



UNIVERSIDAD NACIONAL DEL SUR

Departamento de Química

Tesis de Doctor en Química

Aspectos estructurales y dinámicos en la relajación de sistemas complejos, con énfasis en  
agua líquida, sobreenfriada y nanoconfinada

SEBASTIÁN ROBERTO ACCORDINO

Bahía Blanca

Argentina


2014

## Prefacio

Esta Tesis es presentada como parte de los requisitos para optar por el grado académico de Doctor en Química de la Universidad Nacional del Sur, y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otra. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el área IV (Fisicoquímica), Departamento de Química, dependiente de la Universidad Nacional del Sur (UNS) y en el Instituto de Química del Sur (INQUISUR), dependiente del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), durante el período comprendido entre abril de 2010 y octubre de 2014, bajo la dirección del Dr. Gustavo Adrián Appignanesi, Profesor Adjunto del Dpto. de Química de la UNS e investigador principal del CONICET y la co-dirección del Dr. J. Ariel Rodríguez Fris, Asistente de docencia con dedicación exclusiva del Dpto. de Química de la UNS e investigador adjunto del CONICET.

20 de octubre de 2014

Sebastián R. Accordino

	<p>UNIVERSIDAD NACIONAL DEL SUR Secretaría General de Posgrado y Educación Continua</p>
<p>La presente tesis ha sido aprobada el ....../....../..... , mereciendo la calificación de .....(.....)</p>	

## Agradecimientos

Quisiera agradecer a aquellas personas que me ayudaron y guiaron durante estos años para hacer posible este trabajo de Tesis:

En primera instancia a mi Director de Tesis, el Prof. Gustavo A. Appignanesi, por brindarme la posibilidad de trabajar junto a él y a su grupo de investigación e introducirme en esta temática de estudio tan apasionante. Quiero también agradecerle por su apoyo incondicional y su dedicación constante hacia mi trabajo, como así también por el ambiente de calidez humana y libertad en la que se desarrolló mi trabajo. No solo encontré en él a un excelente profesional sino que también a una gran persona.

En segundo lugar quiero agradecer a mi co-director de Tesis, el Dr. J. Ariel Rodríguez Fris, por todo el conocimiento, apoyo y confianza que me brindó desde el primer día de trabajo. Trabajar junto a él es un privilegio y un placer, tanto desde lo académico como desde lo humano.

Quiero agradecerles también a mis compañero/as del grupo de investigación, el Dr. Laureano M. Alarcón, el Dr. David C. Malaspina, la Prof. Marcela A. Morini, la Dra. M. Belén Sierra, el Lic. Joan M. Montes de Oca y la Lic. Cintia A. Menéndez, quienes con su trabajo aportaron significativamente al desarrollo de esta Tesis y por el apoyo y la grata compañía que me brindan.

También quiero agradecer a la Universidad Nacional del Sur, al Departamento de Química, y al Instituto de Química del Sur (INQUISUR) por el espacio que me brindaron para poder desarrollar este trabajo.

Mi agradecimiento también para el CONICET por las becas otorgadas, las cuales fueron muy importantes y me permitieron poder desarrollar este trabajo de Tesis.

Quiero agradecer también a los profesores que contribuyeron tanto a mi formación de grado como de posgrado.

También les quiero agradecer a todos mis amigos de la carrera de Lic. en Química, por su apoyo incondicional.

Finalmente siento que no tengo palabras que puedan expresar el agradecimiento y el reconocimiento hacia Romina, por hacerme feliz y por todo el apoyo y cariño sin los cuales todo sería más difícil; hacia mi familia; mis padres y en especial hacia mi madre, a quien le dedico de

corazón todos mis esfuerzos y mis compromisos, y a quien le estaré agradecido toda mi vida.



## Resumen

En esta Tesis estudiamos mediante simulaciones de dinámica molecular el comportamiento estructural y dinámico del agua pura o *bulk* como así también en diferentes contextos de nanoconfinamiento, particularmente en sistemas biológicos y en superficies modelo como lo son una placa de grafeno, nanotubos de carbono y monocapas autoensambladas de alcanos (SAMs). Específicamente nos interesa estudiar la relevancia del agua en la unión de proteínas y en el diseño de drogas, como así también en procesos de autoensamble de materiales sintéticos en ciencia de materiales y química supramolecular en medio acuoso.

Nuestras simulaciones de agua pura (Capítulo 2) o *bulk* avalan y profundizan el hecho de que, tanto en el régimen de temperaturas de líquido normal como de líquido sobreenfriado, el agua está formada por zonas de alta y baja densidad local (HDL y LDL respectivamente) [140, 145, 133, 111, 142, 134, 135, 27]. Específicamente, en colaboración con el Prof. Francesco Sciortino de la Universidad de Roma, La Sapienza, exploramos la robustez y validez (para diferentes densidades y temperaturas) de un método para determinar ambos tipos de moléculas (HDL y LDL) (Subsección 2.2.1), como así también determinamos la calidad geométrica de la estructura local de las mismas. Estudiamos la dependencia con la temperatura de la fracción de moléculas estructuradas (Subsección 2.2.4) y observamos que la misma puede ajustarse con una exponencial negativa, lo cual brinda un sustento al modelo de doble estado o *two state* para agua propuesto por el Prof. H. Tanaka [142, 134, 135].

Por otro lado estudiamos la estructura y orientación, en función de la temperatura, del agua de hidratación de la proteína lisozima y de una superficie modelo como lo es una placa de grafeno (Capítulo 3). Observamos que el ordenamiento local tetraédrico del agua mejora con el descenso de la temperatura, y en el caso de la placa de grafeno, las primeras capas de hidratación muestran una tendencia a asemejarse estructuralmente al hielo hexagonal Ih (Subsección 3.2.2), hecho similar al que ocurre en la interfase agua-aire pero en el sentido contrario a lo largo del eje *c* del plano basal [86]. En cambio, en el caso de las primeras capas de hidratación de la lisozima encontramos indicios de ambos tipos de comportamientos mencionados, junto con la tendencia de algunas moléculas de agua más próximas a la superficie a formar puentes de hidrógeno con la

proteína (Subsección 3.2.2). Mostramos que estas primeras capas de hidratación presentan una mejor estructura local que el *bulk* a temperatura ambiente, sin embargo la interfase introduce una restricción geométrica que reduce tanto la coordinación tetraédrica típica del agua *bulk* como la densidad local de las moléculas de agua adyacentes a la superficie. Por lo tanto dichas moléculas de agua tienden a formar tres puentes de hidrógeno con sus moléculas vecinas con el fin de lograr la máxima coordinación posible por puente de hidrógeno.

Abordamos también la hipótesis de la existencia de una transición de primer orden entre las dos fases líquidas del agua propuestas (agua líquida de baja densidad local y alta densidad local, LDL y HDL respectivamente). Esta supuesta transición líquido-líquido del agua estaría dada por la continuación de la línea de coexistencia de las fases HDA y LDA (donde HDA y LDA son simplemente las correspondientes formas vítreas) en el diagrama de fases del agua líquida sobreenfriada y la cual terminaría en un segundo punto crítico del agua, un punto crítico líquido-líquido [140, 108]. La extensión de la supuesta línea de coexistencia entre estas dos fases líquidas más allá del supuesto punto crítico líquido-líquido, se conoce con el nombre de línea de Widom [87]. Así nuestras simulaciones evidencian, a través de diferentes parámetros estructurales, la existencia de una transición dinámica en el agua de hidratación, compatible con la transición líquido-líquido, incorrectamente asociada con una transición vítrea de proteínas (Subsección 3.2.3).

El conocimiento desarrollado sobre el agua de hidratación nos condujo naturalmente a estudiar específicamente el papel que juega el agua en diferentes sistemas biológicos. De esta manera y en colaboración con el Prof. Ariel Fernández, estudiamos la dinámica del agua de hidratación en proteínas en base a un enfoque, basado en sus desarrollos previos [2, 48, 113, 114, 79], que codifica el comportamiento del agua nanoconfinada en un motivo estructural, los llamados dehidrones, los cuales representan defectos de empaquetamiento dados por puentes de hidrógeno, de la cadena principal de la proteína, desprotegidos o expuestos al efecto disruptivo del agua de hidratación. Los dehidrones resultan ser “pegajosos” al requerir del “arropamiento” o protección intermolecular por parte de grupos hidrofóbicos promoviendo así la asociación con ligandos. En este sentido demostramos (Capítulo 4) que dichos motivos estructurales en proteínas muestran una propensión a la deshidratación local y por ende representan regiones de alta hidrofobicidad

local en proteínas. Es esta tendencia a la labilización del agua circundante que poseen dichos motivos la que promueve la asociación de ligandos (otras proteínas o drogas) que deben reemplazar al agua de hidratación y así completar la desolvatación o secado de estas interacciones. Este efecto, en el cual básicamente un tercer cuerpo no polar “protege” a una interacción electrostática contribuyendo a la deshidratación o exclusión del agua de la misma se denomina cooperativismo molecular (Sección 4.2).

En consecuencia el tratamiento y estudio de las interfases proteína-proteína en base a este enfoque presentado anteriormente, implica un cambio de paradigma o, cuanto menos, un cambio en el nivel de análisis de este tipo de sistemas. Así, en el marco de esta colaboración desarrollamos en el seno de nuestro grupo de investigación un *alanine scanning in silico* o computacional (Subsección 5.2.1), basado en estos conceptos, y así logramos predecir satisfactoriamente los *hot spots* de un conjunto de complejos proteína-proteína y verificar la importancia de las interacciones de tres cuerpos o cooperativas en la asociación de proteínas [18]. Específicamente logramos racionalizar la naturaleza de los *hot spots* y así de las interacciones en las interfases proteína-proteína, un problema que desde hace mucho tiempo no ha podido resolverse dado que solo se han considerado enfoques basados en la aditividad de las interacciones de a pares [2].

Estos conceptos son también de vital importancia para el diseño racional de pequeñas moléculas disruptivas de interfases proteína-proteína, las cuales, en base a un diseño racional, pueden imitar la capacidad “protectora” de los *hot spots*. Consecuentemente verificamos la importancia de las interacciones de tres cuerpos o cooperativas en la interacción entre drogas y proteínas (Capítulo 6) al verificar que un conjunto de drogas disruptivas de interfases proteína-proteína tienden a mimetizar las interacciones cooperativas intermoleculares de la proteína que reemplazan [19]. En base a estos conceptos también estudiamos el caso particular de la mutación oncogénica Y220C de la proteína p53 (Sección 6.2) [59, 149, 137, 115, 90, 69] conocido como el “guardián del genoma”, el cual se encuentra mutado en la mitad de los casos de cáncer en seres humanos, como así también estudiamos y optimizamos la función de rescate sobre el p53 de la molécula PhiKan083 desarrollada por el grupo del Prof. Alan Fersht.

Finalmente en la última parte de esta Tesis extendemos preliminarmente las nociones anteriores al ámbito de la ciencia de materiales en medio acuoso (Capítulo 7). Sin embargo, en

estos contextos mucho más amplios la existencia de un motivo estructural prevalente no emerge claramente dada la heterogeneidad de sistemas de estudio. Por lo tanto, como primer paso, es necesario para este contexto desarrollar medidas eficaces de la hidrofobicidad local. Así, resulta de interés estudiar modelos simples de los cuales extraer principios generalizables. Estas cuestiones señalan un camino de continuidad de todo el trabajo expuesto hasta esta etapa, donde esperamos que el desarrollo de enfoques más generales, sean de gran utilidad en el ámbito de la ciencia de materiales y la bioingeniería, en donde también operan otras interacciones no covalentes no contempladas en los temas expuestos anteriormente y en donde se necesite un grado de detalle o resolución mayor al que hemos utilizado anteriormente. En ese sentido estudiamos diferentes parámetros como las fluctuaciones de densidad y tiempos de residencia de agua alrededor de una placa de grafeno, en el exterior de nanotubos de carbono de pared simple (Subsección 7.2.1), en superficies de monocapas autoensambladas de alcanos (SAMs) y en poros o cavidades de diferente tamaño talladas en la superficie de las SAMs (Subsección 7.2.2). En ese sentido verificamos el concepto de que las fluctuaciones de densidad de agua representan una buena medida de la hidrofobicidad/hidrofilicidad de las superficies [70]. Por último, en el marco de una colaboración con el Prof. Pablo Debenedetti de Princeton University demostramos que las fluctuaciones de densidad de agua alrededor de una placa de grafeno (un material ampliamente considerado como prototipo hidrofóbico) son prácticamente idénticas a las fluctuaciones sobre una superficie hidrofílica y muy diferentes de aquellas para una superficie hidrofóbica como las SAMs, lo cual denota un comportamiento hidrofílico de las superficies gráficas [37]. Mostramos que dicho comportamiento hidrofílico del grafeno se traduce en un comportamiento de interés práctico: previene la agregación de dos placas de grafeno.

## Abstract

In this Thesis we studied, by means of molecular dynamic simulations, the structural and dynamic behavior of both bulk water and water within different nanoconfinement conditions. Particularly, we studied biological systems like proteins and model surfaces like graphene sheets, carbon nanotubes and self-assembled monolayers (SAMs). We focused on the role of water in protein binding, in drug design, and also in the self-assembling of synthetic materials within an aqueous environment.

Our simulations of bulk water (Chapter 2) both support and deepen the picture of water as being composed by zones of high density liquid water and low density liquid water (HDL and LDL respectively) for a range of temperatures between normal and supercooled liquid [140, 145, 133, 111, 142, 134, 135, 27]. Specifically, in collaboration with Prof. Francesco Sciortino (Università di Roma, La Sapienza), we studied the robustness and validity (for different temperatures and densities) of a method to determine both kind of molecules (HDL and LDL) (Subsection 2.2.1). We also determined the geometrical quality of the local structure of these kinds of molecules. In this context, we investigated the temperature dependence of the fraction of structured molecules (Subsection 2.2.4) and we observed that it might be described by a negative exponential function, which provides support to the two state model for water of Prof. H. Tanaka [142, 134, 135].

On the other hand, we studied the structure and orientation, as a function of temperature, of the hydration water of the lysozyme protein and of a model surface, namely a graphene sheet (Chapter 3). We observed an improvement of the tetrahedral local order of water with decreasing temperature. In the case of the graphene sheet, we demonstrated that the first hydration layers exhibit a tendency to resemble hexagonal ice Ih (Subsection 3.2.2), similarly to the water-air interface but in the opposite way along the  $c$ -axis of the basal plane [86]. In turn, in the case of the first hydration layers of the lysozyme protein we found evidence of both kinds of behaviors, along with a clear trend of a fraction of the water molecules close to the surface to form hydrogen bonds with the protein (Subsection 3.2.2). We showed that these first hydration layers present a better local structure than the bulk at room temperature. However, the interface introduces a geometrical restriction that reduces both the tetrahedral coordination typical of

bulk water and the local density of the water molecules adjacent to the surface. Therefore these molecules tend to form three hydrogen bonds with neighboring water molecules to achieve the maximum coordination possible.

