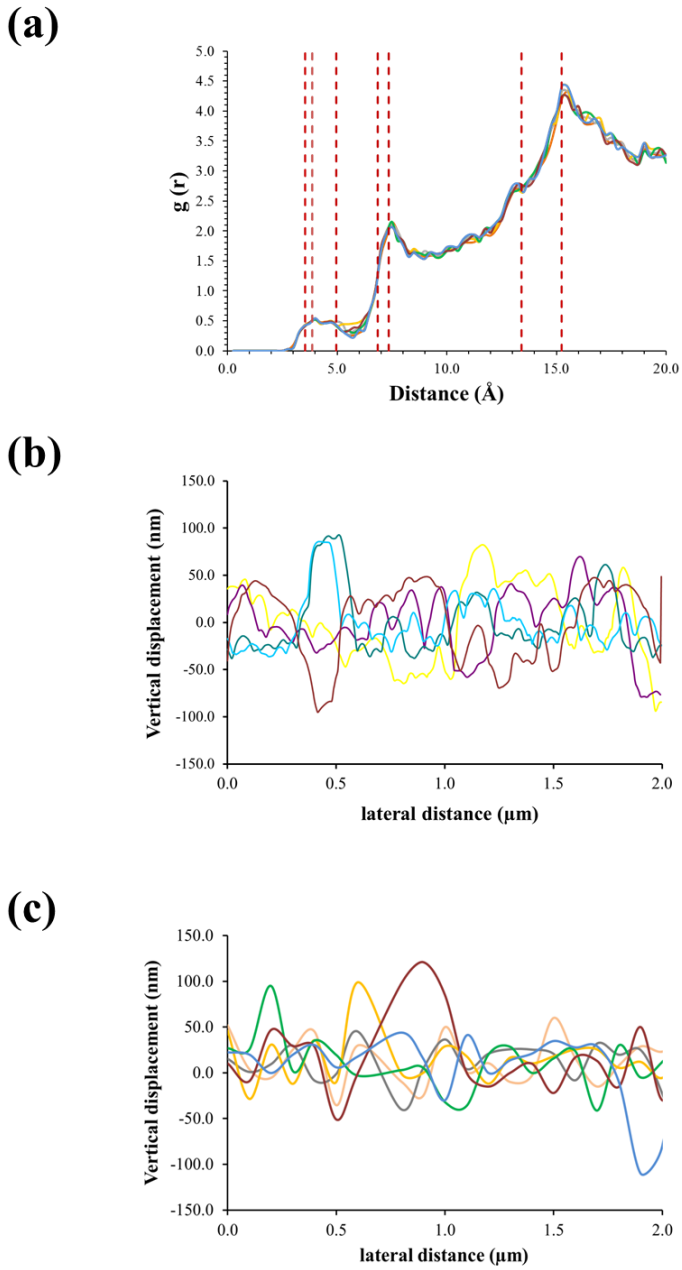


Figure 7.3.- (a) Accumulated radial distribution function of all inter chain EDOT–EDOT distances, measured from the centers of mass, for the 6 built macromodels. Dashed red lines indicate experimentally detected distances characteristic of highly crystalline PEDOT films deposited on steel

(reference 36) **(b)** Height topographic profiles obtained from AFM images of deposited PEDOT after 1s of electropolymerization. **(c)** Height profile obtained from each generated macromodel. In both panels *(a)* and *(c)*, each colored line corresponds to a single macromodel.

Once the energy of each micromodels was relaxed, all micromodels were combined in 5 single macromodels (3244800 RUs each), following the strategy introduced earlier in which the potential energy of every newly added micromodel was in turn relaxed. **Figure 7.3** compares the experimental profile obtained with the topography modeled using our approach. The similarities are significantly remarkable. Our approach of microscopically divide the studied surface in small plots and further rebuild the whole surface successfully describes the main structural and morphological features of the studied polymer film deposited steel. Not only the topographic magnitudes computed are in agreement with the experimental observations but the overall shape of the topographic details that are observed at short polymerization times are reproduced once all the micromodels are arranged in lateral clusters of 260 micromodels each.

7.5 CONCLUSIONS

In this work we have shown that using an effective modeling strategy is possible to represent the atomistic details of the bulk organization in coatings constituted by polymer depositions over metal surfaces. We developed a new synergic strategy that compiles the experimental information of the topographic features of PEDOT coatings on stainless steel and

makes possible to build atomistic models that reproduce the observed experimental features.

From an experimental standpoint, we reduced the size of the target polymer coating to the lesser expression by decreasing the electropolymerization to 1 s, which allowed making extremely thin coatings. The topographic features of those polymer layers are analogous to those observed after longer polymerization times and provided an excellent model system to test our modeling strategy. In order to find a feasible way to obtain a reliable microscopic description of PEDOT coatings, the studied surface is parceled into small independent plots, as many as was required to cover a specific surface direction. By reducing the size of the area to be explored at atomistic level, it was possible to implement a generation algorithm that produces rigid atomistic models of the bulk polymer over those parceled areas. This way, the vertical extend of the polymer coating can reach hundreds of nanometers, which represent the realistic scale in which the PEDOT coatings are organized.

Despite the generation process built unrelaxed structures, their structures are close enough to achieve a realistic description, as relaxation of geometry constrictions is easily achieved using a combination of energy minimization and short runs of Molecular Dynamics simulations. The reliability of the relaxed models is not only demonstrated when they can provide accurate values for magnitudes used to describe the topography of surfaces but by providing correct description of the polymer ultra-structure. This is possible because the

generation of micromodels ensures that all of them are uncorrelated and they can be combined to cover the surface previously parceled.

The results presented in this work represent a leap forward in the comprehension of the molecular organization on polymer coatings, because they provide a molecular description of structures that previously had to be inferred from visualization techniques that did not reach the atomistic scale. Moreover, it provides a new path to systematically study and comprehend the molecular basis of the functionalities of those coatings that generally are discovered after a new polymeric surface is developed. Using our divide and conquer approach is possible to envisage in a close future the development fine new applications of polymer coating based on specific modifications on the interface of the surface. On that region, any molecular description of an unmodified polymer can be changed *in silico* and then tested within a time frame not yet available through computational brute force.

7.6 REFERENCES

1. Li, D., Zheng, Q., Wang, Y. & Chen, H. Combining surface topography with polymer chemistry: Exploring new interfacial biological phenomena. *Polym. Chem.* **5,** 14–24 (2014).
2. Tretyakov, N. & Müller, M. Directed transport of polymer drops on vibrating superhydrophobic substrates: A molecular dynamics study. *Soft Matter* **10,** 4373–4386 (2014).
3. Yin, J., Yagüe, J. L., Boyce, M. C. & Gleason, K. K. Biaxially mechanical tuning of 2-D reversible and irreversible surface topologies through simultaneous and sequential wrinkling. *ACS Appl. Mater. Interfaces* **6,** 2850–2857 (2014).
4. Long, Y. Z. *et al.* Recent advances in synthesis, physical properties and applications of conducting polymer nanotubes and nanofibers. *Prog. Polym. Sci.* **36,** 1415–1442 (2011).
5. del Valle, L. J., Estrany, F., Armelin, E., Oliver, R. & Alemán, C. Cellular adhesion, proliferation and viability on conducting polymer substrates. *Macromol. Biosci.* **8,** 1144–1151 (2008).
6. Groenendaal, L., Jonas, F., Freitag, D., Pielartzik, H. & Reynolds, J. R. Poly(3,4-ethylenedioxythiophene) and its derivatives: past, present, and future. *Adv. Mater.* **12,** 481–494 (2000).
7. Kirchmeyer, S. & Reuter, K. Scientific importance, properties and growing applications of poly(3,4-ethylenedioxythiophene). *J. Mater. Chem.* **15,** 2077–2088 (2005).
8. Tamburri, E., Orlanducci, S., Toschi, F., Terranova, M. L. & Passeri, D. Growth mechanisms, morphology, and electroactivity of PEDOT layers produced by electrochemical routes in aqueous medium. *Synth. Met.* **159,** 406–414 (2009).
9. Aradilla, D., Estrany, F. & Alemán, C. Symmetric supercapacitors based on multilayers of conducting polymers. *J. Phys. Chem. C* **115,** 8430–8438 (2011).
10. Aradilla, D. *et al.* Hybrid polythiophene-clay exfoliated nanocomposites for ultracapacitor devices. *J. Mater. Chem.* **22,** 13110–13122 (2012).
11. Xuan, Y., Sandberg, M., Berggren, M. & Crispin, X. An all-polymer-air PEDOT battery. *Org. Electron. physics, Mater.*

Appl. **13,** 632–637 (2012).

12. Aradilla, D. *et al.* Nanometric ultracapacitors fabricated using multilayer of conducting polymers on self-assembled octanethiol monolayers. *Org. Electron. physics, Mater. Appl.* **14,** 1483–1495 (2013).
13. Aradilla, D., Estrany, F. & Alemán, C. Synergy of the I-/I3- redox pair in the capacitive properties of nanometric poly(3,4-ethylenedioxythiophene). *Org. Electron.* **14,** 131–142 (2013).
14. Aradilla, D., Estrany, F., Armelin, E. & Alemán, C. Ultraporous poly(3,4-ethylenedioxythiophene) for nanometric electrochemical supercapacitor. *Thin Solid Films* **520,** 4402–4409 (2012).
15. Aradilla, D., Estrany, F., Casellas, F., Iribarren, J. I. & Alemán, C. All-polythiophene rechargeable batteries. *Org. Electron. physics, Mater. Appl.* **15,** 40–46 (2014).
16. Zanuy, D. & Alemán, C. Resolving the subnanometric structure of ultrathin films of poly(3,4-ethylenedioxythiophene) on steel surfaces: A molecular modeling approach. *Soft Matter* **9,** 11634–11644 (2013).
17. Ahumada, O. *et al.* Sensitive thermal transitions of nanoscale polymer samples using the bimetallic effect: Application to ultra-thin polythiophene. *Rev. Sci. Instrum.* **84,** (2013).
18. Torras, J. *et al.* Close contacts at the interface: Experimental-computational synergies for solving complexity problems. *Phys. Sci. Rev.* **3,** (2018).
19. Teixeira-Dias, B. *et al.* Influence of the Doping Level on the Interactions between Poly(3,4- ethylenedioxythiophene) and Plasmid DNA. *Macromol. Chem. Phys.* **211,** 1117–1126 (2010).
20. Curcó, D., Zanuy, D., Nussinov, R. & Alemán, C. A simulation strategy for the atomistic modeling of flexible molecules covalently tethered to rigid surfaces: Application to peptides. *J. Comput. Chem.* **32,** 607–619 (2011).
21. Duan, Y. *et al.* A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **24,** 1999–2012 (2003).
22. Preat, J., Zanuy, D., Perpete, E. A. & Aleman, C. Binding of cationic conjugated polymers to DNA: Atomistic simulations of adducts involving the Dickerson's dodecamer.

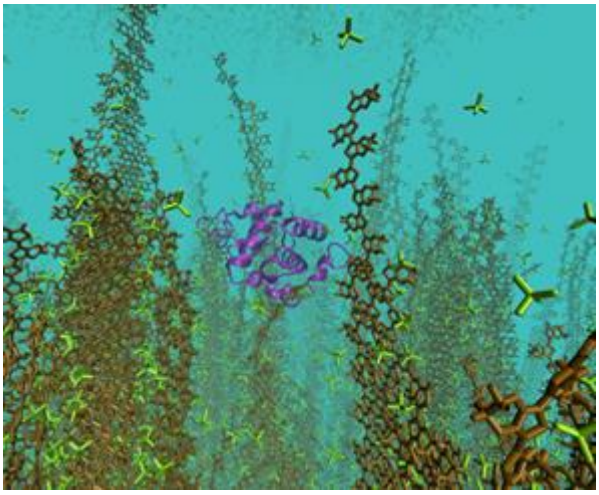
Biomacromolecules **12,** 1298–1304 (2011).

23. Baaden, M., Burgard, M., Boehme, C. & Wipff, G. Lanthanide cation binding to a phosphoryl-calix[4]arene: The importance of solvent and counterions investigated by molecular dynamics and quantum mechanical simulations. *Phys. Chem. Chem. Phys.* **3,** 1317–1325 (2001).
24. Toukmaji, A., Sagui, C., Board, J. & Darden, T. Efficient particle-mesh Ewald based approach to fixed and induced dipolar interactions. *J. Chem. Phys.* **113,** 10913–10927 (2000).
25. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).
26. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81,** 3684–3690 (1984).
27. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101,** 4177–4189 (1994).
28. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **103,** 4613–4621 (1995).
29. Toxvaerd, S. Molecular dynamics calculation of the equation of state of alkanes. *J. Chem. Phys.* **93,** 4290–4295 (1990).
30. Wennberg, C. L., Murtola, T., Hess, B. & Lindahl, E. Lennard-Jones lattice summation in bilayer simulations has critical effects on surface tension and lipid properties. *J. Chem. Theory Comput.* **9,** 3527–3537 (2013).
31. Gadelmawla, E. S., Koura, M. M., Maksoud, T. M. A., Elewa, I. M. & Soliman, H. H. Roughness parameters. *J. Mater. Process. Technol.* **123,** 133–145 (2002).
32. Maione, S. *et al.* Electro-biocompatibility of conjugates designed by chemical similarity. *J. Pept. Sci.* **20,** 537–546 (2014).
33. Heinz, H. & Ramezani-Dakhel, H. Simulations of inorganic-bioorganic interfaces to discover new materials: Insights, comparisons to experiment, challenges, and opportunities. *Chem. Soc. Rev.* **45,** 412–448 (2016).
34. Janke, W. & Paul, W. Thermodynamics and structure of macromolecules from flat-histogram Monte Carlo simulations. *Soft Matter* **12,** 642–657 (2016).

35. Martin, D. C. *et al.* The morphology of poly (3, 4-ethylenedioxythiophene). *Polym. Rev.* **50,** 340–384 (2010).
36. Ocampo, C., Oliver, R., Armelin, E., Alemán, C. & Estrany, F. Electrochemical synthesis of poly(3,4-ethylenedioxythiophene) on steel electrodes: Properties and characterization. *J. Polym. Res.* **13,** 193–200 (2006).
37. Niu, L., Kvarnström, C., Fröberg, K. & Ivaska, A. Electrochemically controlled surface morphology and crystallinity in poly(3,4-ethylenedioxythiophene) films. *Synth. Met.* **122,** 425–429 (2001).

- Chapter 8 -

Impact of Protein-Polymer Interactions in the Antimicrobial Activity of Lysozyme/Poly(3,4-ethylenedioxythiophene) Biocapacitors



8 IMPACT OF PROTEIN-POLYMER INTERACTIONS IN THE ANTIMICROBIAL ACTIVITY OF LYSOZYME/POLY(3,4-ETHYLENEDIOXYTHIOPHENE) BIOCAPACITORS[§]

8.1 ABSTRACT

Biocapacitors constructed by combining lysozyme (LYZ) and poly(3,4-ethylenedioxythiophene) (PEDOT) retained the bactericidal activity of the protein when this was encapsulated within the polymeric matrix but lost the antimicrobial behaviour when the LYZ was adsorbed onto the polymer. In this work we use atomistic Molecular Dynamics simulations to examine the influence of protein···polymer interactions in the bactericidal activity of LYZ-containing biocapacitors. Results show that the anisotropic forces exerted by oxidized PEDOT chains on the adsorbed protein induce small structural changes that locally affect at the active centre, breaking the intra-residue interactions associated with the antibacterial mechanism. Conversely, isotropic polymer···protein interactions in biocapacitors with encapsulated LYZ do not affect the stability of the active centre. These observations suggest that the strong repulsive or attractive forces between p-doped polymer chains and the positively or negatively charged LYZ residues, respectively, are the only ones responsible for the protein activity.

[1] This work has been has been published as Triguero, J., Zanuy, D. & Alemán, C. Impact of Protein-Polymer Interactions in the Antimicrobial Activity of Lysozyme/Poly(3,4-ethylenedioxythiophene) Biocapacitors. *ChemistrySelect* **3,** 9714–9724 (2018).

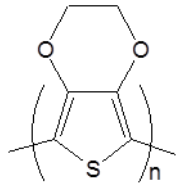
8.2 INTRODUCTION

Biocapacitor is a recently coined term used for biodevices inspired by proteins to store and/or generate energy.[1,2] More specifically, single electrodes in which proteins are combined with synthetic materials, as for example conducting polymers (CPs), rechargeable devices able to store large amounts of electrical energy (supercapacitors) or to store the potential generated by a charge pump (enzymatic fuel cells), have been fabricated.[1–6] The potential of biocapacitors in the biomedical field is enormous since the protein can help not only to improve the biocompatibility of the devices for in vivo implants,[2,5] but also can be used to develop bioinspired sensing actuators for drug delivery/pumping systems.[6] In recent studies we reported the preparation of electrodes for biocapacitors using chicken egg white lysozyme (LYZ), also known as muramidase, and poly(3,4-ethylenedioxythiophene) (PEDOT). [2,5]

Although LYZ is mostly known for its hydrolase activity on the peptidoglycan of Gram-positive bacteria, recent studies have evidenced that this small protein of 129 residues (14.4 kDa) also exhibits activity on Gram-negative bacteria.[7,8] In fact, there are numerous reports on the usage of LYZ as antimicrobial agent in food processing[9–12] and biomedical applications, as for example biosensing devices and therapeutic.[13–15,]This natural antibiotic exhibits an ellipsoidal shape in which ordered secondary structures (~40% α-helix and ~10% β-sheet) coexist with a large amount of random coils and turns.[16,17]

Among conjugated polymers CPs, PEDOT (Scheme 8.1) is one of the most important because of its great environmental stability,

electrical conductivity, electrochemical activity, thermoelectric behavior and high specific capacitance.[18–22] Additionally, PEDOT offers two fundamental advantages as compared to unsubstituted polythiophene (PTh) or other PTh derivatives: (1) the fused dioxane ring blocks the β-position of the thiophene ring preventing from the formation of α-β linkages during the polymerization, which makes the resulting polymer well-defined from a regiochemical point of view; and (2) the fused dioxane ring and the electron-donating effects provided by the oxygen atoms cause the self-rigidification of polymer chains, resulting in a gain of aromaticity and in a reduction of the oxidative doping potential.[23]



PEDOT

Scheme 8.1. Chemical structure of PEDOT

LYZ-containing PEDOT electrodes for biocapacitors were prepared conserving two approaches that differ in the strategy used to immobilize the protein. In the first one, the protein was simply adsorbed onto the surface of CP films.[5] The resulting system, named PEDOT/LYZ, consisted on a biphasic system able to retain the electroactive and electrical properties of PEDOT but unable to protect against bacterial growth. In the second approach, LYZ was homogeneously distributed inside the PEDOT matrix, the resulting material being denoted P(EDOT-LYZ) displaying excellent

antibacterial activity. The latter was prepared by *in situ* polymerization using a medium containing both the protein, which acted as dopant agent, and the 3,4-ethylenedioxythiophene (EDOT) monomer.[2,5] Thus, as CPs behave as dynamical systems able to exchange dopant ions with the medium, LYZ was slowly escaping from the polymer matrix to the culture medium while other ions (including other proteins) entered into it.[2,5] After optimization of the polymerization conditions, P(EDOT-LYZ) hybrid electrodes exhibited higher electrochemical activity and biocompatibility than PEDOT.[5] These results evidenced the potential of P(EDOT-LYZ) as electrodes for electrochemical biocapacitors useful for *in vivo* implants.

In this work we use atomistic Molecular Dynamics (MD) simulations to explain the influence of protein⋯polymer interactions in the antimicrobial activity of LYZ-containing PEDOT electrodes. For this purpose, simulations have been conducted on two models that contain all the elements used for the fabrication PEDOT/LYZ and P(EDOT-LYZ) electrodes. Specifically MD simulations have been performed in absence and presence of electric fields, the latter mimicking the one originated by the voltage cell difference in biocapacitors. Interestingly, results indicate that the directional effects of the forces exerted by the CP on LYZ are responsible for the loss and retention of the biological activity in PEDOT/LYZ and P(EDOT-LYZ), respectively, while unexpectedly the influence of the electric field in the loss of protein activity is negligible.

8.3 CONSTRUCTION OF PEDOT/LYZ AND P(EDOT-LYZ) MODELS AND APPROACH FOR THE ELECTRIC FIELD

The construction of both models started by placing one LYZ molecule in its crystal structure, 1LKS entry[17] (Hen egg white lysozyme nitrate) from the Protein Data Bank (PDB), at the centre of a cubic simulation box (a= 190 Å) fulfilled with 19508 previously equilibrated (1 atm and 298 K) explicit water molecules. The charge of all potential titratable residues (6 Lys, 11 Arg, 7 Asp and 2 Glu) was fixed to the values corresponding to neutral pH (*i.e.* Asp and Glu side chains were negatively charged, while Lys and Arg side groups were represented in their positively charged forms). Then, eight ClO_4^- anions were added to the simulation box, close to positively charged residues, to reach the neutrality. This system, hereafter denoted LYZ/water, was thermalized, equilibrated and relaxed using a three-step process, which can be summarized as follows: 1) 2×10^5 steps of energy minimization; 2) 2.1 ns of NVT MD at 298 K; and 3) 35 ns of NPT MD at 298 K and 1.01325 bars. Figure 8.1a, which displays the superposition of the LYZ as obtained from the last snapshot of the NPT MD simulation (blue) and in the crystal (red), indicates that the simulated protein maintains the structural characteristics observed in the solid state. Thus, the root mean square deviation (RMSD) between such two structures was of 2.162 Å only.

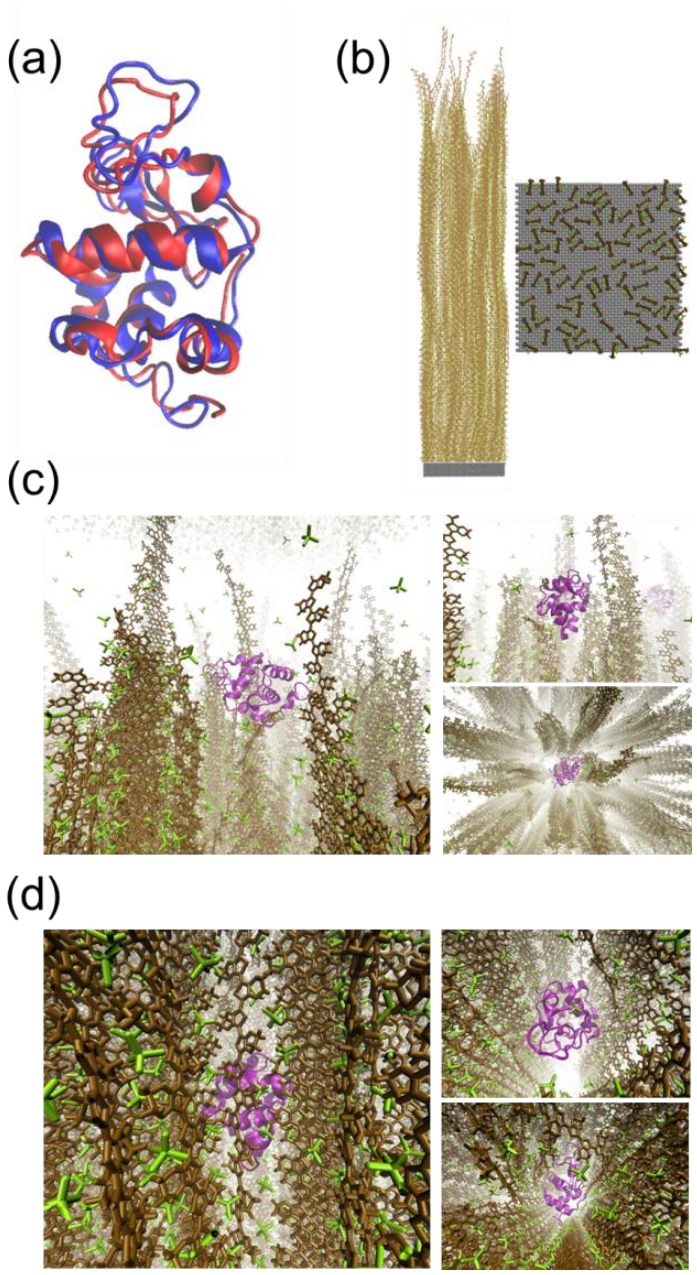


Figure 8.1. (a) Superposition of the LYZ in the crystal (1LKS entry in PDB; in red) and as obtained from the MD relaxation of the LYZ/water model (last snapshot from a 35 ns NPT MD simulation; in blue). (b) Lateral view (left) of the model constructed to mimic ultra-thin PEDOT films deposited onto stainless steel. The relative position of the PEDOT chains onto the iron substrate is also

displayed (right). (c,d) Different views of the (c) PEDOT/LYZ and the (d) P(EDOT-LYZ) models illustrating the position of the protein molecule relative to the CP chains. The two models were constructed combining the results derived from the LYZ/water simulation and the PEDOT film model, as it is described in the text. The protein, PEDOT chains and ClO_4^- anions are represented in purple, brown and yellowish green, respectively, while the iron substrate and the water molecules have been omitted for clarity.