Also, we addressed the hypothesis of the existence of a first order transition between two phases of liquid water (HDL and LDL). This supposed liquid-liquid transition of water would be given by the continuation of the coexistence line of the HDA and LDA phases (where HDA and LDA are the respectively vitreous forms) in the supercooled liquid phase diagram of water which would end in a second critical point of water, a liquid-liquid critical point [140, 108]. The extension of this coexistence line beyond the supposed liquid-liquid critical point is known as the Widom line [87]. Within this context, our simulations showed, by means of different structural parameters, the existence of a dynamic transition in hydration water, compatible with the proposed liquid-liquid transition and usually incorrectly associated with a “protein glass transition” (Subsection 3.2.3).

The knowledge developed for hydration water led us to study the role of water for different biological systems. Thus, and in collaboration with Prof. Ariel Fernandez, we studied the dynamical behavior of protein hydration water based on an approach derived by his previous studies [2, 48, 113, 114, 79] which encodes the behavior of nanoconfined water in a structural motif, the so-called dehydron, which represents a packing defect given by backbone hydrogen bonds exposed to the disruptive effect of water. Dehydrons are found to be “sticky” since they require additional intermolecular “wrapping” (protection by hydrophobic groups) thus promoting protein association. In this sense, we showed (Chapter 4) that such protein structural motifs exhibit a local dehydration propensity, thus representing regions of high local hydrophobicity. This labilization of vicinal water promotes the association of ligands (other proteins or drugs) which must replace hydration water in order to bind and which thus complete the dehydration of the underwrapped intramolecular interactions of the target protein. This effect, in which a third non polar body “protects” an electrostatic interaction from the disruptive effect of water is defined as a molecular cooperativity process (Section 4.2).

Thus, the study of protein-protein interfaces based on this approach implies the introduction of a new paradigm or, at least, a change in the level of analysis of this kind of systems. In the

context of a collaboration with Prof. Ariel Fernandez, we designed an *in silico* alanine scanning scheme (Subsection 5.2.1) which succeeded in predicting the hot spots of a group of protein-protein complexes and verified the relevance of the three-body interactions for protein association [18]. Specifically, we rationalized the nature of the protein hot spots and, thus, of the protein-protein interactions, a problem that had been unsolved for many years due to the fact that it had only been considered in terms of the additivity of pairwise interactions [79].

These concepts are also very important for the rational design of small molecules disruptive of protein-protein interfaces which could mimic the protective function of the hot spots. Therefore we studied the importance of the three-body interactions for drug binding (Chapter 6), thus verifying that a group of drugs disruptive of protein-protein interfaces do in fact tend to mimic the cooperative interactions of the protein that they replace [19]. Within this framework we also studied the particular case of the oncogenic mutation Y220C in protein p53 (Section 6.2) [59, 149, 137, 115, 90, 69], a protein known as the “genome guardian” which is mutated in half of the cases of human cancer. Besides, we studied and optimized the rescue function of the small molecule PhiKan083 for the Y220C mutation developed by the group of Prof. Alan Fersht.

Finally, the last part of this Thesis is devoted to a preliminary study to extend the above-mentioned notions to the field of materials science within an aqueous environment (Chapter 7). However, the existence of a prevalent structural motif is not evident for this wider context because of the heterogeneity of the systems under study. Thus, as a first step it becomes necessary to develop effective measures of local hydrophobicity of relevance for this context. Thus, we studied simple models from which to extract generalizable principles. These issues point out a way of continuity for our work, where we expect that the development of more general approaches would be useful for the fields of materials science and bioengineering where other non-covalent interactions might be operative and where a resolution higher than the one we used in previous works would be demanded. In this sense we studied different parameters like water density fluctuations and water residence times at a graphene sheet, outside single-wall carbon nanotubes (Subsection 7.2.1), on the surface of alkane SAMs and inside pores of different size carved on the surface of the SAMs (Subsection 7.2.2). In that way we verified that the quantification of water density fluctuations represents a good measure of the surface hydrophobicity/hydrophilicity [70].

Within this context, and in collaboration with Prof. Pablo Debenedetti of Princeton University, we calculated water density fluctuations for graphene (a material widely regarded as a prototype hydrophobe) to show a typical hydrophilic behavior [37], and we demonstrated that parallel graphene sheets present a tendency to remain fully hydrated and thus, they do not autoensemble to undergo a hydrophobic collapse like model hydrophobic surfaces we used as a control.



## Publicaciones suscitadas por la Tesis:

El presente trabajo de Tesis dió lugar a los siguientes trabajos de investigación en revistas profesionales de circulación internacional:

1. *Hydrophobicity and geometry: Water at curved graphitic-like surfaces and within model pores in self-assembled monolayers,*

L. M. Alarcón, J. M. Montes de Oca, S. R. Accordino, J. A. Rodríguez Fris y G. A. Appignanesi, *Fluid Phase Equil.* **362**, 81 (2014).

2. *Wrapping Effects within a Proposed Function-Rescue Strategy for the Y220C Oncogenic Mutation of Protein p53,*

S. R. Accordino, J. A. Rodríguez Fris, G. A. Appignanesi, *PLoS ONE* **8**, e55123 (2013).

3. *Protein packing defects “heat up” interfacial water,*

M. B. Sierra, S. R. Accordino, J. A. Rodríguez Fris, M. A. Morini, G. A. Appignanesi y A. Fernández, *Eur. Phys. J. E.* **36**, 62 (2013).

4. *Temperature dependence of the structure of protein hydration water and the liquid-liquid transition,*

S. R. Accordino, D. C. Malaspina, J. A. Rodríguez Fris, L. M. Alarcón y G. A. Appignanesi, *Phys. Rev. E* **85**, 031503 (2012).

5. *A unifying motif of intermolecular cooperativity in protein associations,*

S. R. Accordino, J. A. Rodríguez Fris, G. A. Appignanesi y A. Fernández, *Eur. Phys. J. E.* **35**, 59 (2012).

6. *Wrapping mimicking in drug-like small molecules disruptive of protein-protein interfaces*,  
S. R. Accordino, M. A. Morini, M. B. Sierra, J. A. Rodríguez Fris, G. A. Appignanesi y A.  
Fernández,  
Proteins: Struct. Funct. and Bioinf. **80**, 1755 (2012).

7. *Comment on “Glass Transition in Biomolecules and the Liquid-Liquid Critical Point of Water”*,  
S. R. Accordino, D. C. Malaspina, J. A. Rodríguez Fris y G. A. Appignanesi,  
Phys. Rev. Lett. **106**, 029801 (2011).

8. *Quantitative investigation of the two-state picture for water in the normal liquid and the supercooled regime*,  
S. R. Accordino, J. A. Rodríguez Fris, F. Sciortino y G. A. Appignanesi,  
Eur. Phys. J. E. **34**, 48 (2011).

9. *Hydrophilic behavior of graphene and graphene-based materials*,  
S. R. Accordino, J. A. Rodríguez Fris y G. A. Appignanesi,  
*en proceso*.

# Índice general

<b>1. Introducción, antecedentes y objetivos</b>	<b>18</b>
1.1. Introducción, antecedentes y organización temática . . . . .	18
1.2. Contextos de estudio . . . . .	28
1.2.1. El agua: estructura y modelos . . . . .	29
1.2.2. Sistemas vítreos y líquidos sobreenfriados . . . . .	32
1.2.3. La estructura de las proteínas . . . . .	38
1.3. Objetivos . . . . .	43
<b>2. Agua bulk</b>	<b>45</b>
2.1. Agua líquida, agua vítrea y agua sobreenfriada . . . . .	45
2.2. Análisis cuantitativo de un modelo de doble estado para agua . . . . .	46
2.2.1. Detectando moléculas estructuradas y desestructuradas . . . . .	46
2.2.2. Robustez del índice de estructura local $I$ . . . . .	52
2.2.3. Parámetros de calidad de HBs para moléculas estructuradas y desestructuradas . . . . .	54
2.2.4. Dependencia con la temperatura de la fracción de moléculas estructuradas $\xi$ : testeando las predicciones del modelo de doble estado . . . . .	55
2.3. Conclusiones . . . . .	59
<b>3. Agua de hidratación y la transición vítrea en proteínas</b>	<b>61</b>
3.1. Agua de hidratación o biológica . . . . .	61
3.2. Estudio del agua de hidratación de la lisozima . . . . .	62

<i>ÍNDICE GENERAL</i>	16
3.2.1. Modelado y equilibración de los sistemas lisozima-agua y grafeno-agua . . .	63
3.2.2. Estructura, orientación y puentes de hidrógeno del agua de hidratación . .	64
3.2.3. El cambio en la estructura del agua de hidratación con la temperatura y la transición líquido-líquido del agua . . . . .	75
3.3. Conclusiones . . . . .	80
<b>4. Agua de hidratación en los defectos de empaquetamiento de proteínas</b>	<b>81</b>
4.1. La asociación de proteínas y el papel del agua de hidratación . . . . .	82
4.2. Cuantificando las interacciones cooperativas: definición del parámetro $\rho$ . . . . .	83
4.3. El desplazamiento cuadrático medio como medida de la movilidad del agua de hidratación . . . . .	85
4.4. La movilidad del agua en los dominios de desolvatación de BHBs . . . . .	87
4.5. Conclusiones . . . . .	92
<b>5. La exclusión del agua de hidratación y las interacciones cooperativas en la asociación de proteínas</b>	<b>93</b>
5.1. La necesidad de un concepto de cooperatividad biomolecular . . . . .	93
5.2. Identificación y caracterización de <i>hot spots</i> en interfases proteína-proteína . . . . .	95
5.2.1. <i>Alanine scanning</i> computacional basado en el concepto de cooperatividad	96
5.3. Conclusiones . . . . .	103
<b>6. Interacciones droga-proteína: hacia el diseño racional y la reingeniería de drogas</b>	<b>105</b>
6.1. Drogas disruptivas de interfases proteína-proteína y la mimetización de interac- ciones intermoleculares cooperativas . . . . .	106
6.1.1. El concepto de cooperativismo intermolecular en complejos proteína-droga	107
6.1.2. Resultados del cooperativismo molecular en la disrupción de complejos proteína-proteína por parte de pequeñas moléculas . . . . .	108
6.2. El concepto de cooperativismo como estrategia de recuperación de la función de la proteína p53 con la mutación oncogénica Y220C . . . . .	119

<i>ÍNDICE GENERAL</i>	17
6.2.1. El supresor tumoral p53: Un objetivo terapéutico clave en el tratamiento del cancer . . . . .	119
6.2.2. Reactivación del p53 mutante . . . . .	120
6.3. Conclusiones . . . . .	132
<b>7. Agua en superficies hidrofóbicas modelo y superficies gráficas: la hidroflicidad del grafeno</b>	<b>133</b>
7.1. Cuantificando la hidrofobicidad local mediante el cálculo de las fluctuaciones de densidad de agua . . . . .	135
7.2. Hidratación y geometría . . . . .	136
7.2.1. Agua en superficies gráficas: grafeno y nanotubos . . . . .	138
7.2.2. Agua en las monocapas autoensambladas de alcanos (SAMs) y el llenado de cavidades . . . . .	140
7.3. El comportamiento hidrofílico del grafeno y de los materiales gráficas . . . . .	143
7.4. Conclusiones . . . . .	152
<b>8. Conclusiones generales</b>	<b>154</b>

# Capítulo 1

## Introducción, antecedentes y objetivos

### 1.1. Introducción, antecedentes y organización temática

Si bien el agua es probablemente la sustancia más familiar para el ser humano, la comprensión de su comportamiento fisicoquímico aún dista de ser completa, a pesar de que ello resulte esencial para numerosos campos de enorme relevancia. Por ejemplo, el agua constituye “la matriz de la vida” y, por lo tanto, solo un pleno conocimiento de su comportamiento nos permitirá comprender a la biología a nivel molecular. En este sentido, en esta Tesis nos interesarán no sólo aspectos propios del agua pura sino también del agua interfacial y nanoconfinada, con su consiguiente relevancia en los procesos de unión de proteínas y en el diseño de drogas, como así también en procesos de autoensamble en ciencia de materiales.

El agua es quizás el material más estudiado de la tierra. Sin embargo, sorprendentemente, la ciencia detrás de su comportamiento y función todavía es relativamente pobre, no sólo para la gente en general, sino también para los científicos que trabajan en el tema. El pequeño tamaño de su molécula oculta la complejidad de su comportamiento y sus capacidades singulares. Las propiedades únicas del agua líquida y la naturaleza “camaleónica” de la misma parecen encajar perfectamente, como ninguna otra molécula, en los requisitos para la vida, aunque el rol específico que desempeña el agua en muchos procesos biológicos todavía no ha sido entendido

en su totalidad [9]. El agua es ubicua en nuestro planeta y además está considerada como el químico más importante en la vida humana. Por otro lado, se cree que la mayoría del agua en el universo es vítrea [126], la cual formaría parte de cometas y otros cuerpos celestes. En el ámbito científico, el agua es una fuente continua de fascinación debido a sus inusuales y contraintuitivas propiedades [108]. Por ejemplo, el agua líquida, en condiciones suficientemente frías, se expande, se vuelve más compresible al enfriar y a su vez su viscosidad disminuye cuando se comprime.

Unas de las cuestiones que abordaremos en esta Tesis es que el agua puede existir en al menos dos formas vítreas distintivas, agua amorfa (vítrea) de alta y baja densidad [140], HDA y LDA respectivamente (recientemente se ha encontrado que, en ciertas condiciones, también existe como agua amorfa de muy alta densidad, VHDA [88]), un fenómeno conocido como poliamorfismo. Las formas HDA y LDA son de gran interés en el estudio de sistemas vítreos, ya que al parecer presentan el primer ejemplo conocido de una transición de primer orden entre dos sólidos amorfos [53, 61]. En la primera parte de esta Tesis abordaremos la hipótesis de la existencia de una transición líquido-líquido en el agua. Una transición de primer orden entre dos fases líquidas diferentes: agua líquida de baja densidad y alta densidad (LDL y HDL respectivamente). Esta transición líquido-líquido estaría dada por la continuación de la línea de transición HDA-LDA (donde HDA y LDA son simplemente las correspondientes formas vítreas [108]), la cual terminaría en un segundo punto crítico del agua, un punto crítico líquido-líquido. La extensión de la supuesta línea de coexistencia entre las fases líquidas HDL y LDL más allá del supuesto punto crítico  $C_2$ , se conoce con el nombre de línea de Widom. En la Fig. 1.1 podemos observar la ubicación del supuesto segundo punto crítico del agua  $C_2$  en un diagrama de fases  $T - p$  [140]. La existencia de la transición líquido-líquido ha sido muy recientemente un tema de gran debate con visiones encontradas [32, 36, 33].