Besides, an ultra-thin PEDOT film deposited over the surface of austenitic stainless steel, which was the substrate experimentally used to prepare both PEDOT/LYZ and P(EDOT-LYZ) electrodes,[2,5] was modelled using a previously described approach.[24] In brief, the steel substrate was represented using a 3D organization of frozen iron atoms organized in the face-centred cubic (fcc) crystal lattice observed for steel (cell parameter: a= 3.65 Å).[25] Thus, as the metallic substrate acted as a simple anchoring base for PEDOT chains, the austenitic steel was represented as an iron lattice, its minor components being omitted from the model. Specifically, 13372 iron atoms were distributed in nine layers (c= 14.6 Å) according to the (111) crystallographic plane, which corresponds to a surface of 98.55 × 100.35 $Å^2$. After this, 100 PEDOT chains with a number of EDOT units ranging between 68 and 138 were grown parallel to the crystallographic direction c (z-axis) onto the iron substrate using a generation strategy that combines stochastic methods and MD simulations.[24] This strategy was proved to reproduce the main nanofeatures (*i.e.* topography and morphology) experimentally observed at the surface of ultra-thin PEDOT films. The total charge of every EDOT unit was kept at +0.5 and, as

determined experimentally in previous work.[26] Therefore, one ClO_4^- anion, which was the dopant agent experimentally used for the preparation of PEDOT/LYZ and P(EDOT-LYZ) electrodes,[2,5] was introduced every two EDOT units to neutralize the system. ClO_4^- ions were placed within the polymer matrix using an early strategy envisaged to identify free volume regions within any polymeric solid.[27] Obviously, the density calculated for the PEDOT model, which contained 11000 EDOT units and 5500 perchlorate anions, ρ=1.60 g/cm^3, fits the experimental density as determined by the flotation method, ρ= 1.65 g/cm^3.[26] The resulting PEDOT model, which was equilibrated and relaxed using the protocol reported in our previous work,[24] is displayed in Figure 8.1b.

PEDOT/LYZ and P(EDOT-LYZ) models were constructed placing the PEDOT model in orthorhombic cells of 98.55 × 100.35 × 559.0 Å^3 and 98.55 × 100.35 × 601.0 Å^3, respectively, and using the results from the LYZ/water simulation. Specifically, the PEDOT/LYZ model was built by transferring the LYZ protein, the eight ClO_4^- anions and the water molecules from the first hydration shell (*i.e.* water molecules with the oxygen atom at a distance shorter than 2.5 Å from any polar atom of the protein or the ClO_4^- anions), as obtained from the last snapshot of the LYZ/water simulation, onto the PEDOT surface. The remaining free space of the simulation box was filled with water molecules from a previously equilibrated simulated solvent box that reproduced the experimental density of this liquid. This PEDOT/LYZ model, which is depicted in Figure 8.1c, contained 13372 iron atoms, 11000 EDOT units, one LYZ molecule, 5508 ClO_4^- anions and 112102

water molecules (*i.e.* a total of 522286 explicit atoms). Finally, the whole system was relaxed and equilibrated using the protocol described in the Methods section.

For the construction of the P(EDOT-LYZ) model a four-step strategy was engineered to encapsulate the protein inside the CP matrix. First, a very efficient approach[27] was used to identify all regions with unoccupied volume within the PEDOT film model (Figure 8.1b). After this, the largest unoccupied region was selected and filled with four spherical fictitious atoms of radius R= 10 Å and well depth ε= 0.20 kcal/mol. Third, the whole system, constituted by the PEDOT film model plus the four fictitious atoms was minimized and, subsequently, relaxed using a short NVT MD (5 ns). This step was repeated 15-times increasing slowly both the radius and well depth of fictitious atoms (*i.e.* ΔR= 2 Å and $\Delta\varepsilon$= 0.1 kcal/mol) each time until reach the following parameters: R= 40 Å and well depth ε= 0.35 kcal/mol. Finally, the four fictitious atoms were replaced by the LYZ, the eight ClO_4^- anions and the water molecules from the first hydration shell as obtained from the last snapshot of the LYZ/water simulation. Finally, the P(EDOT-LYZ) model was completed by filling the free space of the simulation box with water molecules from a previously equilibrated simulated solvent box. This model, which was subjected to relaxation and equilibration as described in the Methods section, is displayed in Figure 8.1d and contained 13372 iron atoms, 11000 EDOT units, one LYZ molecule, 5508 ClO_4^- anions and 124996 water molecules (*i.e.* a total of 560968 explicit atoms).

The most direct approach for simulating an electric field consists of constructing a layer of explicit positive charges on one side of the simulation box and another layer of explicit negative charges on the opposite side of the simulation box (*i.e.* the system comprises two parallel layers with fixed charges separated by the bulk phase). Different charge concentrations in each layer can be introduced to create a realistic potential. However, the dual-layer approach considerably increases the size of the simulated systems and the computational cost. Thus, in order to avoid unrealistic border effects derived from artificially created layers, the bulk phase is required to be large enough. Although the dual-layer method might become more popular as computational power increases and simulations of larger systems become accessible, the constant electric field approach allows to avoid such undesirable effects. In this strategy, a uniform electric field, $\overrightarrow{E_z}$, is introduced throughout the entire simulated periodic cell. This gives rise to a force $q_i \cdot \overrightarrow{E_z}$ that applies to all charges q_i in the simulation. The electric field is directed perpendicular to the box plane and must have a magnitude $\overrightarrow{E_z} = V/L_z$, where L_z is the length of the simulation box in that direction. Thus, the value of the applied voltage V is known *a priori* and the magnitude of the applied field depends only on the size of the simulation box L_z, with no need for a dual-layer with an enlarged system. This approach, which has been employed in this work, was successfully used in many applications studying ion conduction, voltage-regulated water flux, and induced conformational changes of membrane proteins.[28–30]

8.4 RESULTS AND DISCUSSION

8.4.1 Structural analysis of the adsorbed and encapsulated protein

Figures 8.2a and 8.3a display the temporal evolution of the RMSD for the protein in PEDOT/LYZ and P(EDOT-LYZ) models, respectively, which was calculated with respect to the initial conformation and considering all backbone and side chain non-hydrogen-atoms. In absence of electric field the RMSD of the protein exhibits small fluctuations in both models. Considering the last 20 ns of each trajectory, the average RMSD is 2.52±0.91 and 2.04±0.08 Å for PEDOT/LYZ and P(EDOT-LYZ), respectively, suggesting that the backbone remains structurally stable during the whole simulations while side chains experience some deviations. It is worth noting that the structural distortion is higher when protein is adsorbed onto the CP surface than when it is embedded inside the CP matrix.

Amazingly, the electric field does not cause relevant changes in the RMSD (Figures 8.2a and 8.3a), independently of its strength and sense. Indeed, in all cases the RMSD is lower in presence than in absence of electric field. For example, the averaged RMSD considering the last 20 ns of the 60 ns production run for PEDOT/LYZ and P(EDOT-LYZ) is 2.01±0.13 and 1.75±0.08 Å, respectively, when $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$), which represents a reduction of around 15-20% with respect to the averages obtained in absence of electric field. This tendency is clearly maintained when the length the trajectories was increased to 250 ns (*i.e.* the averaged RMSD considering the last 200 ns of the 250 ns run for

PEDOT/LYZ and P(EDOT-LYZ) with when $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$) is 2.31±0.18 and 1.65±0.14 Å, respectively), as is reflected in Figure 8.4.

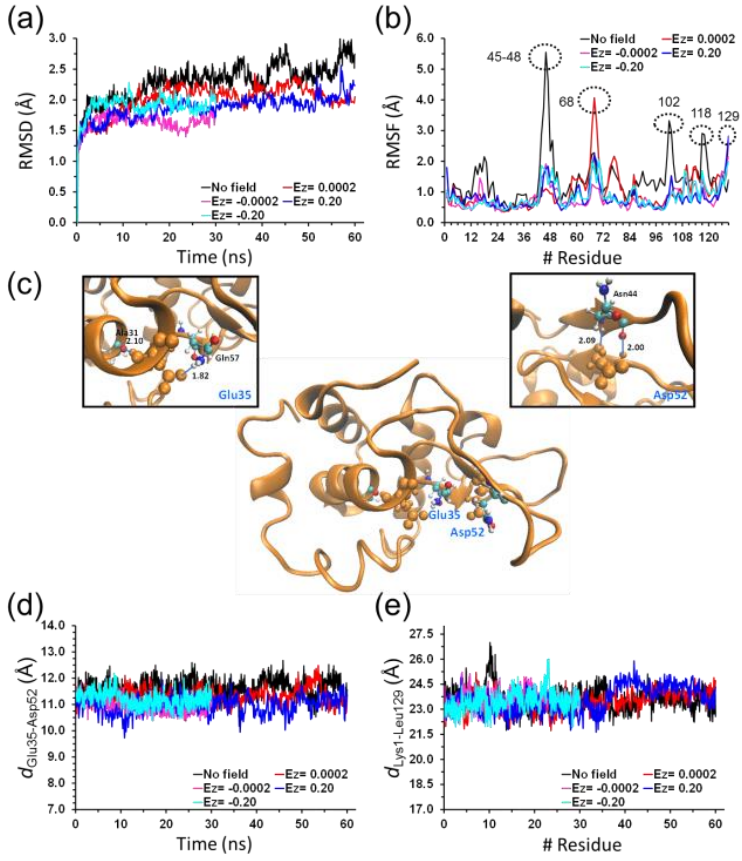


Figure 8.2. For PEDOT/LYZ simulations in absence and presence of electric field: (a) Temporal evolution of the RMSD, which was calculated considering all the protein non-hydrogen-atoms; (b) Residue based RMSF at the end of the simulation, which was calculated considering all the protein non-hydrogen-atoms; (c) Details about the interactions formed by the Glu35 and Asp52 residues (represented by balls in light brown), which are responsible for the antimicrobial activity of LYZ. Magnification of the Glu35- and Asp52-containing regions is displayed (left and right, respectively). (d,e) Distances (in Å) correspond to

specific inter-residue interactions; and temporal evolution of the distance between the Cα-atoms of (d) Glu35 and Asp52 ($d_{\text{Glu35-Asp52}}$), and (e) Lys1 and Leu129 ($d_{\text{Lys1-Leu129}}$).