En los subsiguientes capítulos mostraremos como nuestras simulaciones, si bien no apuntan a resolver la controversia sobre la existencia de una transición líquido-líquido, avalan y profundizan el hecho de que, tanto en el régimen de temperaturas de líquido normal como de líquido sobreenfriado, el agua líquida está de hecho formada por zonas de alta y baja densidad local (HDL y LDL) (ver Fig. 1.2), apoyando así al modelo *two state* o de doble estado para agua propuesto por el Prof. H. Tanaka [142, 134, 135], el cual básicamente se refiere al agua como

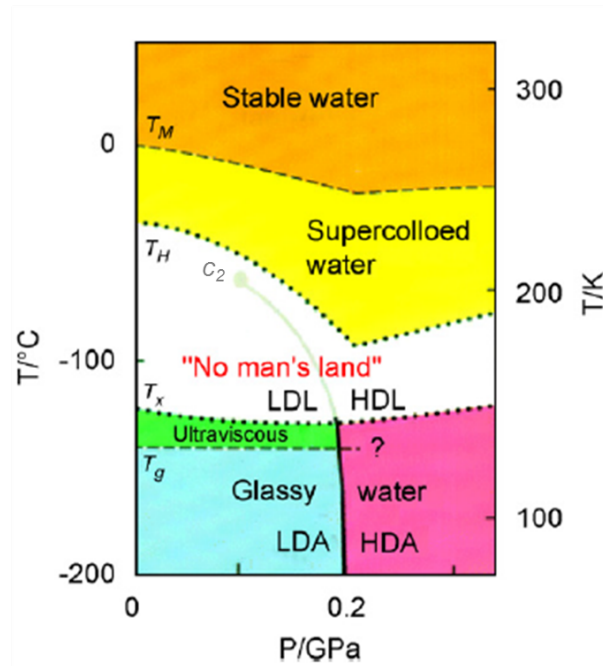


Figura 1.1: Diagrama de fases  $T-p$  de las distintas fases del agua a bajas temperaturas alrededor del régimen sobreenfriado. Adaptado de la Ref. [75]

una mezcla de unos pocos arreglos moleculares estructurados y de un estado desestructurado que puede tener muchas configuraciones locales diferentes [142, 134, 135].

A temperaturas elevadas las fases HDL y LDL son indistinguibles, como lo son las fases gaseosa y líquida por encima del punto crítico líquido-gas del agua. Así a temperaturas por debajo del supuesto punto crítico líquido-líquido del agua, existiría una transición de fase de primer orden entre dos formas diferentes de agua líquida sobreenfriada (HDL y LDL). Sin embargo, dado a que el supuesto punto crítico líquido-líquido del agua, determinado teóricamente y computacionalmente, se sitúa entre la temperatura de nucleación homogénea ( $T_H$ ) y la temperatura de cristalización del hielo cúbico ( $T_X$ ) (zona de color blanco denominada “*No man's land*” o “Tierra de nadie”, ver Fig. 1.1), se hace extremadamente difícil observar esta transición experimentalmente, ya que no se puede obtener agua líquida para poder verificar experimentalmente esta hipótesis [108] (ver Fig. 1.1).

Asociado con cualquier punto crítico existen “fluctuaciones críticas”, las cuales son más pronunciadas a temperaturas por encima de la correspondiente temperatura crítica. Por lo tanto,



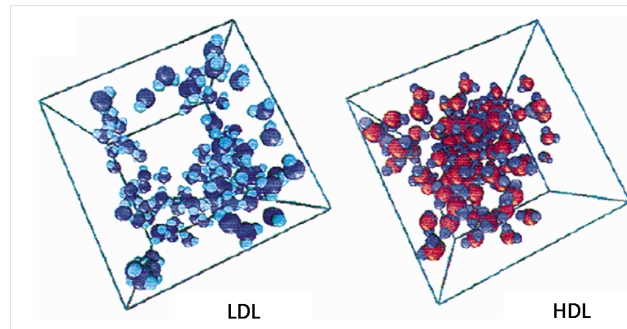


Figura 1.2: Configuraciones de una dinámica molecular de agua *bulk* a bajas temperaturas. Algunas moléculas de agua parecieran presentar una estructura similar a la estructura del agua amorfa o vítrea de baja densidad (izquierda), mientras que otras presentan una estructura parecida al agua vítrea de alta densidad (derecha). Adaptado de la Ref. [140].

el hipotético punto crítico líquido-líquido del agua explicaría el aumento dramático de ciertas propiedades como lo son: la compresibilidad, el calor específico, y el coeficiente de expansión térmico [108].

Otro de los temas tratados en este trabajo de Tesis es la teoría de la existencia de una supuesta transición vítrea de biomoléculas, la cual como veremos se debe, en realidad, a una transición dinámica del solvente, es decir, del agua.

Tanto los experimentos como simulaciones por computadora muestran que proteínas hidratadas sufren una especie de transición vítrea a temperaturas cercanas a los 200 K [87], donde por encima de esta temperatura presentan un movimiento difusivo y por debajo están atrapadas en modos armónicos [87]. Experimentos y simulaciones por computadora sugieren que cuando una proteína es solvatada, la transición observada en la proteína está fuertemente acoplada al solvente, llevandonos así a preguntarnos, si en realidad la supuesta transición vítrea de proteínas es una consecuencia directa de una transición dinámica del solvente que la rodea [87].

Como vimos anteriormente, se ha planteado la hipótesis de que en el diagrama de fases del agua, en la región de agua líquida sobreenfriada, existe una transición de primer orden entre agua de alta densidad (HDL) y baja densidad (LDL) [87]. La Fig. 1.3 es muy similar a la Fig. 1.1. Aquí podemos observar que la supuesta línea de coexistencia entre ambas fases líquidas posee pendiente negativa en el diagrama de fases  $p - T$  y termina con un punto crítico  $C_2$  [87], cuya

localización es  $T_{C2} = 200 \pm 20$  K y  $P_{C2} = 250 \pm 100$  MPa [87]. Al cruzar la línea de la transición de fase de primer orden, las funciones de estado termodinámicas cambian discontinuamente. Por debajo de la presión crítica  $P_{C2}$ , estas funciones de estado cambian rapido pero de manera continua durante el enfriamiento.

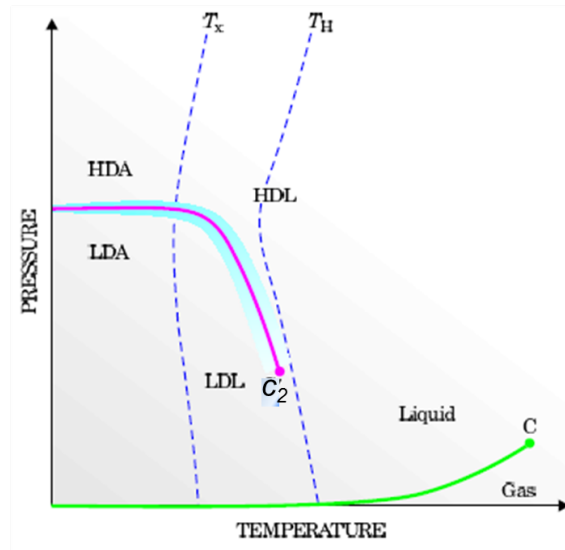


Figura 1.3: Similar a 1.1 pero en este caso se representa el diagrama de fases  $p - T$ . Adaptado de la Ref. [108].

De hecho, simulaciones por computadora de los modelos de agua TIP5P y ST2 muestran que las funciones termodinámicas, como el calor específico isobárico  $C_p$  y la compresibilidad isotérmica  $\kappa_T$ , presentan un máximo en función de la temperatura si el sistema es enfriado isobáricamente a  $p < p_{C2}$  [87].

Por otro lado, estudios experimentales en agua supercrítica [87] de hecho muestran que varias funciones de respuesta termodinámica poseen un máximo agudo en la región del diagrama de fases por encima del punto crítico líquido-vapor  $C_1$ . Sin embargo, como mencionábamos anteriormente, no se puede llevar a cabo dicho experimento en la región del punto crítico líquido-líquido  $C_2$  debido a la inevitable cristalización del agua *bulk* [87].

En cambio, se ha encontrado que el agua no cristaliza en poros nanométricos a temperaturas  $T > 200$  K [87]. Más aun, cuando este sistema se enfría a una presión constante  $p < p_{C2}$ , se observa un cambio en el comportamiento dinámico del agua de no Arrhenius a Arrhenius a  $T = T_x(p)$ . La línea  $T_x(p)$  (ver Fig. 1.3) se encuentra localizada en el rango de temperaturas de 200-230 K y posee pendiente negativa en el diagrama de fases  $p-T$ . En ese sentido, simulaciones por computadora sugieren que esta línea debe estar asociada en realidad con  $T_W(p)$ , la línea de Widom (la continuación de la línea de coexistencia entre las fases líquidas HDL y LDL más allá del supuesto punto crítico), cerca de la cual las características dinámicas locales cambiarían rápidamente, de aquellas que se asemejan a las propiedades del líquido HDL (a elevadas temperaturas), a aquellas semejantes a las propiedades del líquido LDL (a bajas temperaturas)[87]. Por lo tanto todos estos indicios nos conducen a pensar en que la supuesta transición vítrea de proteínas es un problema doblemente mal formulado, ya que no pareciera tratarse de una transición vítrea, y además dicha transición no sería consecuencia de una transición dinámica de la proteína misma, sino que de una transición propia del solvente.

Las inferencias, conceptualizaciones y herramientas, tanto estructurales como dinámicas a que arribaremos en nuestros trabajos con agua *bulk*, resultarán a su vez de gran interés en el contexto del comportamiento del agua de hidratación y nanoconfinada. En tal sentido, existen numerosas evidencias experimentales y computacionales [129, 97] que indican que el agua de hidratación y nanoconfinada suele presentar una mayor estructuración y/o disminución de la densidad local, a la par de exhibir signos de una dinámica más lenta (con reminiscencias vítreas). La primera cuestión ha sido sugerida experimentalmente mediante las técnicas de análisis de estructura fina por absorción de rayos x [*extended x-ray absorption fine structure* (EXAFS)] [124], reflectividad de neutrones (*neutron reflectivity*) [116], reflectividad de rayos x [*x-ray reflectivity* (XRR)] [117] y espectroscopía vibracional de frecuencia de suma [*vibrational sum frequency spectroscopy* (VSFS)] [91]. Por su parte, la aparición de una dinámica lenta, incluso a temperatura ambiente, ha sido revelada en estudios del agua de hidratación de proteínas, tal como indican las contribuciones de H. Frauenfelder [99, 94], M. Vogel [64], H. E. Stanley [87] y A. Bizarri junto a S. Cannistraro [121, 132]. Por ejemplo en la Ref. [121], el desplazamiento cuadrático medio [*mean squared displacement* (MSD)], de las moléculas de agua de la capa de hidratación de una pro-

teína (plastocianina) a temperatura ambiente, se ha determinado que sigue un comportamiento subdifusivo en las proximidades de la superficie,  $\langle \Delta r^2 \rangle \approx t^\alpha$ , donde  $\alpha$  adquiere valores menores a 1 ( $\alpha = 1$  rgimen difusivo). Este comportamiento del agua de hidratación se puede asociar a una dinámica compleja con un comportamiento de relajación de tipo exponencial extendido o ley de Kohlrausch, típico de sistemas vítreos en general, y en particular del agua *bulk* sobreenfriada. De hecho, trabajos previos de nuestro grupo de investigación sobre agua de hidratación en interfases con sistemas hidrofóbicos modelo, identificaron claramente la existencia de “agua lenta” a temperatura ambiente [24, 22, 23].

Dentro de este contexto, en la presente Tesis estudiaremos el comportamiento tanto estructural como dinámico del agua de hidratación de proteínas, contrastándolo con superficies y cavidades o poros modelo con una química y geometría controladas. Estos estudios del agua biológica, a su vez, nos llevarán a adentrarnos en contextos biofísicos más específicos. En este sentido, resulta cada vez más evidente que una plena comprensión del comportamiento del agua en la nano y la meso-escalas resultaría esencial para entender la biología a nivel molecular, desde por ejemplo, las interacciones proteína-proteína, hasta el funcionamiento de las membranas biológicas [35, 39]. En los procesos de organización biológica, el agua oficia de mediador entre complejas superficies que se relacionan generalmente por interacciones de tipo no-covalente [40, 62, 63, 46, 54, 22, 26, 23], generando ambientes de nanoconfinamiento para los cuales nuestro conocimiento del comportamiento del agua masiva suele tener valor escaso (o incluso resultar engañoso), por lo que se requiere del desarrollo de una nueva intuición. Por ejemplo, en la asociación de superficies hidrofóbicas, el nanoconfinamiento afecta las propiedades termodinámicas del agua de hidratación de modo de producir el “secado” promotor del colapso hidrofóbico. Adicionalmente, en el caso de la unión de un ligando a una proteína, se espera que éste desplace agua de hidratación fácilmente removible [63, 46]. De hecho, se han diseñado programas computacionales comerciales, como el “*WaterMap*” [63], para asistir al diseño de drogas basados en la determinación de regiones con solvente fácilmente removible, es decir, con alta hidrofobicidad o con energía libre “desfavorable”. Asimismo, el conocimiento de la existencia de regiones con agua lábil y con agua fuertemente unida se ha complementado con estudios de clusterización de fragmentos (en el espíritu del “*fragament-based drug discovery*”, o diseño de drogas por fragmen-

tos [46]). Paralelamente, los relevantes estudios del Prof. A. Fernández [48, 113] han determinado que las proteínas poseen defectos de empaquetamiento estructurales, llamados “dehidrones”, caracterizados por regiones en las que el *backbone* o cadena principal de una proteína se encuentra expuesta al solvente. Las proteínas solubles intentan proteger del ataque del agua a sus puentes de hidrógeno del *backbone*, “arropándolos” intramolecularmente con los grupos no-polares de sus cadenas laterales. Por su parte, los puentes de hidrógeno mal arropados o desprotegidos (dehidrones) resultan “pegajosos” al promover, intermolecularmente, la remoción del agua circundante, y en consecuencia han sido determinados como centrales para los fenómenos de asociación de proteínas de los que dependen la mayor parte de las funciones biológicas [48, 113]. Se supone además, que dichos defectos de empaquetamiento representan a su vez defectos de hidratación, al presentar una marcada propensión a la labilización y remoción del agua que requieren los procesos de asociación intermolecular [48, 113]. En este punto, y en el marco de una colaboración con el Prof. Ariel Fernández, estudiaremos en primer lugar como es la dinámica del agua de hidratación de dichos motivos. Estos estudios desembocaron naturalmente luego en la investigación de las interacciones entre proteínas y el diseño de fármacos.

Casi todas las funciones biológicas se basan en las interacciones proteína- proteína [143, 77, 153, 122]. Sin embargo, a pesar del reconocimiento de su importancia, la comprensión de los mecanismos moleculares que gobiernan la unión de proteínas está lejos de ser satisfactoria [143, 77, 153, 122]. De hecho, los fundamentos físicos de la selectividad y la especificidad no están totalmente elucidados. En este contexto, un descubrimiento de gran interés tanto desde el punto de vista del conocimiento básico como de las aplicaciones prácticas, fue que sólo una pequeña fracción de los residuos de la interfase de interacción de las proteínas contribuyen decisivamente a la unión (a estos residuos se los denomina “*hot spots*”) [143, 77, 153, 122]. Normalmente, las dimensiones de las interfases de unión en un complejo formado por dos proteínas son superiores a unos  $600 \text{ \AA}^2$  (lo que implica más de 40 residuos), mientras que aproximadamente el 5% de los residuos comprendidos (lo que significaría no mucho más de 1-3 residuos para las superficies más pequeñas) contribuyen por más de 2 kcal/mol a la energía libre de unión. En tal sentido, la técnica denominada “*alanine scanning*” [153] constituye una herramienta principal de experimentación. En esta técnica se realizan mutaciones sobre los residuos interfaciales por

alanina (Ala) y se calcula la diferencia entre el  $\Delta G$  del complejo mutado *vs.* el complejo sin mutar (“*wild type*”). Por lo tanto, este procedimiento determina la contribución de un residuo determinado a la unión de las proteínas. Este procedimiento podría ser entendido como el corte de la cadena lateral del residuo a la altura del carbono- $\beta$ , dado a que la cadena lateral de la alanina consiste simplemente en un metilo que a su vez es el carbono- $\beta$  (ver Fig. 1.10), mientras que aminoácidos más complejos poseen cadenas laterales más largas que le aportan su carácter específico (por ej. su naturaleza hidrofóbica, polar, etc.). Esta técnica ha determinado satisfactoriamente los “*hot spots*” en muchos complejos proteína-proteína [143, 77, 153, 122]. También análogos computacionales de este enfoque (empleando simulaciones de dinámica molecular con potenciales de interacción de energía libre sumamente completos e incluyendo asimismo los efectos de solvatación) [122] y/o redes neuronales computacionales con grandes bases de datos [77] han demostrado cierto éxito predictivo. Sin embargo, estos procedimientos no han sido capaces de identificar los principios más importantes subyacentes en las interacciones proteína-proteína (qué es lo que hace de un residuo un *hot spot*), ni siquiera dictar reglas de validez general (esta incapacidad ha sido abiertamente reconocida por los principales grupos de trabajo [77, 153, 122]). Y este último punto no sólo es crucial desde un punto de vista básico, sino también dadas las importantes aplicaciones que se basan en este conocimiento, tales como la predicción de la función de proteínas, el rediseño por bioingeniería de la interfase proteica y el descubrimiento de drogas [143, 77, 153, 122]. En particular, una vía novedosa e importante para el diseño de fármacos es la posibilidad de encontrar/diseñar pequeñas moléculas o péptidos con alta especificidad y afinidad que inhiban, bloqueen o alteren las interfases proteína-proteína, generalmente imitando o modulando las interacciones proteína-proteína. Sin embargo, estos esfuerzos casi en su totalidad carecen de elementos de diseño racional. La situación habitual (como por ejemplo en la lucha contra el cáncer, el HIV, la diabetes, el paludismo, etc.) es que una vez descubierta una interacción con una proteína o enzima involucrada en el desarrollo de una enfermedad, se lleva a cabo una revisión exhaustiva de muchas moléculas pequeñas (cientos de miles de compuestos e incluso millones de compuestos) con el fin de encontrar un medicamento que se una a la proteína/enzima, inhibiendo por tanto su unión a su ligando natural e interrumpiendo así la señal que desencadena la enfermedad [100]. Este procedimiento es de alto costo y no produce

los resultados optimizados que un diseño racional alcanzaría, a la vez de muchas veces ser susceptible de presentar graves efectos colaterales no deseados. Sin embargo, el diseño racional de compuestos capaces de interferir con las interacciones proteína-proteína es casi inexistente al día de hoy. Es de capital importancia entonces la comprensión de las bases de la afinidad y la especificidad interfacial. En este sentido, basados en los desarrollos previos del Prof. Ariel Fernández y trabajando en colaboración con él, nuestra hipótesis es que no sólo se deberían considerar las interacciones proteína-proteína al nivel de contactos de dos cuerpos (*“two-body”* o *“pair-wise”*), sino al nivel de interacciones de tres cuerpos (*“three-body interactions”*). De hecho, esperamos que las interacciones de tres cuerpos, destinadas a excluir el agua, constituyan los principales determinantes de *“hot-spots”*.

Por su parte, los principios que gobiernan la organización biológica a nivel molecular son los mismos que, en solución, actúan en el autoensamblado de estructuras moleculares para crear materiales sintéticos supramoleculares de dimensiones nanométricas [40, 63]. El solvente juega, por lo tanto, un papel fundamental en el diseño de nuevos materiales [40, 109, 55]. En ese contexto, en trabajos previos de nuestro grupo de investigación se han estudiado distintos sistemas con polaridad y curvatura controladas (de modo de poder separar la química de la geometría), determinando así tanto el comportamiento estructural como dinámico del agua de hidratación [22, 24]. Por ejemplo, nuestro grupo ha determinado que el tamaño mínimo de nanotubo de carbono que permite penetración del agua posee la mitad de diámetro que cavidades o túneles hidrofóbicos modelo [22, 14], lo que anticipa una cierta hidrofiliidad de la superficie (cabe notar que las superficies gráficas han sido estudiadas por varios años como sistemas hidrofóbicos paradigmáticos). De hecho, el consenso sobre su naturaleza de prototipo hidrofóbico [71] ha fomentado su uso generalizado para modelar el comportamiento del agua en condiciones de confinamiento no polar [97, 130, 105, 70, 71].

Por su parte, es evidente que el agua es el solvente natural en bioingeniería y en el diseño de drogas, a la vez de estar adquiriendo gran relevancia en la química supramolecular. Más allá del incremento en la importancia de la “química verde”, es interesante notar que los investigadores que trabajan en el ámbito de la síntesis orgánica están encontrando dificultades en lograr imitar el tipo de procesos que la naturaleza lleva a cabo todo el tiempo, como ser en los procesos

enzimáticos [109, 55] (obviamente en ese sentido la naturaleza tiene mucho que enseñarnos). Además, diversos trabajos de investigación han tratado de emular ciertas reglas de la biología para poder generar estructuras a escala nanométrica, trabajando con bloques de construcción (*building blocks*) para conseguir sintetizar materiales de gran complejidad y funcionalidad, distribuyendo de este modo un complejo problema estructural y funcional en diferentes niveles de integración [109, 55]. Sin embargo, para que esto sea efectivo deberíamos entender completamente los principios que gobiernan las interacciones, generalmente de naturaleza no-covalente, entre aquellas unidades modulares operando en ambientes de nanoconfinamiento en el cual el solvente es obviamente discreto y no un medio continuo.

Por lo tanto, el camino recorrido durante todo este trabajo de Tesis nos conduce finalmente a estudiar las propiedades de hidratación de superficies modelo, como monocapas autoensambladas de alcanos (SAMs) y superficies gráficas, en función de su geometría local, como así también las propiedades hidrofílicas de las superficies gráficas, las cuales durante varios años han sido consideradas superficies hidrofóbicas. Este conocimiento en sistemas simples que pretenden dilucidar principios generalizables y encontrar elementos de diseño, se espera que sea de utilidad para abordar procesos complejos en biología y en ciencia de materiales, para los cuales es evidente que resulta fundamental desarrollar una teoría de la hidrofobicidad al nivel de la nanoescala.

## 1.2. Contextos de estudio

Durante este trabajo de Tesis estudiaremos diferentes contextos o sistemas en los cuales el agua representa un protagonista esencial. El agua es el líquido más ubicuo y quizás el único que existe en los tres estados de la materia (sólido, líquido y gaseoso) en condiciones normales en nuestra biósfera. Aunque la composición del agua es muy simple, las propiedades físicas de la misma la hacen única frente a otras sustancias. El agua participa en la mayoría de los procesos biológicos, aunque sin embargo, el rol específico que desempeña en muchos de estos procesos no es todavía entendido en su totalidad [1]. En este trabajo de Tesis estudiaremos sistemas de particular interés en las áreas de fisicoquímica, biología como así también en ciencia de materiales.



Dada la variedad de sistemas que estudiaremos, a continuación realizaremos una breve introducción general a los sistemas de estudio. Sin embargo, en los sucesivos capítulos introduciremos algunos conceptos adicionales necesarios para el abordaje de cada sistema en particular. Por último enumeraremos cuales son los objetivos particulares de este trabajo de investigación.

### 1.2.1. El agua: estructura y modelos

La estructura del agua, en sus estados líquido o sólido, consiste en la asociación de retículos tetraédricos constituidos por moléculas de agua. Cada uno de estos tetraedros (Fig. 1.4) está formado por 5 moléculas de agua, donde cuatro moléculas se ubican en cada uno de sus 4 vértices mientras que la quinta molécula se aloja en el centro, de modo que la molécula central se encuentra coordinada por enlaces puente de hidrógeno (HB) (hydrogen bonds) con sus cuatro moléculas vecinas [3, pág. 61].

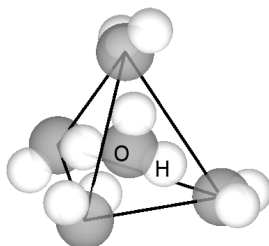


Figura 1.4: Un tetraedro de agua, el cual se forma imaginándose la unión entre los centros de los 4 oxígenos más cercanos a cada molécula de agua. Las esferas oscuras indican oxígenos mientras que las esferas blancas chicas indican átomos de hidrógeno. Adaptado de [3, pág. 77].

El enlace por puente de hidrógeno (HB), resulta de la interacción electrostática entre la carga parcial positiva expuesta de un protón (núcleo del átomo de hidrógeno) enlazado a un átomo de oxígeno inherentemente sustractor de electrones (más electronegativo que el hidrógeno), y la carga negativa de un par de electrones no enlazados de la capa de valencia de un segundo átomo de oxígeno perteneciente a otra molécula de agua [5, pág. 468]. Este orden que propone la Fig. 1.4 es solo perfecto en el hielo, mientras que en el agua líquida la red de tetraedros (que obviamente es dinámica) no es perfecta. Además, tal como se ha indicado en la Sección 1.1 y como veremos en el Capítulo 2, el agua presentaría regiones con moléculas de mayor o menor

densidad local, similares a las estructuras del agua amorfa de alta o baja densidad HDA y LDA respectivamente.

La primera simulación de agua líquida data del año 1969 realizada por Baker y Watts [169], la cual se llevó a cabo mediante la técnica de simulación Monte Carlo. Luego en 1971 Rahman y Stillinger realizan la primera simulación de dinámica molecular de agua líquida [167]. En este sentido a lo largo de los años se han propuesto distintos modelos de agua. La importancia del agua como sustancia vital puede atribuirse a la gran cantidad de anomalías en sus propiedades físicas y químicas en comparación con sustancias similares. Por lo tanto el modelado teórico del agua reviste la difícil tarea de reproducir estas anomalías. Hoy en día no existe un modelo tan complejo que reproduzca todas las particularidades del agua (o por lo menos no es viable computacionalmente). Es así como surgen distintos modelos de agua, cada uno de ellos con algunos aciertos y algunas falencias en intentar reproducir las propiedades de esta sustancia tan particular [1, pág. 54].

En los subsiguientes capítulos utilizaremos los modelos de agua explícita SPC/E [161], TIP3P [131] y TIP5P [131]. Los tres modelos son rígidos, es decir que cada molécula de agua se trata como un set rígido de masas puntuales. Los modelos SPC/E y TIP3P poseen tres puntos de interacción correspondientes a los tres átomos de cada molécula de agua, por lo que se conocen como modelos de tres sitios (ver Fig. 1.5). Cada sitio posee una carga puntual, y el sitio correspondiente al átomo de oxígeno posee también los correspondientes parámetros de Lennard-Jones. Dado a que los modelos de tres sitios logran una gran eficiencia computacional, estos son usados ampliamente en simulaciones de dinámica molecular. Por otro lado el modelo TIP5P posee cinco puntos de interacción (ver Fig. 1.6). Los dos puntos de interacción adicionales corresponden a los dos pares de electrones libres del átomo de oxígeno, los cuales se modelan como dos átomos ficticios con carga negativa enlazados al átomo de oxígeno con una geometría tetraédrica. Uno de los primeros modelos de este tipo fue el modelo BNS de Ben-Naim y Stillinger, el cual fue diseñado en 1971, luego le siguió el modelo ST2 propuesto por Stillinger y Rahman en 1974. Principalmente debido al elevado costo computacional, los modelos de cinco sitios no se desarrollaron mucho hasta el año 2000, cuando el modelo TIP5P de Mahoney y Jorgensen fue publicado [131]. En comparación con los modelos de agua anteriores, el modelo TIP5P posee mejoras en la

geometría de dímeros de agua, una estructura más tetraédrica que reproduce mejor las funciones de distribución radial, experimentales, de difracción de neutrones, como así también logra reproducir mejor la temperatura de máxima densidad del agua.

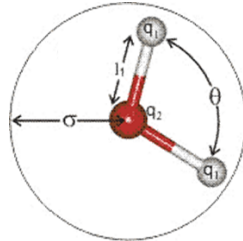


Figura 1.5: Diagrama de la geometría de los modelos de tres sitios SPC/E y TIP3P, los valores de los parámetros  $q_1$ ,  $q_2$ ,  $\theta$ ,  $l_1$  y  $\sigma$  se detallan para cada modelo en la tabla 1.1. Adaptado de la Ref. [10]

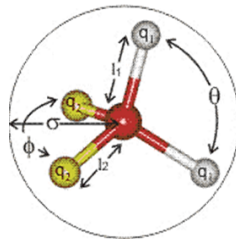


Figura 1.6: Diagrama de la geometría del modelo de cinco sitios TIP5P, los valores de los parámetros  $q_1$ ,  $q_2$ ,  $\theta$ ,  $l_1$  y  $\sigma$  se detallan en la tabla 1.2. Adaptado de la Ref. [10]

Modelo	$\sigma$ [Å]	$\epsilon$ [kJ/mol]	$l_1$ [Å]	$q_1$	$q_2$	$\theta$
SPC/E	3,166	0,650	1,000	0,4238	-0,8476	109,47°
TIP3P	3,15061	0,6364	0,9572	0,4170	-0,8340	104,52°

Cuadro 1.1: Parámetros de los modelos de tres sitios SPC/E y TIP3P, el detalle de la geometría del modelo y de los parámetros enumerados en esta tabla se pueden observar en el Fig. 1.5. Los parámetros  $\sigma$  y  $\epsilon$  son los correspondientes al potencial de Lennard-Jones para la interacción entre los átomos de oxígeno de las moléculas de agua. La cargas eléctricas parciales,  $q_1$  y  $q_2$ , coinciden con los centros de los átomos correspondientes (ver Fig. 1.5). La magnitud de estas cargas es  $q_1$  veces la carga elemental (la carga eléctrica del electrón) para cada uno de los átomos de hidrógeno, mientras que para cada uno de los átomos de oxígeno la carga es -2 veces la carga parcial de un hidrógeno,  $q_2 = -2q_1$ . Adaptado de la Ref. [10]

Modelo	$\sigma$ [Å]	$\epsilon$ [kJ/mol]	$l_1$ [Å]	$l_2$ [Å]	$q_1$	$q_2$	$\theta$	$\phi$
TIP5P	3,12	0,6694	0,9572	0,7	0,241	-0,241	104,52°	109,47°

Cuadro 1.2: Parámetros del modelo de cinco sitios TIP5P, el detalle de la geometría del modelo y de los parámetros enumerados en esta tabla se pueden observar en el Fig. 1.6. Los parámetros  $\sigma$  y  $\epsilon$  son los correspondientes al potencial de Lennard-Jones para la interacción entre los átomos de oxígeno de las moléculas de agua. La cargas eléctricas parciales,  $q_1$  y  $q_2$ , coinciden con los centros de los átomos correspondientes (ver Fig. 1.6). La magnitud de estas cargas es  $q_1$  veces la carga elemental (la carga eléctrica del electrón) para cada uno de los átomos de hidrógeno. En este caso la carga parcial negativa no se encuentra sobre el átomo de oxígeno sino que se encuentra equitativamente distribuída en los dos átomos ficticios enlazados al átomo de oxígeno central. Estos últimos poseen la carga parcial  $q_2 = -q_1$ . Adaptado de la Ref. [10]

### 1.2.2. Sistemas vítreos y líquidos sobreenfriados

La primera mitad de esta Tesis está dedicada al estudio de la estructura y la dinámica del agua en rangos de temperaturas que abarcan los regímenes de líquido normal y sobreenfriado. El comportamiento del agua sobreenfriada y vítrea es de gran interés en diversos contextos tanto en estado puro como en soluciones (por ejemplo, soluciones de agua y azúcar tienen un gran interés como criopreservantes [67, 74]). En estado sobreenfriado o vítreo, el agua posee una estructura similar a la del líquido, por lo que en estos estados también resulta esencial en muchos procesos biológicos [3, pág. 61]. De hecho, como hemos mencionado anteriormente, el

agua de hidratación o interfacial y nanoconfinada suele presentar ciertas reminiscencias vítreas, como una dinámica más lenta o subdifusiva [22, 24, 23, 129, 97, 124, 116, 117, 91] y una mayor estructuración o más baja densidad local [22, 24, 23, 99, 94, 87, 121, 132, 129, 97]. Por lo tanto, a continuación daremos una somera explicación del fenómeno de la transición vítrea y de los líquidos sobreenfriados. Este tema ha sido objeto de trabajo para nuestro grupo de investigación en los últimos años (para ello, el lector puede referirse por ejemplo a las Refs. [31, 29, 27] y a la Tesis Doctoral de J. A. Rodríguez Fris [3]).