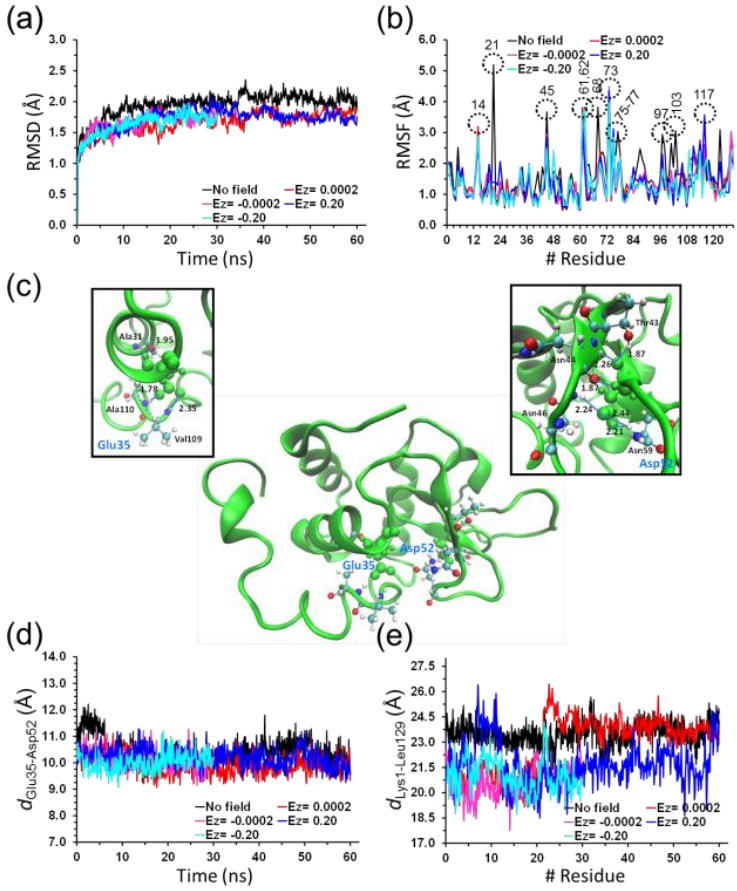


Figure 8.3. For P(EDOT-LYZ) simulations in absence and presence of electric field: (a) Temporal evolution of the RMSD, which was calculated considering all the protein non-hydrogen-atoms; (b) Residue based RMSF at the end of the simulation, which was calculated considering all the protein non-hydrogen-atoms; (c) Details about the interactions formed by the Glu35 and Asp52 residues (represented by balls in green), which are responsible for the antimicrobial activity of LYZ. Magnification of the Glu35- and Asp52-containing regions is displayed (left and right, respectively). (d,e) Distances (in Å) correspond to

specific inter-residue interactions; and temporal evolution of the distance between the Cα-atoms of (d) Glu35 and Asp52 ($d_{Glu35\text{-}Asp52}$), and (e) Lys1 and Leu129 ($d_{Lys1\text{-}Leu129}$).

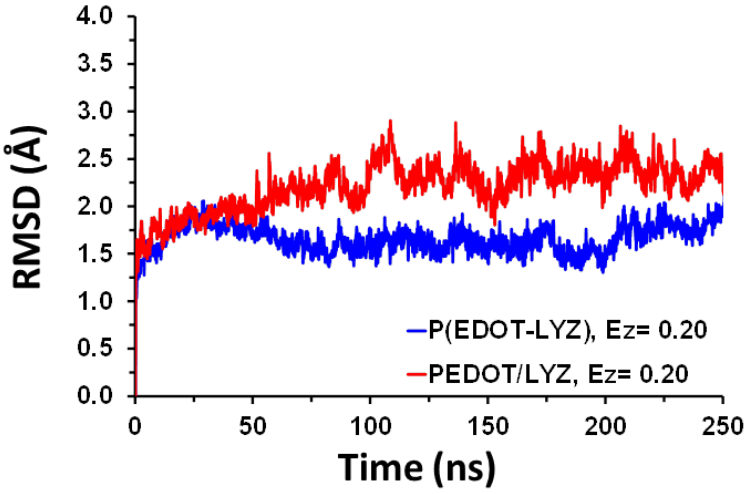


Figure 8.4. Temporal evolution of the RMSD, which was calculated considering all the protein non-hydrogen-atoms, for PEDOT/LYZ and P(EDOT-LYZ) simulations (250 ns) applying $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$).

To assess the location of the major distortions of the constituting residues, the root mean square fluctuations (RMSF) of all non-hydrogen atoms from the protein in PEDOT/LYZ and P(EDOT-LYZ) was measured with respect to their positions in the crystal (Figures 8.2b and 8.3b, respectively). According to the Phillips mechanism,[31] Glu35 and Asp52 residues play an essential role in the antibacterial mechanism of LYZ. The protein damages bacterial cells by cleaving the β(1→4) glycosidic linkage between alternating units of *N*-acetylmuramic acid and *N*-acetylglucosamine, which are building blocks of bacterial cell walls. The terminal proton of Glu35

is transferred to the oxygen atom of the glucosidic bond between two neighbouring sugar residues, leading to the cleavage of the glucosidic bond and the formation of the carbenium ion. The positive charge of this carbenium ion is stabilized by the negative charge of Asp52 until a hydroxyl ion binds to the positive carbon atom and Glu35 is reprotonated. Another reaction mechanism, proposed by Koshland,[32] involves a glycosyl-LYZ intermediate in which the substrate is covalently bonded to Asp52. Vocadlo *et al.*[33] concluded from electrospray-ionization mass-spectrometric studies that the catalytic mechanism of LYZ involves an intermediate complex in which one sugar ring is covalently bonded to Asp52. However, the crystallized structure of a trisaccharide bound to the active-site of LYZ did not present such covalent intermediate, while the presence of hydrogen bonds in the environment of Glu35 and Asp52 supported the Phillips mechanism.[34] Inspection of the LYZ crystal structure[17] shows that the carboxylate side group of Glu35 interacts with the following residues: Gln57 (N_{bb} / 1.76 Å), Val109 (N_{bb} / 1.82 Å) and Ala110 (N_{bb} / 1.84 Å), where N_{bb} refers to the nitrogen atom from the backbone and the distance corresponds to the N–H···O interaction. Besides, the carboxylate of the Asp52 residue interacts with Arg45 (N_{sc} / 1.81 Å; N_{sc} / 2.10 Å; double salt bridge), Asn46 (N_{bb} / 1.91 Å), Thr47 (N_{bb} / 2.09 Å) and Asn59 (N_{sc} / 1.82 Å), where N_{sc} denotes the nitrogen atom from the side chain. In addition, the backbone N–H and C=O groups of Glu35 and Asp52 also participate in specific hydrogen bonding interactions: (Glu35)N–H···O=C(Ala31 / 1.96 Å), (Glu35)C=O···H–N_{sc}(Gln57 / 2.22 Å), (Asp52)N–H···O_{sc}(Thr43 / 2.04 Å) and (Asp52)C=O···H–

O_{sc}(Thr43 / 1.62 Å). These hydrogen bonds and salt bridges, which are responsible for the antimicrobial activity of LYZ are depicted in Figure 8.5.

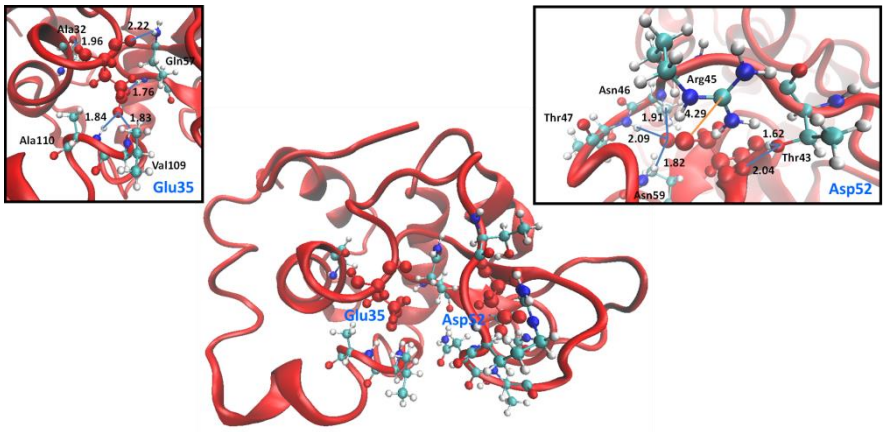


Figure 8.5. For crystallized LYZ (1LKS entry in the PDB): Details about the interactions formed by the Glu35 and Asp52 (represented by balls in red), which are responsible of the antimicrobial activity of LYZ. Distances (in Å) correspond to specific inter-residue interactions. Magnification of the Glu35- and Asp52-containing regions is displayed (left and right, respectively).

In absence of electric field, the RMSF of Glu35 and Asp52 in PEDOT/LYZ are relatively low: 0.89 and 1.65 Å, respectively (Figure 8.2b). However, in this system PEDOT···LYZ interactions cause drastic distortions in surrounding regions that are crucial for the activity of the protein. Figure 8.2c shows that Glu35 only preserves the side chain···backbone interaction with Gln57 (1.82 Å) and the backbone···backbone hydrogen bond with Ala31 (2.10 Å), which is associated to the α-helix secondary structure. This reduction is even more drastic for Asp52 since all the interactions mentioned above disappear upon adsorption onto PEDOT surface.

Instead two new N–H···O interactions with the backbone of Asn44 are formed due to the apparition of a β-sheet. Figure 8.2b shows the very high RMSF values of Arg45 (3.72 Å), Asn46 (5.55 Å) and Asp47 (4.61 Å), which explains the drastic change in the interaction pattern of the active site with respect to the crystal.

Independently of its strength and sign, the RMSFs decreases in presence of $\overrightarrow{E_z}$ (Figure 8.2b), which is consistent with the RMSD behaviour. In spite of this, the greater distortions and, by extension, the lost inter-residue interactions involving Glu35 and Asp52 are the same in absence and presence of electric field. This is illustrated in Figure 8.6 for $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$). Overall, MD simulations on PEDOT/LYZ indicate that the loss of antimicrobial activity is due to the distortions induced by the PEDOT···LYZ interactions formed upon adsorption rather than to the applied electric field. Although the alterations caused by these interactions are not very drastic, as is reflected by the superposition of the crystal structure and the last snapshot from the simulation (Figure 8.7), they affect the side groups of key polar residues, altering significantly the interaction pattern with respect to the crystal (Figures 8.5 and 8.2c).

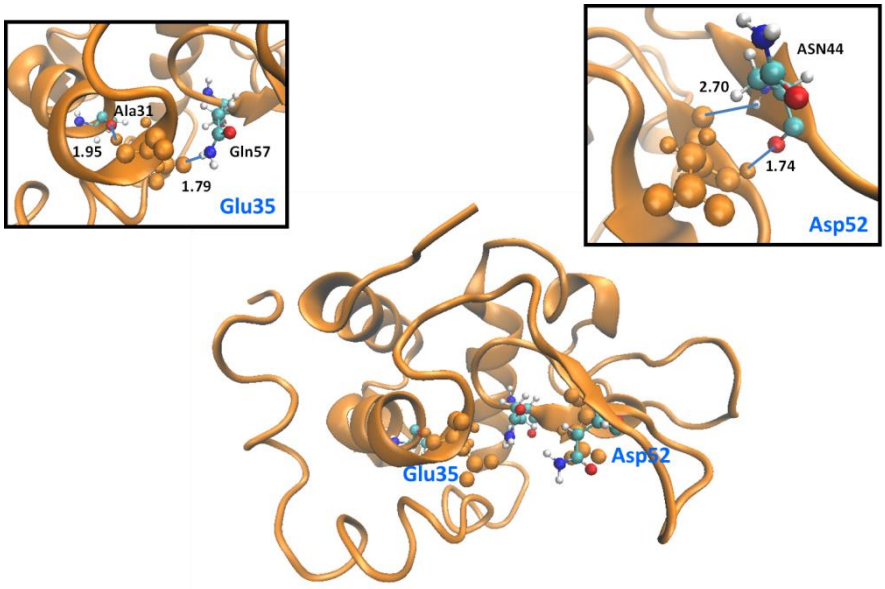


Figure 8.6. For PEDOT/LYZ with $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$): Details about the interactions formed by the Glu35 and Asp52 residues (represented by balls in light brown). Distances (in Å) correspond to specific inter-residue interactions. Magnification of the Glu35- and Asp52-containing regions is displayed (left and right, respectively).