En general podemos definir a un líquido sobreenfriado como aquel sistema que permanece en estado líquido y no ha cristalizado por debajo de su correspondiente temperatura de fusión (temperatura de *melting*  $T_m$ ) [1]. Además si este sistema tiene facilidad para alcanzar el régimen sobreenfriado se dice que es un buen formador de vidrios. Definiremos aquí también a los sistemas vítreos o vidrios como sólidos metaestables que carecen de orden a largo alcance. Recordemos que los sistemas sólidos se pueden clasificar de acuerdo a su estructura en sólidos cristalinos o en sólidos amorfos [1]. Por un lado los sólidos cristalinos se caracterizan por presentar una estructura periódica de largo alcance, mientras que en el caso de los sólidos amorfos estos presentan una estructura que solo es local y no periódica. Así, dentro del grupo de los sólidos amorfos encontramos a los sistemas vítreos.

En la Fig. 1.7 mostramos en forma esquemática la dependencia del volumen  $V$  o la entalpía  $H$  para una sustancia en función de la temperatura  $T$ , a presión  $p$  constante. La sustancia permanece en estado líquido por encima de su temperatura de fusión termodinámica  $T_m$ . Por debajo de  $T_m$  tiene, en general, dos opciones: puede cristalizar (camino  $lc \rightarrow c$ ), lo cual ocurre cuando los constituyentes (átomos, iones o moléculas) del sistema forman una estructura ordenada a corto y largo rango, o por otro lado si es enfriado lo suficientemente rápido como para no permitir que las moléculas o átomos puedan ordenarse en una estructura cristalina, el sistema puede continuar en estado líquido en un régimen denominado sobreenfriado (camino  $ls \rightarrow a$  ó  $ls \rightarrow b$ ). En este último estado, y al disminuir la temperatura, el líquido alcanza eventualmente una temperatura denominada de transición vítrea calorimétrica  $T_{g,c}$  la cual depende de la tasa de enfriamiento: cuanto mayor sea esta tasa, mayor será  $T_{g,c}$  [6, pág. 242]. En la figura podemos observar dos de estas temperaturas,  $T_{g,c,a}$  y  $T_{g,c,b}$ , donde la primera se

obtuvo con una mayor tasa de enfriamiento que la segunda. La forma experimental de estimar  $T_{g,c}$  es cuando el coeficiente de expansión térmica  $(\partial V/\partial T)_p$  o la capacidad calorífica isobárica  $C_p = (\partial H/\partial T)_p$  cambia bruscamente en función de la temperatura [159, pág. 2409], tal como sucede en los puntos  $a$  y  $b$  de la figura. La  $T_{g,c}$  ocurre cuando el tiempo característico para las reacomodaciones de los constituyentes (la denominada *relajación*) se vuelve comparable con la escala de tiempo del experimento, por lo que el líquido sobreenfriado parecerá estructuralmente arrestado cinéticamente en dicha escala de tiempo, vale decir se ha vitrificado, *e. g.* es un vidrio.

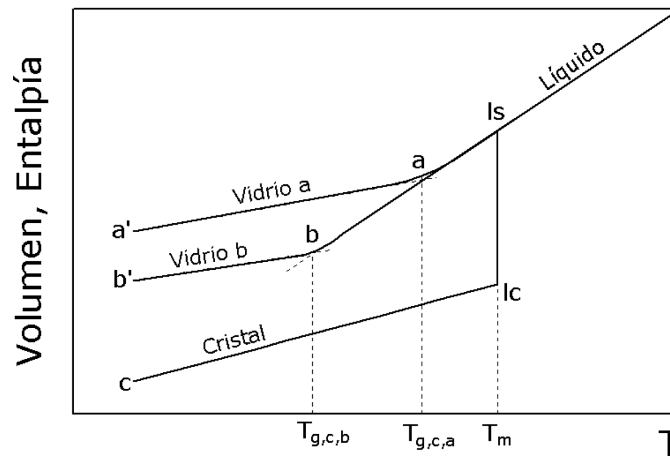


Figura 1.7: Esquema del comportamiento del volumen  $V$  o la entalpía  $H$  para una sustancia *versus* la temperatura  $T$  a presión  $p$  constante. Al enfriar el líquido por debajo de su temperatura de fusión termodinámica  $T_m$ , este puede cristalizar (camino  $lc \rightarrow c$ ) o sobreenfriarse (camino  $ls \rightarrow a$  ó  $ls \rightarrow b$ ). Este último es un estado metaestable, el estado de equilibrio real es el cristal. Al enfriarse el líquido sobreenfriado eventualmente sufre una transición (vítrea) a la temperatura  $T_{g,c}$  donde el líquido parecerá arrestado cinéticamente, se ha convertido en un vidrio. Esta  $T_{g,c}$  depende de la tasa de enfriamiento  $dT/dt$  ( $t$  es tiempo): el vidrio  $b$  (camino  $b \rightarrow b'$ ) se obtiene con menor  $dT/dt$  que el vidrio  $a$  (camino  $a \rightarrow a'$ ). El coeficiente de expansión térmica  $(\partial V/\partial T)_p$  y la capacidad calorífica isobárica  $C_p = (\partial H/\partial T)_p$  cambian bruscamente aunque continuamente a  $T_{g,c}$ . Adaptado de [126, Fig. 1] y [6, Figs. 4.3 y 4.4].

Los fenómenos que ocurren durante el sobreenfriamiento son el resultado de la interacción entre dos tiempos característicos: el tiempo de relajación  $\tau$ , relacionado con la movilidad o reacomodación del líquido, y el tiempo de cristalización  $t_c$  [6, pág. 236] al cual el líquido inexo-

rablemente cristaliza. La competencia entre cristalización y vitrificación se da por el hecho que  $\tau$

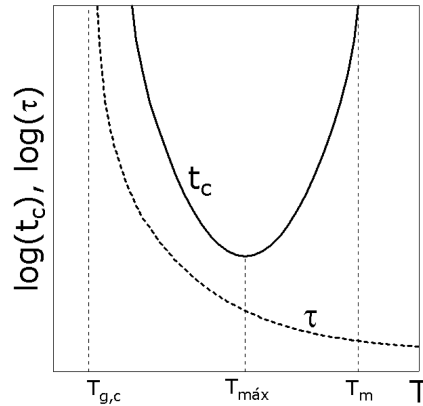


Figura 1.8: Esquema de la dependencia con la temperatura del tiempo de cristalización del sistema  $t_c$  (línea sólida) y el comportamiento típico del tiempo de relajación  $\tau$  para un líquido (curva punteada gruesa). Notamos que esta última aumenta rápidamente al aproximarse a la temperatura de transición vítrea  $T_{g,c}$ . Por debajo de la temperatura de fusión termodinámica  $T_m$  podremos estudiar el líquido sobreenfriado dependiendo del valor  $|T - T_{máx}|$  (ver texto para más detalles), donde  $T_{máx}$  es la temperatura a la cual  $t_c$  es mínimo. Adaptado de la Fig. 4.2 de la Ref. [6].

aumenta con el sobreenfriamiento, mientras que  $t_c$  tiene una dependencia con  $T$  no-monotónica, como mostramos en la Fig. 1.8 en forma esquemática. En este contexto se dan las siguientes posibilidades,

- Si el sobreenfriamiento es débil ( $T_{máx} \ll T < T_m$ ), donde  $T_{máx}$  es la temperatura a la cual  $t_c$  es mínimo, el líquido puede permanecer en este estado metaestable (es decir, un estado de no-equilibrio donde la energía de Gibbs <sup>1</sup> no es la mínima dados los parámetros  $T$  y  $p$ ) por un tiempo relativamente largo debido a que  $t_c$  es muy alto (nótese la escala logarítmica en el eje de ordenadas).
- Cuando  $T \sim T_{máx}$ , sólo podremos estudiar el líquido sobreenfriado en un tiempo del orden de  $t_c$ . Este hecho es debido a que alrededor de  $T_{máx}$  la fuerza impulsora para la nucleación (y subsecuente cristalización) es alta por lo cual  $t_c$  es pequeño y el líquido tiene una alta tendencia a cristalizar.

<sup>1</sup>La energía de Gibbs es aún conocida por su antiguo nombre, la “energía libre” [5, pág. 92].

- Para  $T_{g,c} < T \ll T_{\text{máx}}$ , el líquido puede nuevamente permanecer en estado sobreenfriado por un tiempo relativamente largo, debido a que la tasa de nucleación disminuye y entonces  $t_c$  aumenta nuevamente. Este efecto se produce porque las partículas deben moverse para formar los núcleos (que posteriormente formarán los cristales) y como mencionamos, la movilidad (a través de la relajación  $\tau$ ) se encuentra disminuida por la baja  $T$ .
- Por debajo de  $T_{g,c}$ , el tiempo  $\tau$  es muy grande en la escala de tiempo considerada, por lo que el líquido se ha vitrificado.

Una consecuencia importante respecto de estos últimos cuatro items es que, en principio, cualquier líquido puede sobreenfriarse o vitrificarse. El único requisito es enfriarlo suficientemente rápido por encima o debajo de  $T_{g,c}$  respectivamente para evitar su cristalización.

El desplazamiento cuadrático medio (*Mean square displacement*, MSD) es una de las propiedades dinámicas más simples utilizada para el análisis de la dinámica de líquidos formadores de vidrios, como así también para el análisis del comportamiento subdifusivo del agua de hidratación (en el Capítulo 4 estudiaremos la dinámica del agua de hidratación de un grupo de proteínas, para ello utilizaremos el MSD como medida de dicho comportamiento dinámico, por lo cual resulta útil introducir brevemente dicho parámetro). El mismo se define como,

$$\langle r^2(t) \rangle = N^{-1} \sum_{i=1}^N \langle [\mathbf{r}_i(t) - \mathbf{r}_i(0)]^2 \rangle \quad ,$$

donde  $\mathbf{r}_i(t)$  es la posición del átomo de oxígeno de la molécula de agua  $i$  al tiempo  $t$ , por lo tanto el término entre  $[\dots]$  representa el desplazamiento de la molécula  $i$  en el intervalo de tiempo  $[0,t]$ .

A modo de ejemplo en la Fig. 1.9 representamos la dependencia de  $\langle r^2(t) \rangle$  con  $t$  para el promedio sobre las partículas  $A$  de un sistema binario  $A-B$  de Lennard-Jones. Nuestro grupo ha trabajado con este tipo de sistemas en el estudio de los mecanismos generales de la relajación vítrea en líquidos sobreenfriados (para más detalles sobre los estudios desarrollados por nuestro grupo en este tema recomendamos al lector dirigirse a las Refs. [3, 1]).

El análisis de una curva de desplazamiento cuadrático medio nos proporciona información muy útil para entender mejor el comportamiento dinámico de los líquidos sobreenfriados. En la



Fig. 1.9 podemos apreciar los distintos regímenes de comportamiento en el movimiento de las partículas. A temperaturas altas  $\langle r^2(t) \rangle$  muestra dos regímenes. A tiempos cortos, cuando la partícula aún no siente la presencia de sus vecinas, se establece un régimen inercial en el cual  $\langle r^2(t) \rangle \propto t^2$ . A tiempos largos el régimen es difusivo y  $\langle r^2(t) \rangle \propto t$ . En este régimen difusivo, la “colisión” constante de las partículas con sus compañeras hace en promedio para tiempos suficientemente largos (notar la escala de tiempo logarítmica) un movimiento al azar caracterizado por un proceso aleatorio, comunmente denominado caminata aleatoria, el cual es Gaussiano [3].

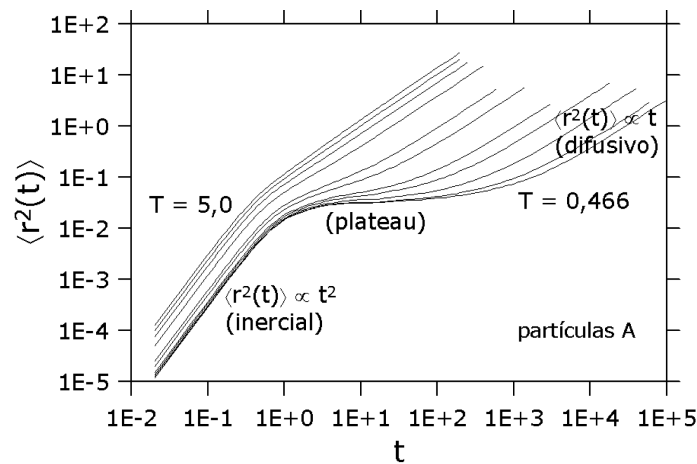


Figura 1.9: Desplazamiento cuadrático medio en función del tiempo para las partículas A del sistema binario de Lennard-Jones. Las curvas corresponden a diferentes temperaturas (de izquierda a derecha respectivamente): 5,0; 4,0; 3,0; 2,0; 1,0; 0,8; 0,6; 0,55; 0,50; 0,475 y 0,466. Se aprecian los regímenes: inercial (o balístico), de plateau y difusivo [3]. Tomado de [152].

Para temperaturas bajas la dependencia de  $\langle r^2(t) \rangle$  con  $t$  es más compleja (curvas inferiores en Fig. 1.9): aunque uno encuentra nuevamente los regímenes balístico y difusivo a tiempos cortos y largos, respectivamente, se hallan separados por una ventana temporal en la cual  $\langle r^2(t) \rangle$  muestra una muy débil dependencia con  $t$  (plateau). Este régimen intermedio es debido al denominado “efecto de caja”, hecho por el cual a tiempos intermedios la partícula rebota contra sus vecinas por lo que su movimiento es básicamente de vibración [3].