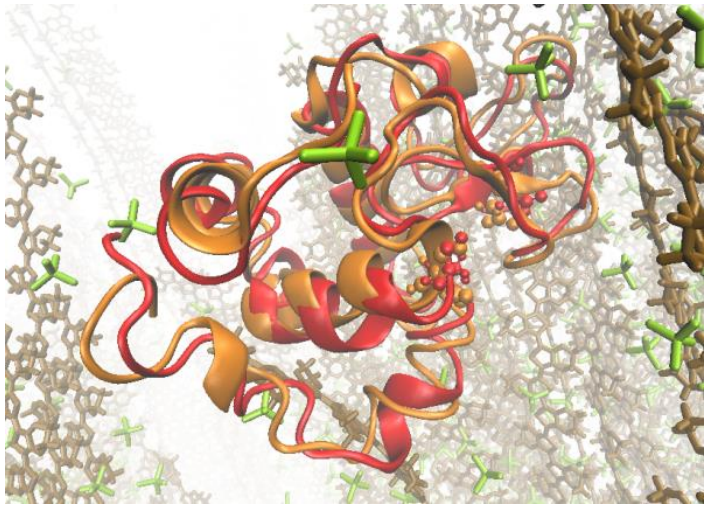


Figure 8.7. Superposition of the LYZ crystal structure (in red) and the last snapshot from PEDOT/LYZ MD simulation in absence of electric field (in light

brown). Residues Glu35 and Asp52, which play a crucial role in the antimicrobial activity of LYZ have been represented by balls.

Regarding to P(EDOT-LYZ), many of the interactions involving Glu35 and Asp52 remain practically unaltered, as it is shown in Figure 8.3c. More specifically, the Glu35 maintains the side chain···backbone interactions with Val109 (1.82 Å) and Ala110 (2.35 Å) and the backbone···backbone interaction with Ala31 (1.95 Å), losing the two interactions with Gln57. The Asp52 preserves the side chain···backbone interaction with Asn46 (2.24 Å), the side chain···side chain with Asn59 (this becomes stronger, involving the two oxygen atoms of the carboxylate moiety: 2.41 and 2.20 Å), and one of the backbone···side chain interactions with Thr43 (1.87 Å). Thus, Asp52 loses the interactions with Arg45 and Thr47, even though new interactions with Asn44 (Figure 8.3c), similar to those found for PEDOT/LYZ, are detected. Moreover, many of the interactions found in the crystallized active centre are also detected in P(EDOT-LYZ) when the $\overrightarrow{E_z}$ is applied, independently of the strength and sign. This is illustrated in Figure 8.8 for $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$), which shows the interactions of Glu35 with Ala31 (1.84 Å), Val109 (2.03 Å) and Ala110 (2.04 Å) and of Asp52 with Thr43 (1.83 Å), Asn44 (2.02 and 2.04 Å), Asn46 (2.49 Å) and Thr47 (1.55 and 1.90 Å).

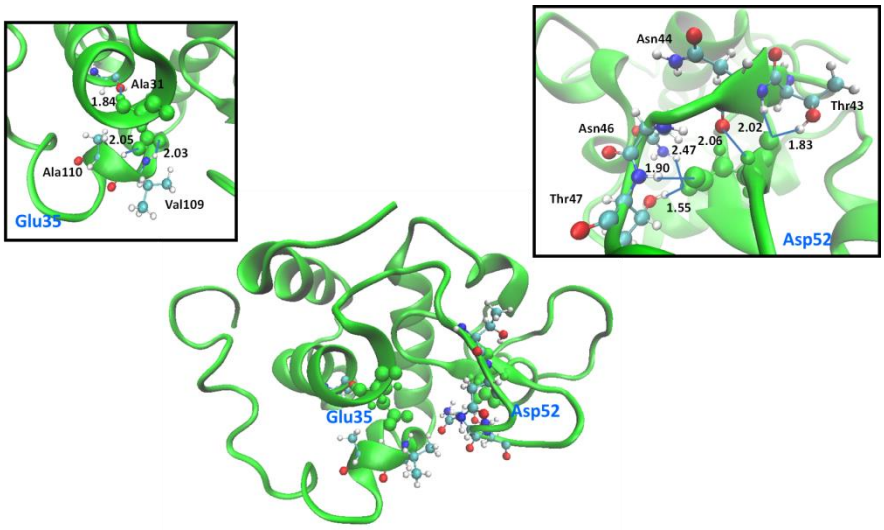


Figure 8.8. For P(EDOT-LYZ) with $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$): Details about the interactions formed by the Glu35 and Asp52 residues (represented by balls in green). Distances (in Å) correspond to specific inter-residue interactions. Magnification of the Glu35- and Asp52-containing regions is displayed (left and right, respectively).

Figure 8.9 summarizes the comparison among the side chain⋯backbone, backbone⋯side chain, side chain⋯side chain and backbone⋯backbone interactions involving Glu35 and Asp52 in 1LKS (crystal), PEDOT/LYZ and P(EDOT-LYZ), both simulations in absence of electric field and applying $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$) being considered for the latter two. Unfortunately, from these results it is not possible to ascertain which residue, Glu35 or Asp52, plays the most important role in the antimicrobial activity of LYZ since almost all interactions are lost in PEDOT/LYZ.

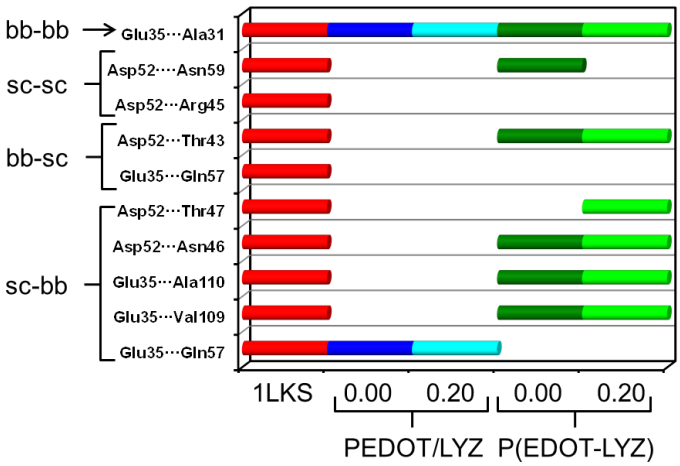


Figure 8.9. Comparison among the side chain···backbone (sc-bb), backbone···side chain (bb-sc), side chain···side chain (sc-sc) and backbone···backbone (bb-bb) interactions involving Glu35 and Asp52 in 1LKS (crystal structure), PEDOT/LYZ and P(EDOT-LYZ). Observed interactions are represented by coloured bars. Results from simulations in absence of electric field and applying $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$) are considered for the two composites with PEDOT.

Overall, results indicate that the anisotropy of PEDOT···LYZ interactions in PEDOT/LYZ are responsible for the loss of antimicrobial activity experimentally observed.[2] Although such anisotropic interactions does not cause drastic changes in the conformation of LYZ, as is reflected by the RMSD and RMSF profiles displayed in Figures 8.2a and 8.2b, respectively, they are large enough to affect the local interactions that preserves the stability of the active centre. This is clearly evidenced in Figure 8.2d, which represents the variation of the distance between the Cα-atoms of Glu35 and Asp52 ($d_{\text{Glu35-Asp52}}$). The $d_{\text{Glu35-Asp52}}$ values averaged over the last 20 ns of each simulation are 11.69±0.36,

11.00±0.35, 11.38±0.38, 10.84±0.22, and 11.09±0.47 Å for $\overrightarrow{E_z}$= 0.00, 0.000.2, 0.20, -0.0020 and -0.020 kcal/(mol·Å·e$^-$), respectively. These values are approximately 15% larger than the one observed in the crystallized structure (*i.e.* 9.66 Å). Figure 8.2e depicts the temporal evolution of the distance between Lys1 and Leu129 ($d_{\text{Lys1-Leu129}}$), which can be interpreted as the breathing movements of the LYZ protein during the trajectory. As it can be seen, fluctuations are very small, suggesting that the polymer stretches the part of the protein in contact with the surface. This avoids the small movements associated with the thermal effects and cause local distortions. The average $d_{\text{Lys1-Leu129}}$ values (last 20 ns of each simulation), which are 23.55±0.52, 23.89±0.44, 24.58±0.51, 23.61±0.41, and 23.59±0.64 Å for $\overrightarrow{E_z}$= 0.00, 0.0002, 0.20, -0.0002 and -0.20 kcal/(mol·Å·e$^-$), respectively, are greater than the crystallographic value (*i.e.* 21.07 Å). This overestimation (~13%) is fully consistent with the features discussed above.

A completely different situation is found for P(EDOT-LYZ), in which PEDOT⋯LYZ interactions are much more isotropic. In this case, the protein is kept enclosed in a cavity, which acts as a mold. The temporal evolution of $d_{\text{Glu35-Asp52}}$ (Figure 8.3d) indicates values very close the crystallographic one, even when the electric field is applied. Specifically, the average $d_{\text{Glu35-Asp52}}$ values (last 20 ns) of each simulation are 10.31±0.40, 9.87±0.34, 10.07±0.39, 10.11±0.39, and 11.07±0.34 Å, or $\overrightarrow{E_z}$= 0.00, 0.002, 0.20, -0.002 and -0.20 kcal/(mol·Å·e$^-$), respectively, which differ by less than 6.7% with respect to the crystallographic value. On the other hand, this cage-like effect observed for P(EDOT-LYZ) does not affect the

breathing movement, fundamentally associated with the thermal energy, as is displayed in Figure 8.3e. Thus, $d_{\text{Lys1-Leu129}}$ experiences occasional deviations, recovering after a few ns values close to the average value. This behaviour should be attributed to the fact that cavity that contains the protein is large enough to prevent compressing affects but small enough to avoid structural distortions.

$\overrightarrow{E_z}$ [kcal/(mol·Å·e$^-$)]	R_g / $R_{g/20}$ (Å)	R_H / $R_{H/20}$ (Å)
	PEDOT/LYZ	
0.00	15.2±0.1 / 15.2 ± 0.1	19.9±0.2 / 19.8 ± 0.2
0.0002	15.3±0.1 / 15.3 ± 0.1	20.2±0.2 / 20.1 ± 0.1
0.20	15.3±0.1 / 15.3 ± 0.1	20.2±0.2 / 20.2 ± 0.2
-0.0002	15.4±0.1 / 15.3 ± 0.1	20.3±0.1 / 20.3 ± 0.1
-0.20	15.4±0.1 / 15.4 ± 0.1	20.3±0.2 / 20.3 ± 0.1
	P(EDOT-LYZ)	
0.00	15.1±0.1 / 15.2 ± 0.1	19.8±0.1 / 19.8 ± 0.1
0.0002	15.1±0.1 / 15.1 ± 0.1	19.8±0.1 / 19.8 ± 0.1
0.20	15.0±0.1 / 15.0 ± 0.1	19.7±0.1 / 19.6 ± 0.1
-0.0002	15.1±0.1 / 15.1 ± 0.1	19.7±0.1 / 19.7 ± 0.1
-0.20	15.1±0.1 / 15.1 ± 0.1	19.8±0.1 / 19.8 ± 0.1
	Crystal (1LKS)	
-	14.7	19.1

Table 8.1. Average values of the radius of gyration (R_g) and the hydrodynamic radius (R_H) for PEDOT/LYZ and P(EDOT-LYZ) in absence and presence of electric field. For each case, the average values and the corresponding standard deviations were calculated considering the whole production trajectory (R_g and R_H values) and the last 20 ns of the trajectory ($R_{g/20}$ and $R_{H/20}$ values). The values obtained considering the coordinates of the crystallized structure (1LKS entry in the PDB) are displayed in the last row.