### 1.2.3. La estructura de las proteínas

En esta Tesis estudiaremos diferentes sistemas de proteínas de relevante interés terapéutico. Estudiaremos el comportamiento estructural y dinámico del agua de hidratación de proteínas, como así también abordaremos temas de índole más biológico o biofísico como lo son la interacción entre proteínas y el diseño de fármacos, en los cuales como veremos, el agua de hidratación (de hecho, su remoción) juega un papel esencial. Por lo tanto, resulta oportuno introducir al lector los conceptos básicos necesarios para abordar estos temas y así facilitar la lectura e interpretación de los resultados a los que arribaremos en los capítulos posteriores.

#### Propiedades de los aminoácidos

Las unidades fundamentales que conforman a las proteínas son veinte L- $\alpha$ -aminoácidos [El prefijo  $\alpha$  significa que el grupo amino está unido al carbono contiguo al grupo carboxilo (carbono  $\alpha$ ), la letra L hace referencia a la orientación relativa en el espacio de los 4 grupos distintos unidos al carbono  $\alpha$ . Todos los aminoácidos proteicos son L-aminoácidos, pero esta nomenclatura no indica si el compuesto es dextrógiro o levógiro, es decir si desvía el plano de la luz polarizada hacia la derecha o hacia la izquierda]. Sus estructuras, nombres, códigos de una letra y de tres letras, usados como abreviatura, se muestran en la Fig. 1.10. En los veinte casos, excepto en el caso de la Glicina (Gly), el carbono- $\alpha$  es un centro quiral o asimétrico. Como mencionamos anteriormente, las proteínas o péptidos sintetizados en los ribosomas de la célula, poseen la configuración del estereoisómero L, el cual es el que se muestra en la Fig. 1.10. Dada la universalidad en la configuración del carbono- $\alpha$ , el prefijo L generalmente será omitido. Eventualmente se han encontrado D-aminoácidos en algunos pocos péptidos de origen natural. Sin embargo, estos últimos no son sintetizados en los ribosomas [7].

Dos aminoácidos, Isoleucina (Ile) y Treonina (Thr), poseen un segundo centro asimétrico o quiral en el carbono- $\beta$ . Su configuración también es constante, y la correcta esteroquímica es la que se muestra en la Fig. 1.10.

La estructura covalente básica de una proteína es una cadena polipeptídica, tal cual se muestra en la Fig. 1.11.

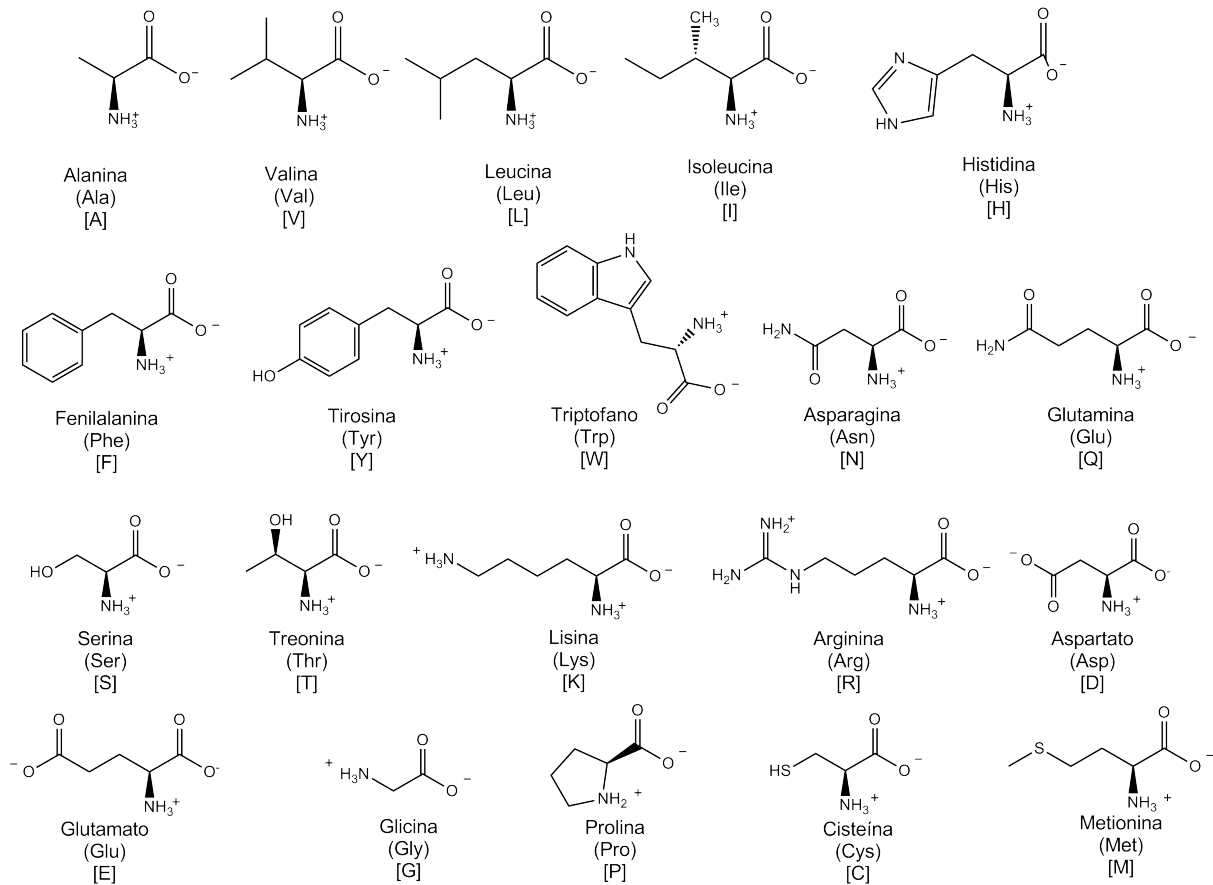


Figura 1.10: Estructura de los veinte aminoácidos a  $\text{pH} = 7$  con sus respectivos códigos de una letra y tres letras.

Los aminoácidos en la Fig. 1.10 se presentan en su forma iónica, la cual predomina a  $\text{pH} = 7$ . El grupo amino se encuentra protonado a este valor de  $\text{pH}$ , mientras que el grupo carboxilo se encuentra desprotonado. Por lo tanto, la mayoría de los aminoácidos no poseen carga neta a  $\text{pH}$  neutro, pero existen como zwitteriones ( $^+\text{NH}_3\text{-CHR-COO}^-$ ) y en consecuencia tendrán un elevado momento dipolar. Cuatro aminoácidos poseen una carga adicional en sus cadenas laterales (R) a  $\text{pH} = 7$ : Asp y Glu son negativos; Lys y Arg son positivos [7].

### Estructura primaria, secundaria, terciaria y cuaternaria

La estructura primaria de una proteína consiste en la forma de organización más básica, es decir que está determinada por la secuencia de aminoácidos de la cadena proteica, el número de aminoácidos presentes y el orden en que están enlazados por medio de enlaces peptídicos.

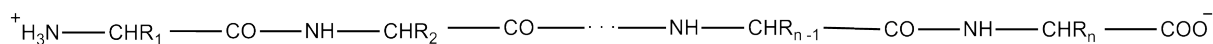


Figura 1.11: Esquematación del enlace peptídico y cadena principal (*backbone*) en una proteína. Los grupos R representan la cadena lateral de cada uno de los aminoácidos que conforman la proteína.

La determinación de la secuencia de aminoácidos de una proteína es una técnica bien establecida [7, pág. 66]. La secuencia de aminoácidos de una cadena polipeptídica determina el tipo de interacciones que se producirán (dentro de la misma proteína como con su entorno) y el grado de libertad de adoptar diferentes conformaciones estables a una temperatura fisiológica. Además el conocimiento de la estructura primaria es fundamental para muchos otros estudios bioquímicos y biofísicos. Un ejemplo de esto es la determinación de la estructura tridimensional de proteínas mediante la técnica de cristalografía de rayos-x. En raras ocasiones solo son suficientes los resultados arrojados por la cristalografía de rayos-x para determinar la estructura tridimensional de la proteína. En general los resultados nunca son del todo claros como para poder determinar inequívocamente la secuencia de aminoácidos de la cadena principal (*backbone*). Sin embargo, la mayor parte del *backbone* usualmente es visible a las resoluciones típicas obtenidas, y algunas de las cadenas laterales más voluminosas también se pueden identificar claramente. Por lo tanto, se utiliza la secuencia de aminoácidos y el mapa cristalográfico de densidad electrónica a la par, para tratar de determinar precisamente la estructura de la proteína. Para más detalles sobre este proceso dirigirse a la Ref. [8, Capítulo 13]. En la referencia [7, pág. 68-86] se describen otras aplicaciones directas del conocimiento de la estructura primaria, como lo son entre ellas la predicción de la estructura secundaria y terciaria.

La estructura secundaria de las proteínas es el plegamiento regular local entre residuos aminoácidos cercanos de la cadena polipeptídica. Este tipo de estructura se adopta gracias a la formación de enlaces de hidrógeno entre los grupos carbonilo (-CO-) y amino (-NH-) de la cadena principal o *backbone*. A este tipo de interacciones por puente de hidrógeno las llamaremos BHB (*backbone hydrogen bond*).

La hélice- $\alpha$  es la estructura secundaria polipeptídica más famosa. Fue postulada por Linus Pauling y Robert Corey seis años antes que una estructura de hélice- $\alpha$  se pueda observar por

primera vez con una resolución del orden molecular en la estructura cristalina de la mioglobina [7, pág. 91]. La Fig. 1.12 muestra un esquema de este tipo de estructura secundaria. Las características básicas son: 3,6 residuos por vuelta; cada residuo  $i$  forma un puente de hidrógeno con un residuo ubicado a 4 unidades de distancia ( $i+4$ ) en la secuencia primaria; los carbonos- $\alpha$  de los residuos contiguos están separados 1,5 Å; la hélice está estrechamente empaquetada y debido a que todos los residuos poseen la configuración L, todas las cadenas laterales apuntan hacia afuera de la hélice. De esta manera, la misma estructura puede acomodar casi cualquier tipo de cadena lateral flexible, siempre y cuando no sea lo suficientemente grande como para interferir estéricamente. Vale decir que la Prolina (Pro) no puede, por cuestiones estéricas formar parte de una hélice- $\alpha$  [7, Capítulo 5]. Además, algunas cadenas laterales energéticamente no preferirán formar parte de una hélice- $\alpha$  [7, pág. 91].

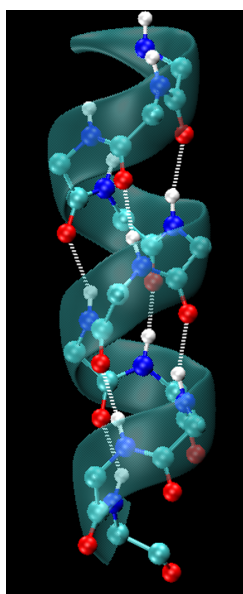


Figura 1.12: La hélice- $\alpha$ : estructura secundaria helicoidal dextrógira.

La lámina- $\beta$ , u hoja plegada- $\beta$ , es otra de las estructuras secundarias posibles adoptadas por proteínas. Se forma por el posicionamiento paralelo de dos cadenas de aminoácidos dentro de la misma proteína, donde los grupos -NH- de una de las cadenas forman enlaces de hidrógeno con los grupos -CO- de la opuesta. Las cadenas laterales (R) de esta estructura están posicionadas por encima y por debajo del plano de las láminas. Estos grupos R no deben ser muy grandes

para no crear un impedimento estérico, ya que en consecuencia se vería afectada la estructura de la lámina. Como en el caso de las hélices- $\alpha$ , donde puede existir una hélice destrógira o levógira, aquí también son dos las posibilidades: lámina- $\beta$  paralela y lámina- $\beta$  antiparalela. En la primera de ellas ambas cadenas van de un grupo amino a un grupo carboxilo en el mismo sentido (ver Fig. 1.13), mientras que en el segundo caso una de las cadenas paralelas va del grupo amino al carboxilo en un sentido determinado y la otra lo hace en el sentido contrario.

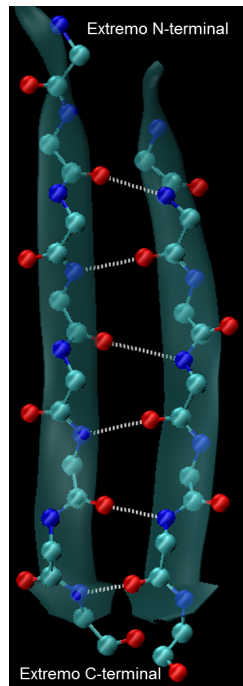


Figura 1.13: La lámina- $\beta$ : estructura secundaria de lámina- $\beta$  paralela

Cada una de estas estructuras existen en las proteínas exactamente como fueron predichas por Pauling y Corey [7, pág. 94].

La estructura terciaria es el modo en que las cadenas polipeptídicas se pliegan en el espacio, básicamente como se “enrolla” una determinada proteína, ya sea globular o fibrosa. Es decir que la estructura terciaria es la disposición de los dominios de la proteína en el espacio 3D.

Finalmente la estructura cuaternaria deriva de la conjunción de varias cadenas peptídicas que, asociadas, conforman un complejo, un multímero, que posee propiedades diferentes a la de sus monómeros componentes. Dichas subunidades se asocian entre sí mediante interacciones no covalentes.

### 1.3. Objetivos

Como ya mencionamos anteriormente, en la primera parte de este trabajo de Tesis comenzaremos estudiando un sistema de agua pura o agua *bulk*. Analizaremos el comportamiento estructural y dinámico del agua en rangos de temperaturas que abarcan los regímenes de líquido normal y sobreenfriado. En este contexto hemos trabajado en colaboración con el Prof. Francesco Sciortino, de la Universidad de Roma, La Sapienza. Así luego, en una segunda parte, a partir del conocimiento obtenido del comportamiento tanto dinámico como estructural del agua líquida en ambos regímenes, nos concentraremos en el estudio del comportamiento del agua en contextos de confinamiento o en superficies, atentos a que el mismo resulta crucial para comprender su rol tanto en el campo de la ciencia de materiales como en biología. Por lo tanto, estudiaremos la estructura y orientación, en función de la temperatura, de moléculas de agua que forman parte de las primeras capas de hidratación de la proteína lisozima y de una superficie modelo como una placa de grafeno. Adicionalmente pondremos en evidencia, a través de diferentes parámetros estructurales, la existencia de una transición líquido-líquido a la temperatura de la transición dinámica incorrectamente asociada con una transición vítrea de proteínas.

En el contexto del agua biológica, nos interesará estudiar el comportamiento del agua de hidratación en los defectos de empaquetamiento de proteínas (dehidrones), los cuales exhiben propensión a la desolvatación, labilizando el agua circundante de modo de promover interacciones intermoleculares. Por lo tanto, el estudio del agua interfacial en proteínas nos conducirá naturalmente a la necesidad de comprender cuales son los principios fundamentales que subyacen a la unión de proteínas. Esta cuestión posee una enorme relevancia intrínseca como problema académico, pero también reviste enorme significación en virtud de los contextos a los que determina, como ser la comprensión de innumerables funciones biológicas que son gobernadas por la unión de proteínas y el diseño racional de fármacos. Aquí también el agua resulta un protagonista fundamental de la escena, ya que la misma actúa como mediador de las interacciones (generalmente no covalentes) entre las superficies complejas interactuantes. En este contexto trabajamos en colaboración con el Prof. Ariel Fernandez. La hipótesis principal se basa en que la imposibilidad de detectar los principios moleculares subyacentes al proceso de unión de proteínas

(*binding*), radica en haber considerado sólo interacciones de a pares o de dos cuerpos, mientras que las interacciones realmente relevantes son las interacciones de tres cuerpos o cooperativas [2]. Ello implica un cambio de paradigma o, cuanto menos, un cambio en el nivel de análisis. Posteriormente utilizaremos dicho conocimiento para desarrollar un concepto de diseño racional de drogas destinadas a bloquear interfases específicas. También utilizaremos estos conceptos para rediseñar compuestos guías y para explicar la naturaleza de una mutación oncogénica en la proteína p53, un sistema de particular relevancia en el estudio del cancer [59].

En la etapa final de esta Tesis estudiaremos el agua de hidratación de superficies modelo como monocapas autoensambladas de alcanos y superficies gráficas. Específicamente, estudiaremos las fluctuaciones de densidad y tiempos de residencia (funciones de autocorrelación) del agua de hidratación de estos sistemas. Estudiaremos el efecto de la curvatura de las superficies sobre estas propiedades, lo cual podría ser de gran utilidad en sistemas reales como los sitios de *binding* de proteínas, los cuales requieren la remoción del agua de hidratación para que el ligando correspondiente ocupe dicho sitio. Finalmente estas medidas nos permitirán estudiar el comportamiento hidrofílico de las superficies gráficas, superficies que por un largo tiempo fueron consideradas como hidrofóbicas. De esta manera, nuestros resultados hacen eco en recientes estudios experimentales que sugieren una mayor interacción grafeno-agua. El estudio de estos sistemas simples con una química y geometría controladas nos permiten dar los primeros pasos para sentar las bases de una teoría de la hidrofobicidad a nivel de la nanoescala y para encontrar elementos de diseño de relevancia tanto en biología como en ciencia de materiales.



## Capítulo 2

# Agua bulk

En este capítulo exploramos la robustez y validez de un método para determinar moléculas de agua estructuradas y desestructuradas, tanto en el régimen de líquido normal como sobreenfriado. Determinamos la fracción de moléculas estructuradas a distintas temperaturas y en particular demostramos que la fracción de este tipo de moléculas puede ser descrita por un factor de Boltzmann, lo cual está de acuerdo con las predicciones de un modelo de doble estado para el agua líquida.

### 2.1. Agua líquida, agua vítrea y agua sobreenfriada

En el estado sólido, el agua presenta polimorfismo y, como vimos en el capítulo 1 sección 1.1, poliamorfismo (polimorfismo amorfo), lo cual es considerado como una de sus tantas anomalías. El primero de estos fenómenos es un concepto importante en sistemas de un solo componente, reconocido hace aproximadamente 200 años a partir de los ejemplos de sales de carbonatos, fosfato y arseniato. El último es un concepto relativamente nuevo, reconocido hace aproximadamente 25 años con el ejemplo del agua. Estos hechos a menudo se relacionan con la hipótesis de la posible existencia de más de una fase líquida de composición  $\text{H}_2\text{O}$ , y la posibilidad de un (segundo) punto crítico líquido-líquido [45] (ver capítulo 1 sección 1.1).

El agua vítrea, también llamada hielo amorfo, puede existir cuando la temperatura disminuye por debajo de la temperatura de transición vítrea  $T_g$  (aproximadamente 130 K a 1 bar)(ver

capítulo 1 subsección 1.2.2). Aunque el agua vítrea es un sólido, su estructura exhibe un arreglo desordenado al estilo de un líquido. Como mencionamos anteriormente, el hecho de que el agua presente poliamorfismo hace que este material puro, en el estado vítreo pueda existir en más de un estado amorfo [45], donde basicamente estos estados son tres: LDA (low density amorphous ice), HDA (high density amorphous ice) y VHDA (very high density amorphous ice). Sin embargo, probablemente un sin fin de diferentes estados relajados, que difieren, por ejemplo, en términos de funciones de distribución radial o densidad, se pueden obtener experimentalmente, aunque como se evidencia en la referencia [45], estos últimos corresponden a subestados de una de las tres formas mencionadas anteriormente. Por otro lado, se pueden encontrar en la literatura, visiones conflictivas sobre la pregunta de cuantas fases o estados amorfos existen, pero no cabe duda que las fases HDA y LDA son de gran interés en el estudio de sistemas vítreos, ya que al parecer representan el primer ejemplo conocido de una transición de primer orden entre dos sólidos amorfos [1].

## 2.2. Análisis cuantitativo de un modelo de doble estado para agua

El agua líquida presenta diversas anomalías que se vuelven más importantes a medida que es sobreenfriada por debajo de su temperatura de cristalización. Estas anomalías fueron asociadas tentativamente a la presencia de dos tipos de estructuras locales preferenciales compitiendo, con moléculas caracterizadas por tener una alta o baja densidad local [140, 145, 133, 111, 142, 134, 135, 27], en un escenario inspirado, como vimos anteriormente, por la existencia de al menos dos formas diferentes de estados vítreos amorfos, llamados LDA y HDA [162, 163, 127, 88, 21].

### 2.2.1. Detectando moléculas estructuradas y desestructuradas

Para determinar las moléculas de agua estructuradas (baja densidad local) y desestructuradas (alta densidad local) estudiaremos en esta Tesis, por medio de simulaciones de dinámica molecular, un sistema de  $N = 216$  moléculas que interactúan mediante el modelo SPC/E [161] (ver sección 1.2.1) en una caja cúbica, usando condiciones periódicas de contorno. Las cargas

eléctricas parciales de los átomos de las moléculas de agua,  $q_1$  y  $q_2$  (ver sección 1.2.1), interactúan mediante el método de reacción de campo [166], siempre y cuando la distancia  $r$  entre ellas sea mayor a  $2,5\sigma$ . Además los pares de átomos de oxígeno del sistema interactúan via la energía potencial de Lennard-Jones con los parámetros  $\sigma$  y  $\epsilon$  detallados en la sección 1.2.1, tabla 1.1 para el modelo de agua SPC/E, siempre y cuando la distancia  $r$  entre ellos sea menor a  $2,5\sigma$ .

En 1996 Shiratani y Sasai [145] notaron, a partir de simulaciones de dinámica molecular de agua líquida y sobreenfriada amorfa, que ciertas moléculas mostraban radialmente la existencia de un espacio no ocupado en el intervalo  $3,2 < r/\text{Å} < 3,8$ . Este efecto puede verse en la Fig. 2.1 con ayuda de la función de distribución radial  $g(r)$  [3, Capítulo 2, Sección 2.2., Subsección 2.2.1.] (entre oxígenos) para una trayectoria de dinámica inherente y  $T = 210$  K ( $T_m = 215$  K para el modelo SPC/E) para la correlación promedio entre cualquier molécula y una molécula sin vecinas en el intervalo  $3,2 < r/\text{Å} < 3,8$  (línea fina negra). Incluimos para su comparación el comportamiento promedio entre cualquier par de moléculas (línea gruesa gris) [3, pág. 71]. La dinámica inherente implica que cada estructura o configuración de una dinámica molecular se la minimiza por medio de algún método como *steepest descent* o gradiente conjugado de modo de encontrar el mínimo local o cuenca de atracción a la que pertenece dicha configuración. Esto permite eliminar el efecto de las distorsiones que provoca la energía térmica, de modo de enfocarnos únicamente en los aspectos estructurales. Como veremos, el uso de la dinámica inherente permite una mejor identificación de moléculas estructuradas y desestructuradas.

Para cuantificar este efecto, estos autores crearon el índice de estructura local  $I(\mu, t)$  de la molécula  $\mu$  en la configuración al tiempo  $t$ . Para calcular dicho índice debe conocerse la distancia  $r_i$  entre el oxígeno de la molécula  $i$  y el oxígeno de la molécula  $\mu$ . Luego las moléculas se numeran en orden de distancia a partir de la molécula  $\mu$ ;  $r_1 < r_2 < \dots < r_i < r_{i+1} < \dots < r_{n(\mu, t)} < r_{n(\mu, t)+1}$ . El valor  $n(\mu, t)$  se elige de manera que  $r_{n(\mu, t)} < 3,7 \text{ Å} < r_{n(\mu, t)+1}$ . Luego el índice de estructura local se calcula como,

$$I(\mu, t) = \frac{1}{n(\mu, t)} \sum_{i=1}^{n(\mu, t)} [\Delta(i; \mu, t) - \bar{\Delta}(\mu, t)]^2 \quad ,$$

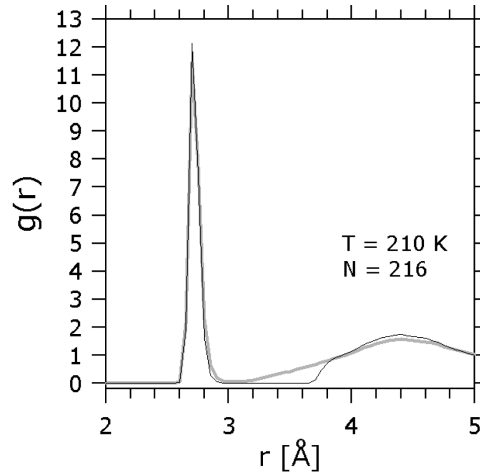


Figura 2.1: Función de distribución radial  $g(r)$  entre oxígenos de moléculas de agua de estructura amorfa de un sistema con  $N = 216$  y  $T = 210$  K en una trayectoria de dinámica inherente. Correlación entre cualquier par de oxígenos (línea gruesa gris) y correlación entre cualquier oxígeno y oxígenos sin vecinos en el intervalo  $3,2 < r/\text{Å} < 3,8$  (línea fina negra). El pico agudo debajo de los 3 Å corresponde a la ubicación preferencial de los primeros cuatro vecinos del tetraedro que forman estos últimos junto con la molécula de agua central. Adaptado de [3, pág. 73].

donde  $\Delta(i; \mu, t) = r_{i+1} - r_i$  y  $\bar{\Delta}(\mu, t)$  es el promedio sobre  $\Delta(i; \mu, t)$ ,

$$\bar{\Delta}(\mu, t) = \frac{1}{n(\mu, t)} \sum_{i=1}^{n(\mu, t)} \Delta(i; \mu, t) \quad .$$

Las moléculas que no poseen moléculas vecinas en la región  $3,2 < r/\text{Å} < 3,8$ , lo cual resulta en un alto valor de  $I(\mu, t)$ , se consideran estructuradas. Por el contrario, moléculas con vecinas en dicha región tienen un bajo valor de  $I(\mu, t)$ , y se consideran no-estructuradas. La presencia de moléculas vecinas en dicha región resulta en una perturbación de la coordinación tetraédrica local. La ausencia de vecinas a partir de cierta molécula entre los mencionados radios implica que la densidad local sea menor y que la estructura tetraédrica con centro en la citada molécula sea más regular [3, pág. 72].

Varios índices se han propuesto, además del índice de estructura local  $I$ , para identificar moléculas estructuradas y desestructuradas con un éxito parcial, hasta que recientemente nuestro grupo introdujo una nueva aproximación, usando el índice propuesto por Shiratani y Sasai

junto con minimizaciones de energía potencial. De esta manera pudimos cuantificar estos dos tipos de “especies” de moléculas de agua [27]. A diferencia de ref. [145], calculamos dicho índice en las estructuras inherentes (IS) (minimizando la energía potencial de las correspondientes estructuras instantáneas o configuraciones de la dinámica real) para deshacernos del efecto aleatorio introducido por las vibraciones térmicas presentes en la trayectoria real, las cuales dificultan la posibilidad de identificar correctamente la estructura local [126].

La Fig. 2.2 muestra la distribución de  $I(\mu, t)$  para la dinámica inherente y real (inset) en el rango de temperaturas  $210 \text{ K} \leq T \leq 320 \text{ K}$  para una densidad constante  $\rho = 1,0 \text{ g/mL}$ . Una inspección directa de dicha figura nos permite observar que para todas las temperaturas  $T$  estudiadas las distribuciones de  $I(\mu, t)$  son claramente bimodales (esta fue la primera vez que se encontró un índice de estructura que genere una distribución claramente bimodal, permitiendo determinar sin ambigüedades, la presencia de dos “clases” de moléculas de agua [27]). El primer pico corresponde a valores de  $I(\mu, t)$  relacionados con moléculas de alta densidad local (estructura tetraédrica distorsionada por la presencia de una quinta molécula); mientras que el segundo se relaciona con moléculas de baja densidad local o estructuradas, cuya amplitud se incrementa a medida que  $T$  decrece. Ambos picos se encuentran separados por un mínimo invariante (independiente de la temperatura, una especie de punto isobéstico) en  $I(\mu, t) \sim 0,14 \text{ \AA}^2$ . Este resultado indica que el agua, tanto en el régimen de líquido normal como sobreenfriado, consiste en una mezcla de moléculas en dos estados, donde la fracción de moléculas estructuradas o de baja densidad local, aumenta con la disminución de  $T$ . Para el caso de la dinámica real (inset Fig. 2.2), no existe una separación neta en dos picos. Solo a las temperaturas más bajas podemos ver un cambio de curvatura cuando  $I(\mu, t) \sim 0,14 \text{ \AA}^2$  [21].

Otro índice, comunmente utilizado en la literatura para detectar ambos tipos de “especies” de moléculas de agua, es el índice tetraédrico de Errington y Debenedetti:

$$q(\mu, t) = 1 - \frac{3}{8} \sum_{j=1}^3 \sum_{k=j+1}^4 \left( \frac{1}{3} + \cos \psi_{jk}(\mu, t) \right)^2 ,$$

donde  $\psi_{jk}(\mu, t)$  es el ángulo formado por las líneas que conectan el oxígeno de la molécula  $\mu$  con aquellos de sus moléculas más próximas  $j$  y  $k$  al tiempo  $t$ . El rango de valores que puede tomar

$q(\mu, t)$  es  $-3 \leq q \leq 1$ . Luego, un alto valor de  $q$  implica una molécula con buen orden tetraédrico local mientras que un valor bajo indica una coordinación tetraédrica local distorsionada [125]. Sin embargo  $q(\mu, t)$  no presenta la bimodalidad que muestra  $I(\mu, t)$ , incluso tampoco al nivel de estructuras inherentes (IS) [27].

Asumiendo que (para cualquier  $T$ ) la distribución de  $I(\mu, t)$  es el resultado de la suma de dos distribuciones diferentes (una para moléculas estructuradas y otra para moléculas desestructuradas), y dadas las pendientes pronunciadas de la curva a ambos lados de  $I(\mu, t) \sim 0,14 \text{ \AA}^2$  ( $I_{\min}$ ), la superposición entre las dos distribuciones sería muy pequeña. Así, el área debajo del segundo pico de la Fig. 2.2 representa una estimación de la fracción de moléculas estructuradas  $\xi$  presentes en la muestra. Esta fracción  $\xi$  se calcula integrando la curva  $P(I)$  entre  $I_{\min}$  e infinito. En la subsección 2.2.4 estudiaremos la dependencia de  $\xi$  con  $T$ .