These conclusions are supported by relevant structural properties such as R_g and R_H, which have been calculated to be 14.7 Å and 19.1 Å, respectively, for the crystallized protein. The values averaged from PEDOT/LYZ and P(EDOT-LYZ) simulations in absence of electric field are [R_g= 15.2±0.1 Å; R_H= 19.9±0.2 Å] and [R_g= 15.1±0.1 Å; R_H= 19.8±0.2 Å], respectively. These values represent an increment with respect to the crystal of 3%-4% only, supporting that the CP does not induce severe structural distortions in the protein. Moreover, the changes caused by the application of the electric field are very small, independently of the strength and sign (Table 8.1).

Analysis of the secondary motifs in the crystal indicates that the content of β-strands and helices amounts to βS= 3.1% and H= 37.2%, whereas loops- and turns-containing unfolded regions reach the highest population L+T= 59.7%. Figure 8.10 represents the temporal evolution and average value of the population of βS, H and L+T motifs for PEDOT/LYZ and P(EDOT-LYZ) in absence of electric field and applying $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$), which has been chosen to illustrate the effect of the electric field (*i.e.* trajectories with $\overrightarrow{E_z}$= 0.002, -0.002 and -0.20 kcal/(mol·Å·e$^-$) provide similar results). Interestingly, in absence of electric field both the adsorption and the encapsulation of LYZ (Figures 8.10a and 8.10b, respectively) result in a significant increment of β-strands, which is possible due to a reduction in the population of unstructured regions, while helices remain approximately at the same value (*i.e.* differences were 1.0-1.2% only). However, these tendencies are more pronounced for PEDOT/LYZ (βS=

11.0%±0.9% and L+T= 50.6%±2.3%) than for P(EDOT/LYZ) (βS= 8.1%±0.9% and L+T= 53.7%±2.2%). Application of $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$) to the system with the adsorbed protein is in detriment of the βS and H populations (Figure 8.10c), which increase with respect to the crystal, causing a reduction in the percentage of L+T that is higher than that obtained in absence of electric field. Conversely, the electric field does not produce any significant change in the distribution of secondary motifs of P(EDOT-LYZ) (Figure 8.10d), the populations of secondary motifs being similar to those achieved in absence of electric field.

On the basis of these results and the previously reported experimental observations,[2] the mechanisms that regulates the antimicrobial activity of P(EDOT-LYZ) is proposed to be based on the dynamics of the PEDOT chains. Thus, CPs are dynamic systems in which dopant ions are continuously exchanged with the medium (*i.e.* access of ions into the polymer matrix and escape of ions from the polymer matrix). This exchange is accompanied by relatively slow structural re-arrangements and local relaxations of the polymer chains, which probably facilitate the slow diffusion and subsequent release the LYZ molecules towards the medium.

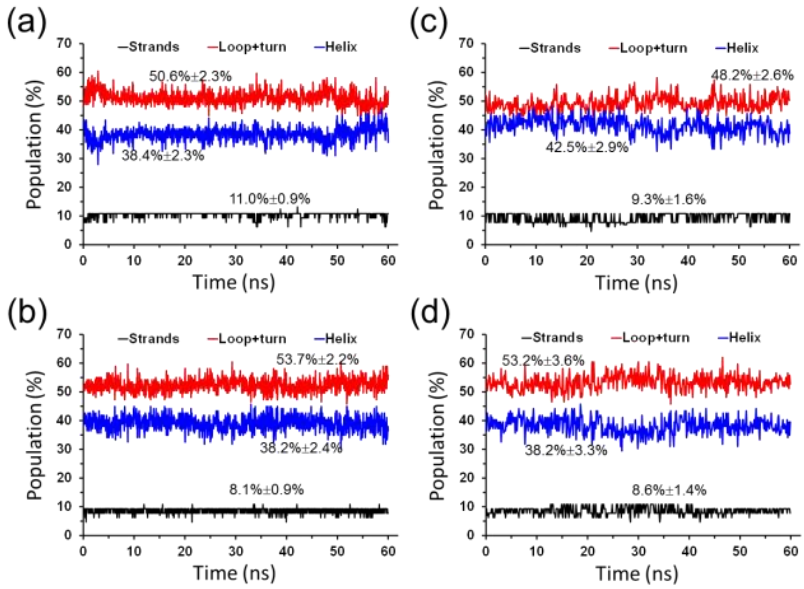


Figure 8.10. Temporal evolution of the population of β-strands (βS), helices (H), and loops and turns (L+T) as derived from (a,c) PEDOT/LYZ and (b,d) P(EDOT/LYZ) MD trajectories in (a,b) absence of electric field and (b,d) applying $\overrightarrow{E_z}$= 0.20 kcal/(mol·Å·e$^-$). The L+T motif is mainly related with unstructured regions of the protein. Average values with the corresponding standard deviations, which were calculated considering the whole trajectories, are also displayed.

8.4.2 PEDOT···LYZ interactions

The interactions between the LYZ protein and PEDOT chains have been analysed by calculating the radial distribution functions (RDFs) between pairs of atoms. Figure 8.11a, which represents the RDFs calculated between all atoms of LYZ and all atoms of PEDOT chains atom ($RDF_{LYZ\text{-}PEDOT}$), reflect the presence of van der Waals contacts in both PEDOT/LYZ and P(EDOT-LYZ). In both cases the profiles start at $r \approx$ 2.3 Å, which is shorter to the sum of the van der Waals radii of hydrogen atoms, and increasing with *r*. This

feature indicates the existence of non-specific electrostatic and van der Waals PEDOT···LYZ interactions. As it is expected, the growing tendency is much more pronounced for the protein embedded into the polymer matrix than for the protein adsorbed onto the surface of PEDOT.

Specific N–H···O, N–H···S, O–H··· O and O–H···O hydrogen bonding interactions were examined through the corresponding LYZ(N or O)···(O or S)PEDOT pairs. Obviously, the O atoms belonging to the backbone C=O from LYZ were excluded from the analysis since they cannot act as hydrogen bond donor. Figures 8.11b and 8.11c display the $RDFs_{LYZ(N)\text{-}(O)PEDOT}$ and $RDFs_{LYZ(N)\text{-}(S)PEDOT}$ profiles, respectively, computed for PEDOT/LYZ and P(EDOT-LYZ). A peak centred at $r\approx$ 4.6 Å is obtained for the $RDF_{LYZ(N)\text{-}(O)PEDOT}$ of the two systems, even though this is significantly higher for PEDOT/LYZ than for P(EDOT-LYZ). The opposite tendency is shown for the first peak of the $RDF_{LYZ(N)\text{-}(S)PEDOT}$ at $r\approx$ 3.8 Å. These features confirm the existence of N–H···O and N–H···S hydrogen bonds in both the adsorbed and encapsulated systems. However, the former hydrogen bonds are more abundant in PEDOT/LYZ than in P(EDOT-LYZ), while the latter interaction is more abundant in P(EDOT-LYZ) than in PEDOT/LYZ. Interestingly, N–H···S hydrogen bonds are stronger than N–H···O ones, as is reflected by the *r* values of the first peaks. This has been attributed to the flexibility of the dioxane ring, which is used to accommodate the protein onto the surface or inside the cavity of the CP film. Thus, the maximization of non-specific electrostatic interactions is in detriment of the strength of N–H···O

hydrogen bonds. In contrast, the strength of N–H···S hydrogen bonds is imposed by the rigidity of the thiophene rings.

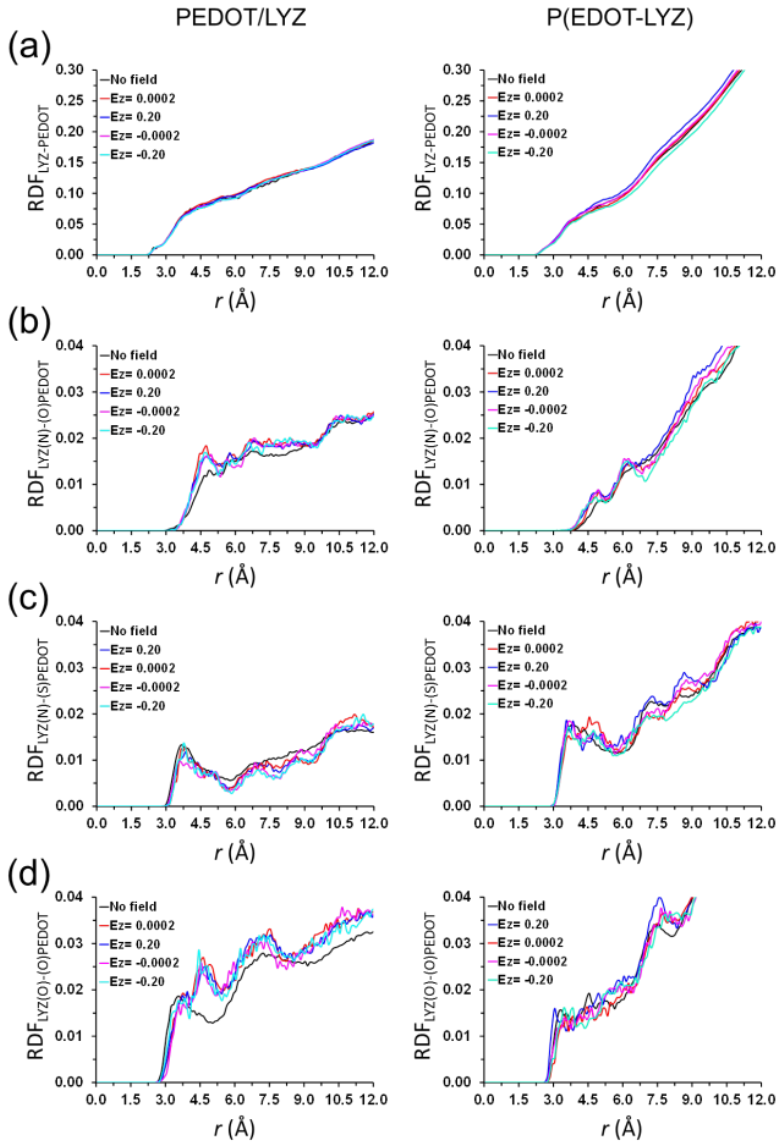


Figure 8.11. Radial distribution functions (RDF) for the following pairs of atoms: a) all atoms of LYZ and all atoms of PEDOT chains atom ($RDF_{LYZ\text{-}PEDOT}$) to identify non-specific electrostatic and van der Waals interactions; b) N(–H) atoms

of LYZ and O atoms of PEDOT ($RDF_{LYZ(N)-(O)PEDOT}$) to identify N–H···O hydrogen bonds; c) N(–H) atoms of LYZ and S atoms of PEDOT ($RDF_{LYZ(N)-(S)PEDOT}$) to identify N–H···S hydrogen bonds; and d) O atoms (excluding those belonging to the C=O groups of the backbone) of LYZ and O atoms of PEDOT ($RDF_{LYZ(O)-(O)PEDOT}$) to identify O–H···O hydrogen bonds. In all cases the profiles calculated for PEDOT/LYZ and P(EDOT-LYZ) are displayed at the left and right, respectively.

Inspection of the $RDF_{LYZ(O)-(O)PEDOT}$ profiles (Figure 8.11d) reveals that the characteristics of O–H···O hydrogen bonds are similar to those discussed for the N–H···O ones. Finally, $RDF_{LYZ(O)-(S)PEDOT}$ profiles (not shown, respectively) evidence that O–H···S interactions are practically negligible, no well-defined peak is detected. In general, results displayed in Figure 8.11 indicate that the electric field has a minor role in the population and strength of hydrogen bonds, causing only small changes.

8.4.3 Perchlorate anions: Interactions and diffusion

Finally, the strength and abundance of ClO_4^-···PEDOT and ClO_4^-···LYZ interactions were analysed by comparing the RDFs involving the following pairs of atoms: 1) the chlorine atom of ClO_4^- and the sulphur atoms of PEDOT chains ($RDF_{Cl-(S)PEDOT}$); 2) the chlorine atom of ClO_4^- and the backbone α-carbon atoms of LYZ ($RDF_{Cl-(C\alpha)LYZ}$). Results, which are displayed in Figures 8.12a-b, indicate that the peaks are much higher for $RDF_{Cl-(S)PEDOT}$ than for $RDF_{Cl-(C\alpha)LYZ}$, independently of both the encapsulated or adsorbed configuration of the protein and the presence or absence of electric field. This feature is consistent with the fact that ClO_4^- ions

remain close to the PEDOT chains and only a few anions interact directly with the protein. Deeper analyses evidenced that such few ClO_4^-···LYZ interactions occur with the nitrogen atoms of Arg and Lys charged residues, as is shown by the RDFs calculated considering pairs involving the chlorine atom of ClO_4^- and the nitrogen at the side chains of such residues ($RDF_{Cl\text{-}(Nsc)LYZ}$ in Figure 8.12c). Moreover, these strong electrostatic interactions are more abundant in P(EDOT-LYZ) than in PEDOT/LYZ, as expected because of the encapsulation of the protein in CP matrix.

The diffusion coefficients (*D*) of the ClO_4^- ions were determined using the Einstein relation once the diffusive regime was reached. For this purpose, the mean square displacement as a function of time of ClO_4^- ions were calculated. The values of *D* calculated for PEDOT/LYZ and P(EDOT-LYZ) are represented in Figure 8.13 as a function of $\overrightarrow{E_z}$. In absence of electric field, the values of *D* were similar for the two systems. Analysis of their three components (*i.e.* D_x, D_y and D_z) indicates that D_z (~$72 \cdot 10^{-5}$ cm/s) is larger than D_x and D_y, which are very similar (~$47 \cdot 10^{-5}$ cm/s). This feature indicates that, as it was expected, the diffusion of ClO_4^- anions is higher in the direction perpendicular to the steel substrate than in the parallel one. The anisotropy in the diffusive behaviour of the dopant anions is guided by the orientation of the positively charged PEDOT chains, which grow perpendicular onto the iron substrate.

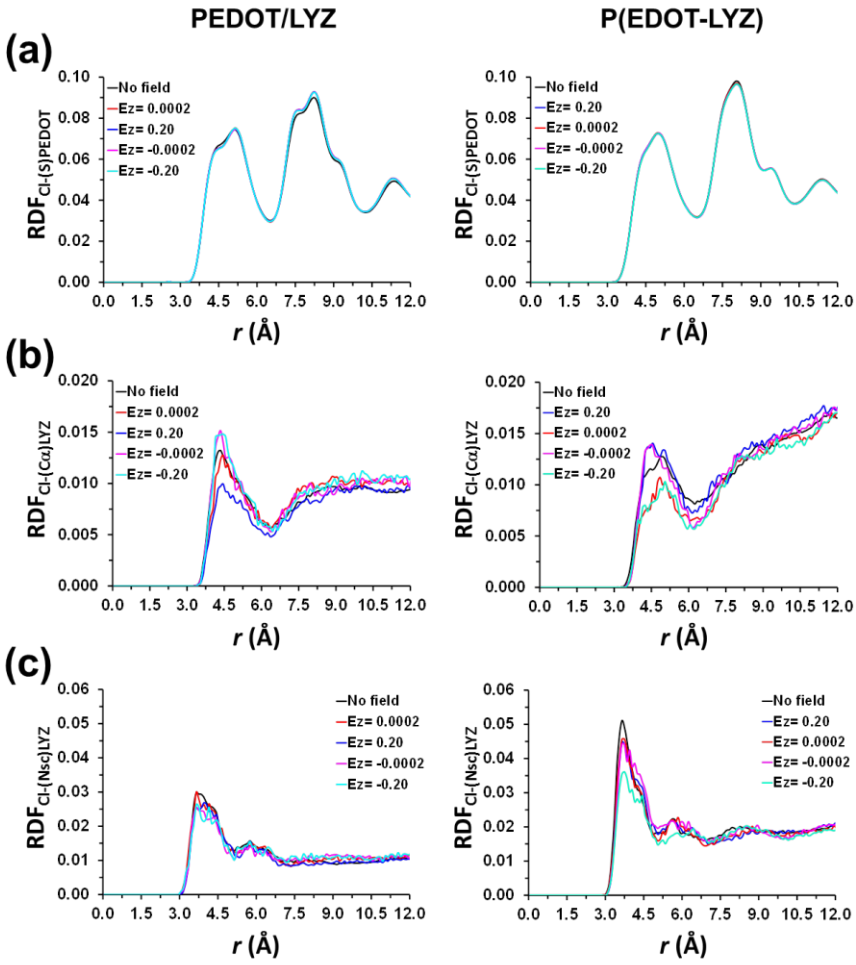


Figure 8.12. Radial distribution functions (RDF) for the following pairs of atoms: a) chlorine atom of ClO_4^- and the sulphur atom of PEDOT chains ($RDF_{Cl\text{-}(S)PEDOT}$); b) chlorine of ClO_4^- and the C^{α}-LYZ pairs of atoms ($RDF_{Cl\text{-}(C\alpha)LYZ}$); and c) chlorine of ClO_4^- and the nitrogen at the side chains LYZ pairs of atoms ($RDF_{Cl\text{-}(Nsc)LYZ}$). In all cases the profiles calculated for PEDOT/LYZ and P(EDOT-LYZ) are displayed at the left and right, respectively.

Application of $\overrightarrow{E_z}$.results in an enhancement of the diffusion coefficient, this effect being considerably more pronounced in the

direction of the applied electric field than in the perpendicular direction. Moreover, such tendency increases with the strength of $\overrightarrow{E_z}$, especially for PEDOT/LYZ, while it is independent of sign of the field. Accordingly, the Arg and Lys residues of the protein immersed in the cavity inside the polymer matrix, remarkably shield the effects of the electric field in the diffusion of the dopant anions. Thus, the electric field is not high enough to disrupt the strong electrostatic interactions between the positive charged side groups and the negatively charged counter-anions. This feature is consistent with the fact that $RDF_{Cl\text{-}(Nsc)LYZ}$ is higher for P(EDOT-LYZ) than for PEDOT/LYZ (Figure 8.12c).

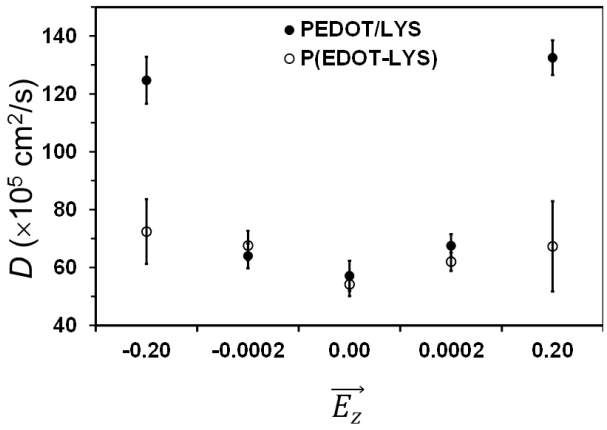


Figure 8.13. Diffusion coefficients of ClO_4^- anions in PEDOT/LYZ and P(EDOT-LYZ) in absence of electric field and applying selected external electric fields ($\overrightarrow{E_z}$, in kcal/(mol·Å·e$^-$)).

8.5 CONCLUSIONS

Our MD study gives support to the experimental observation that biocapacitors in which the LYZ is adsorbed onto the surface of PEDOT films do not present antimicrobial activity. Our results show that oxidized PEDOT chains, which grow onto a steel substrate and exhibit a total charge of +0.5 per EDOT repeat unit, interact with the numerous charged residues of the protein (20.1%). These interactions produce small structural alterations at the protein, affecting the residues at the active centre and breaking the key interactions for the bactericidal activity of LYZ. Unexpectedly, this behaviour is independent of the presence of electric field, even when this is strong. Therefore, the loss of biological activity has been exclusively attributed to the anisotropy of the PEDOT···LYZ electrostatic interactions, which is a phenomenon induced by the alignment of CP chains in the direction perpendicular to the metallic substrate. Conversely, biocapacitors prepared by encapsulating the protein inside the PEDOT matrix retain the most important inter-residue interactions at the LYZ active centre, preserving the antimicrobial activity. In this case, the anisotropic effects are minimized since the surrounding environment exerts forces of similar strength around the whole protein. Consequently, the formed electrostatic PEDOT···LYZ interactions are isotropic or quasi-isotropic and structural distortions are minimized.

The great advantage of the MD simulation technique is the wealth of structural information provided with atomistic detail. Although, this technique has a severe limitation in its high computational cost, which is particularly evident in simulations

conducted in presence of electric field, realizable results can be obtained by exploring conditions adapted to drastically accelerate the investigated physical phenomena. For example, the length of the trajectories conducted using $\overrightarrow{E_z}$= 0.20 and -0.20 kcal/(mol·Å·e$^-$) is enough to conclude unmistakably that the bactericidal activity of PEDOT/LYZ and P(EDOT-LYZ) biocapacitors is independent of the electric field applied or supplied during charge-discharge processes.

8.6 REFERENCES

1. Sode, K., Yamazaki, T., Lee, I., Hanashi, T. & Tsugawa, W. BioCapacitor: A novel principle for biosensors. *Biosens. Bioelectron.* **76,** 20–28 (2016).
2. López-Pérez, D. E., Aradilla, D., Del Valle, L. J. & Alemán, C. Capacitive composites made of conducting polymer and lysozyme: Toward the biocondenser. *J. Phys. Chem. C* **117,** 6607–6619 (2013).
3. Pankratov, D., Blum, Z., Suyatin, D. B., Popov, V. O. & Shleev, S. Self-Charging Electrochemical Biocapacitor. *ChemElectroChem* **1,** 343–346 (2014).
4. Lv, Z. *et al.* Microbial fuel cell as a biocapacitor by using pseudo-capacitive anode materials. *J. Power Sources* **246,** 642–649 (2014).
5. Teixeira-Dias, B. *et al.* Dextrin- and conducting-polymer-containing biocomposites: Properties and behavior as cellular matrix. *Macromol. Mater. Eng.* **297,** 359–368 (2012).
6. Hanashi, T. *et al.* The development of an autonomous self-powered bio-sensing actuator. *Sensors Actuators, B Chem.* **196,** 429–433 (2014).
7. Derde, M. *et al.* Hen Egg White Lysozyme Permeabilizes Escherichia coli Outer and Inner Membranes. *J. Agric. Food Chem.* **61,** 9922–9929 (2013).
8. Derde, M. *et al.* Dry-heating of lysozyme increases its activity against Escherichia coli membranes. *J. Agric. Food Chem.* **62,** 1692–1700 (2014).

9. Liburdi, K., Benucci, I., Palumbo, F. & Esti, M. Lysozyme immobilized on chitosan beads: Kinetic characterization and antimicrobial activity in white wines. *Food Control* **63,** 46–52 (2016).
10. Uygun, M., Uygun, D. A., Altunbaş, C., Akgöl, S. & Denizli, A. Dye Attached Nanoparticles for Lysozyme Adsorption. *Sep. Sci. Technol.* **49,** 1270–1278 (2014).
11. Park, S. I., Daeschel, M. A. & Zhao, Y. Functional properties of antimicrobial lysozyme-chitosan composite films. *J. Food Sci.* **69,** M215–M221 (2004).
12. Bayramoglu, G. *et al.* Lysozyme specific aptamer immobilized MCM-41 silicate for single-step purification and quartz crystal microbalance (QCM)-based determination of lysozyme from chicken egg white. *Microporous Mesoporous Mater.* **207,** 95–104 (2015).
13. Cavalieri, F. *et al.* Antimicrobial and biosensing ultrasound-responsive lysozyme-shelled microbubbles. *ACS Appl. Mater. Interfaces* **5,** 464–471 (2013).
14. Mahalingam, S., Xu, Z. & Edirisinghe, M. Antibacterial Activity and Biosensing of PVA-Lysozyme Microbubbles Formed by Pressurized Gyration. *Langmuir* **31,** 9771–9780 (2015).
15. Liao, A. H., Hung, C. R., Lin, C. F., Lin, Y. C. & Chen, H. K. Treatment effects of lysozyme-shelled microbubbles and ultrasound in inflammatory skin disease. *Sci. Rep.* **7,** 41325 (2017).
16. Weiss, M. S., Palm, G. J. & Hilgenfeld, R. Crystallization, structure solution and refinement of hen egg-white lysozyme at pH 8.0 in the presence of MPD. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56,** 952–958 (2000).
17. Steinrauf, L. K. Structures of monoclinic lysozyme iodide at 1.6 Å and of triclinic lysozyme nitrate at 1.1 Å. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 767–779 (1998).
18. Lei, C., Wilson, P. & Lekakou, C. Effect of poly(3,4-ethylenedioxythiophene) (PEDOT) in carbon-based composite electrodes for electrochemical supercapacitors. *J. Power Sources* **196,** 7823–7827 (2011).
19. Ryu, K. S. *et al.* Poly(ethylenedioxythiophene) (PEDOT) as polymer electrode in redox supercapacitor. *Electrochim. Acta* **50,** 843–847 (2004).
20. Groenendaal, L., Jonas, F., Freitag, D., Pielartzik, H. &

Reynolds, J. R. Poly(3,4-ethylenedioxythiophene) and its derivatives: past, present, and future. *Adv. Mater.* **12,** 481–494 (2000).

21. Aradilla, D., Estrany, F. & Alemán, C. Symmetric supercapacitors based on multilayers of conducting polymers. *J. Phys. Chem. C* **115,** 8430–8438 (2011).
22. Wei, Q., Mukaida, M., Kirihara, K., Naitoh, Y. & Ishida, T. Recent progress on PEDOT-based thermoelectric materials. *Materials (Basel).* **8,** 732–750 (2015).
23. Poater, J., Casanovas, J., Solà, M. & Alemán, C. Examining the planarity of poly(3,4-ethylenedioxythiophene): Consideration of self-rigidification, electronic, and geometric effects. *J. Phys. Chem. A* **114,** 1023–1028 (2010).
24. Zanuy, D. & Alemán, C. Resolving the subnanometric structure of ultrathin films of poly(3,4-ethylenedioxythiophene) on steel surfaces: A molecular modeling approach. *Soft Matter* **9,** 11634–11644 (2013).
25. M.McGuire. *Stainless Steels For Design Engineers*. *US Patent 3,306,736* (Asm International, 2008). doi:10.1361/ssde2008p091
26. Ocampo, C., Oliver, R., Armelin, E., Alemán, C. & Estrany, F. Electrochemical synthesis of poly(3,4-ethylenedioxythiophene) on steel electrodes: Properties and characterization. *J. Polym. Res.* **13,** 193–200 (2006).
27. Curcó, D., Zanuy, D. & Alemán, C. EVEBAT: A fast strategy for the examination of the empty space in polymer matrices. *J. Comput. Chem.* **24,** 1208–1214 (2003).
28. Jensen, M. O. *et al.* Principles of conduction and hydrophobic gating in K+ channels. *Proc. Natl. Acad. Sci.* **107,** 5833–5838 (2010).
29. Wang, S., Tang, Y., Schobert, H. H., Guo, Y. & Su, Y. FTIR and 13C NMR investigation of coal component of late permian coals from southern China. *Energy and Fuels* **25,** 5672–5677 (2011).
30. Khalili-Araghi, F. *et al.* Calculation of the gating charge for the Kv1.2 voltage-activated potassium channel. *Biophys. J.* **98,** 2189–2198 (2010).
31. Phillips, D. C. The three-dimensional structure of an enzyme molecule. *Scientific American* **215,** 78–90 (1966).
32. KOSHLAND, D. E. Stereochemistry and the Mechanism of Enzymatic Reactions. *Biol. Rev.* **28,** 416–436 (1953).

33. Vocadlo, D. J., Davies, G. J., Laine, R. & Withers, S. G. Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature* **412,** 835–838 (2001).
34. Strynadka, N. C. J. & James, M. N. G. Lysozyme revisited: Crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. *J. Mol. Biol.* **220,** 401–424 (1991).

- Chapter 9 -

Conclusions

9 CONCLUSIONS

9.1 General conclusions

1) The conformational preferences of small peptides, as for example the $RG^{E}D$ sequence (where G^{E} refers to an engineered amino acid bearing a EDOT ring as side group), can be successfully determined combining precise DFT calculations with a conformational search strategies.
2) The electronic properties of medium size compounds, as for example small tripeptides, can be predicted by applying sophisticated TD-DFT calculations.
3) Conventional classical MD simulations are very efficient to examine the conformational preferences of small peptides and medium size adsorbed peptide-polymer conjugates. Furthermore, this methodology allows to distinguish the effect of the facet on the conformational preferences as well as the influence of physical processes, such as the desolvation of the assembled peptide-polymer conjugate.
4) Conventional classical MD simulations are very efficient to examine the conformational preferences of small peptides and medium size adsorbed peptide-polymer conjugates. Furthermore, this methodology allows to distinguish the effect of the facet on the conformational preferences as well as the influence of physical processes, such as the desolvation of the assembled peptide-polymer conjugate.
5) Classical MD simulations of a representative OMP in several surrounding environments have provided important insights about the effects of both water and detergent

molecules in the structural stability and self-assembly capacity of such porin. Theoretical results have been extremely useful for the design of new generation of bioinspired protein-integrated devices.

6) An effective modelling strategy, which combines stochastic algorithms and classical MD simulations, has been developed to represent the atomistic details of the bulk organization in PEDOT depositions over metal surfaces. The developed synergic strategy compiles the experimental information of the topographic features of PEDOT deposited on stainless steel and makes possible to build atomistic models that reproduce the observed experimental features.
7) The modelling strategy developed to reproduce the bulk organization in PEDOT depositions has been successfully used to simulate P(EDOT-LYZ) and PEDOT/LYZ biocapacitors. Results obtained for these complex systems have allowed to rationalize their antimicrobial activity experimentally observed in terms of PEDOT⋯LYZ interactions..

9.2 Specific conclusions

1) DFT calculations indicate that the bioactive characteristics of the RGD sequence become unstable in $RG^{E}D$ because of both the steric hindrance caused by the EDOT side group and the repulsive interactions between the oxygen atoms belonging to the backbone amide groups and the EDOT side group.

2) The conformational preferences identified for the $RG^{E}D$ sequence have been used to propose some chemical changes oriented to improve the bioadhesive properties. These are: *(i)* reverse the chirality of the GlE residue; and *(ii)* increase the stiffness of the Arg residue by introducing one of the following constraints $C^{\alpha}=C^{\beta}$ or $C^{\beta}=C^{\gamma}$.
3) Electronic transitions predicted by TD-DFT calculations are consistent with the experimentally observed optical properties for the $RG^{E}D$ tripeptide. Calculation showed the importance of accounting for several DFT conformations of such flexible compound to reproduce the experimental data with computational means.
4) The RPAR CendR sequence, which has promising therapeutic uses as internalization trigger in anticancer nanodevices, binds both the (100) and (111) gold surface facets. Accordingly, gold nanoparticles can be considered as suitable vehicles for the transport and targeted delivery of this CendR peptide.
5) The conformation of the adsorbed RPAR differs considerably from the bioactive conformation. However, the conformations adopted by the adsorbed peptide, which display elongated-like shapes, are not stabilized by strong intramolecular interactions, like salt bridges or specific hydrogen bonds. This feature favors, despite the changes introduced in the conformation, the utilization of gold nanoparticles as carriers for the targeted delivery of the RPAR sequence. Thus, the elongated-like conformation of

the peptide adsorbed onto gold nanoparticles can easily change towards the bioactive conformation once the affinity for the NRP-1 receptor induces their targeted delivery.

6) The assembled cyclic peptide-polymer conjugate studied in this work (*i.e.* cyc[(L-Gln-D-Ala-L-Lys-D-Ala)$_2$] coupled with two poly(n-butyl acrylate) blocks) forms stable nanotube structures consisting on antiparallel stacks of the cyclic peptide units stabilized by inter-strand hydrogen bonds between main chain amide polar groups. Simulations show that the poly(n-butyl acrylate) segments deposited on mica tend to be partially contracted, which is agreement with experimental observations.
7) Using a theoretical approach that simulates a selective and progressive desolvation of the nanotube-like assembly, which mimics the presumed conditions in which the nanotube-like assembled peptide was deposited, we have demonstrated that the solvent presence during deposition is the main responsible for the unexplained conformational preferences of the acrylate blocks. DMF molecules and the acrylate segments always show very attractive interactions. However, when the proportion of DMF molecules drops to a critical point, determined to be between 50 and 40%, the loss of many attractive interactions induces a meaningful increase in the number of torsions that are arranged in conformations different from the *trans*.
8) Despite the attraction between the nanotube-like assembled conjugate and the mica surface, the loss of stability by

desolvation prior to completing the deposition, hampers the adoption of necessary geometrical rearrangements that the acrylate chains require to expose the highest possible interaction area. Only a sudden desolvation, prior to being deposited on mica, would facilitate the acrylate side chains to optimally interact with the ceramic surface.

9) MD simulations on a representative OMP show that the environment affects not only the stability of the β barrel but also the internal and surface exposed loops, which are related with the ion transport and with the assembly of protein units in homotrimers, respectively.
10) Water causes a very rapid and practically global destabilization of the OmpF structure that is accompanied by an increment of the SASA. The new structure is reached through both the spatial re-orientation of disorganized loops and the complete transformation of secondary motifs, as for example the conversion of a disorganized loop into a regular helix. These features indicate a competition between water···OmpF and OmpF···OmpF interactions.
11) DDM molecules in water induce a partial protective effect on the OmpF protein. More specifically, ordered βS and H motifs mostly maintain their structure, while poorly organized L+T motifs undergo severe conformational changes. These changes affect, amongst other things, the association capacity of OmpF, which explains our recent observation in porin-PLA nanodevices.

12) The OmpF protein is perfectly protected by DDM bilayers. This protection refers to ordered motifs but also to intrinsically disordered loops. Consequently, the OmpF maintains not only a well-defined channel structure but also its ability to form homotrimers through the L2 loop.
13) Results derived from OmpF simulations are extremely relevant for the improvement of bioinspired porin-integrated nanodevices, such as the nanoperforated PLA FsNMs for selective ion transport currently in fabrication. Indeed, this is expected to facilitate the fabrication process as well as to improve the efficiency of the membrane in terms of transport.
14) The topographic features of PEDOT deposited onto stainless steel using a polymerization time of 1 s have been studied by parceling the surface into small independent plots. The topographic characteristics of those polymer ultra-thin layers were similar to those observed after longer polymerization times and provided an excellent model system to test the developed modeling strategy. By reducing the size of the area to be explored at atomistic level, it was possible to implement a generation algorithm that produces rigid atomistic models of the bulk polymer over those parceled areas.
15) Despite the generation process built unrelaxed structures, their structures were close enough to achieve a realistic description, as relaxation of geometry constrictions was easily achieved using a combination of energy minimization and short runs of MD simulations.

The reliability of the relaxed models was demonstrated by providing accurate values for magnitudes used to describe the topography of surfaces and a correct description of the polymer ultra-structure. This is possible because the generation of micromodels ensures that all of them are uncorrelated and they can be combined to cover the surface previously parceled.

16) The loss of biological activity in PEDOT/LYZ electrodes has been attributed to the anisotropy of the PEDOT···LYZ electrostatic interactions, which is a phenomenon induced by the alignment of CP chains in the direction perpendicular to the metallic substrate.

17) P(EDOT-LYZ) electrodes, which were prepared by encapsulating the protein inside the PEDOT matrix, retain the most important inter-residue interactions at the LYZ active centre, preserving the antimicrobial activity. In this system, the anisotropic effects are minimized since the surrounding environment exerts forces of similar strength around the whole protein. Consequently, the formed electrostatic PEDOT···LYZ interactions are isotropic or quasi-isotropic and structural distortions are minimized.

18) The bactericidal activity of PEDOT/LYZ and P(EDOT-LYZ) biocapacitors is independent of the electric field applied or supplied during charge-discharge processes.