Shiratani y Sasai [145] notaron la existencia de una transición y de configuraciones inestables junto con ciertas discrepancias entre la información dada por la dinámica inherente y la dinámica real. Ellos notaron que a los tiempos en que una molécula está en una configuración de transición entre ambos estados (estructurado y desestructurado), tiende a alternar, en la dinámica real, entre valores de  $I(\mu, t)$  moderadamente altos y moderadamente bajos. De acuerdo a la ref. [27] es evidente que la minimización de la energía potencial conducirá a estas configuraciones a uno de los dos estados, dependiendo de la proximidad a cada uno de ellos. Como se desprende claramente de la Fig. 2.2, el proceso de minimización conduce a la configuración actual al mínimo local, a la parte inferior de su cuenca de atracción, y así solo los estados estructurados y no estructurados son estables (mínimo de baja energía o alta energía, pero en ambos casos, mínimo local), mientras que los estados de transición son inestables. Así la existencia de configuraciones de transición no es un problema para esta descripción de doble estado.

Para demostrar lo expuesto anteriormente, dividimos arbitrariamente la distribución de  $I(\mu, t)$  para la dinámica real de  $T = 210 \text{ K}$  en tres porciones: moléculas con  $I(\mu, t) < 0,075 \text{ \AA}^2$ ;  $0,075 \text{ \AA}^2 < I(\mu, t) < 0,12 \text{ \AA}^2$  y  $I(\mu, t) > 0,12 \text{ \AA}^2$ . Así podemos hablar de moléculas con valores bajos, medios y altos de  $I(\mu, t)$ . Consideramos luego, para las moléculas de bajo  $I(\mu, t)$ , las correspondientes configuraciones inherentes y calculamos la distribución de  $I(\mu, t)$ . De la misma manera se calculan las distribuciones de  $I(\mu, t)$  de la dinámica inherente para las moléculas

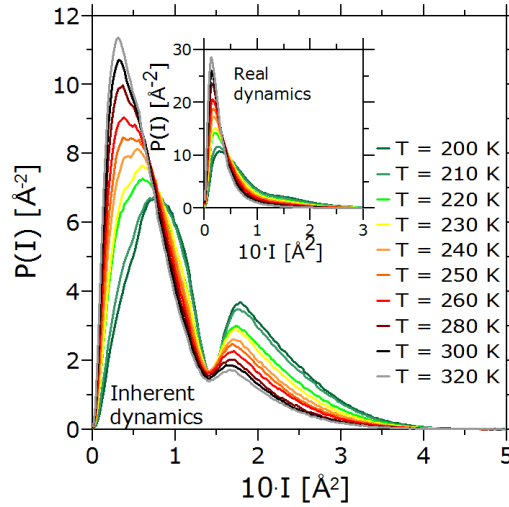


Figura 2.2: Distribución de  $I(\mu, t)$  para la dinámica inherente y real (inset) para un rango de temperaturas  $T$  dentro del régimen de líquido normal y sobreenfriado y densidad  $\rho = 1,0 \text{ g/mL}$ .

las con valores de  $I(\mu, t)$  medios y altos, discriminados previamente en la dinámica real. Estos resultados se pueden observar en la Fig. 2.3, donde la línea sólida corresponde a la distribución de  $I(\mu, t)$  para la dinámica real a  $T = 210\text{K}$ , mientras que las curvas con círculos muestran las distribuciones de  $I(\mu, t)$  de la dinámica inherente para las tres clases de moléculas. La curva con círculos negros representa el caso de las moléculas con bajo  $I(\mu, t)$ , la curva con círculos grises oscuros corresponde a las moléculas con valores medios de  $I(\mu, t)$ , mientras que la de círculos grises claros corresponde al caso de las moléculas con alto  $I(\mu, t)$  [21]. Vale la pena destacar que las moléculas con  $I(\mu, t)$  medio en la dinámica real, están equitativamente particionadas entre ambos estados (estructurado o baja densidad local y desestructurado o alta densidad local) en la dinámica inherente. Las configuraciones de transición corresponden mayoritariamente a este último tipo de moléculas y así representan la porción minoritaria de todas las configuraciones. Por otro lado el 79 % de las moléculas de bajo  $I(\mu, t)$  caen antes del mínimo ( $I_{min}$ ) en la distribución IS de  $I(\mu, t)$ , mientras que el 76 % de las moléculas de alto  $I(\mu, t)$  caen dentro del pico estructurado en la distribución IS.

Finalmente vemos que a pesar de la importancia dinámica de las configuraciones de transición, podemos ver que tanto las moléculas estructuradas como desestructuradas son capturadas tanto por las distribuciones de  $I(\mu, t)$  real e inherente, lo cual avala nuestra descripción de un

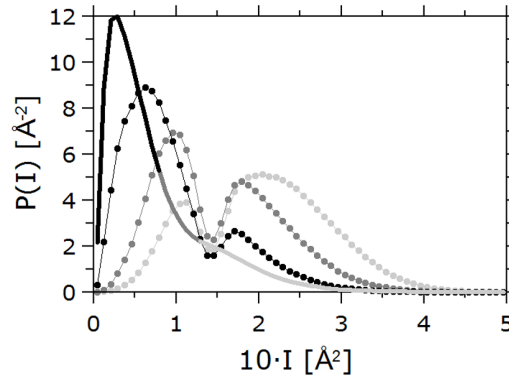


Figura 2.3: Distribución de  $I(\mu, t)$  para  $T = 210\text{K}$  y  $\rho = 1,0\text{g/mL}$ . Consideramos tres grupos de moléculas en la dinámica real con  $I(\mu, t) < 0,075 \text{ \AA}^2$  (línea gruesa negra);  $0,075 \text{ \AA}^2 < I(\mu, t) < 0,12 \text{ \AA}^2$  (línea gruesa gris oscura) y  $I(\mu, t) > 0,12 \text{ \AA}^2$  (línea gruesa gris clara). Luego para cada grupo de moléculas graficamos las correspondientes distribuciones de  $I(\mu, t)$  (misma escala de grises con círculos) en la dinámica inherente.

modelo de doble estado [21].

### 2.2.2. Robustez del índice de estructura local $I$

Este índice usa un radio de corte  $r_{\text{th}} = 3,7 \text{ \AA}$  (subsección 2.2.1) un tanto arbitrario para determinar los vecinos que se van a considerar en orden para realizar el cálculo, incluso cuando es cuidadosamente elegido de la posición del mínimo entre los primeros dos picos en la función de distribución radial  $g(r)$ . Por lo tanto decidimos determinar si los resultados son todavía válidos para cambios moderados en ese radio de corte. En la Fig. 2.4 se muestra la distribución del índice de estructura local para una densidad unitaria y  $T = 210\text{K}$ , pero cuando  $I(\mu, t)$  es calculado con los radios de corte  $r_{\text{th}} = 3,6 \text{ \AA}$  y  $r_{\text{th}} = 3,8 \text{ \AA}$ , además de la distribución original con  $r_{\text{th}} = 3,7 \text{ \AA}$ . Podemos ver que en ambos casos la distribución continua siendo claramente bimodal (esto ocurre también con las otras temperaturas estudiadas). Sin embargo, ahora es interesante recalcar que la posición del mínimo ( $I_{\text{min}}$ ) varía. Esto significa que la estimación de la fracción de moléculas estructuradas y desestructuradas dependerá de el radio de corte  $r_{\text{th}}$ . Aun así, como veremos más adelante, la dependencia funcional de  $\xi$  con  $T$  no depende de este radio de corte.

Otro punto de interés es si la bimodalidad en la distribución de  $I(\mu, t)$  emerge solo en el



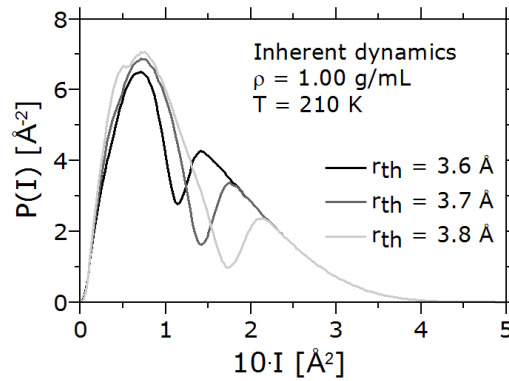


Figura 2.4: Distribución de  $I(\mu, t)$ , en la dinámica inherente, para diferentes valores de  $r_{th}$  para  $T = 210$  K y  $\rho = 1,0$  g/mL.

caso de densidad unitaria o si persiste para otros valores de densidad. Por lo tanto realizamos el análisis para otros valores de densidades  $\rho$  y temperaturas  $T$ . La Fig. 2.5 a) muestra los resultados para las nueve temperaturas estudiadas y  $\rho = 0,9$  g/mL, mientras que la Fig. 2.5 b) muestra datos para  $T = 210$  K y  $\rho$  variable. Nuevamente vemos que la bimodalidad no se ve afectada, aunque la población de ambos estados se modifica notablemente, más aún, vemos que el valor de  $I_{min}$  practicamente no se ve afectado. Como es esperable, la fracción de moléculas estructuradas crece a menor densidad.

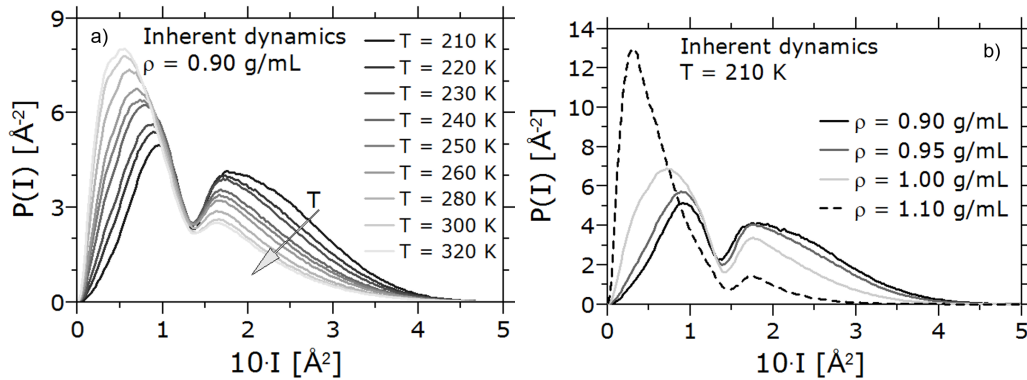


Figura 2.5: a) Distribución de  $I(\mu, t)$  para las nueve temperaturas estudiadas y densidad  $\rho = 0,9$  g/mL y b) Distribución de  $I(\mu, t)$  para las cuatro densidades estudiadas y  $T = 210$  K.

### 2.2.3. Parámetros de calidad de HBs para moléculas estructuradas y desestructuradas

Típicamente cada molécula de agua tiende a estar rodeada por sus cuatro primeras vecinas (subsección 1.2.1) con las cuales forma HBs. Debemos tener en cuenta que estamos utilizando una técnica de dinámica inherente, la cual refina la estructura y así, la geometría local se verá perfeccionada si comparamos con configuraciones instantáneas. Sin embargo, el índice de estructura local  $I(\mu, t)$  claramente discrimina entre moléculas estructuradas y desestructuradas en las configuraciones inherentes, por lo tanto esperamos diferencias en los parámetros geométricos locales para ambos tipos de moléculas. Para ello calculamos, para cada  $\rho$  y  $T$ , las distribuciones de parámetros de puente de hidrógeno, tanto para las moléculas estructuradas como para las desestructuradas. Uno de los parámetros es la distancia en HBs  $r_{\text{HB}}$ . Para calcularla medimos la distancia entre el oxígeno de una molécula de agua  $i$  y el oxígeno de una molécula  $j$ , para  $N$  pares de moléculas de agua. Entonces, si una molécula de agua  $i$  forma un HB a través de uno de sus átomos de hidrógeno con una molécula  $j$ , consideramos el menor ángulo  $A$ , aquel formado entre las líneas que unen a  $O_i$  con  $H_i$  y  $H_i$  con  $O_j$  (un HB perfecto presentará un ángulo de  $0^\circ$ ). Dado que el proceso de minimización perfecciona la geometría local, para calcular las distribuciones de distancias y ángulos consideramos que cada molécula de agua forma HBs con sus primeras cuatro moléculas vecinas (de hecho, esto se ve evidenciado en las distribuciones de  $r_{\text{HB}}$  y  $A$  obtenidas). En la Fig. 2.6 graficamos  $\langle r_{\text{HB}} \rangle$  para moléculas estructuradas y desestructuradas en función de  $T$ , para las cuatro densidades estudiadas. Como era de esperar  $\langle r_{\text{HB}} \rangle$  es menor para densidades mayores. Este valor es claramente dependiente de  $T$  y decrece a medida que  $T$  disminuye, ya que la geometría de los HBs se perfecciona con el enfriamiento. Si prestamos atención a las curvas para densidad unitaria, vemos que los HBs a cualquier  $T$  fija poseen un  $\langle r_{\text{HB}} \rangle$  menor (HBs más fuertes) para las moléculas estructuradas que para las no estructuradas [21].

En la Fig. 2.7 mostramos la desviación estándar  $\sigma_A$  del valor medio del menor ángulo  $A$  de la distribución de moléculas estructuradas y no estructuradas, en función de  $T$ , para las cuatro densidades estudiadas (desviaciones mayores de  $A = 0^\circ$  hacen que el HB sea más débil).

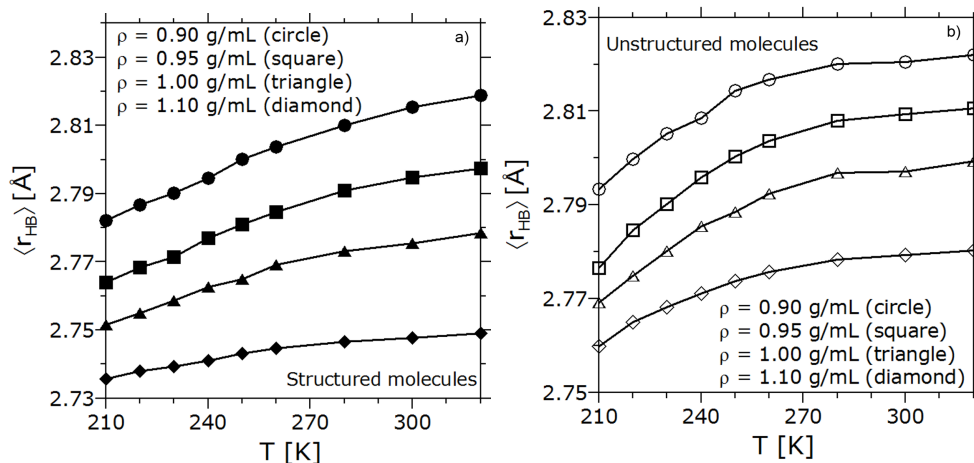


Figura 2.6: Distancia media en HBs,  $\langle r_{\text{HB}} \rangle$ , para moléculas a) estructuradas y b) desestructuradas, en función de  $T$ , para las cuatro densidades estudiadas.

La desviación estándar  $\sigma_A$  se define como  $\{n^{-1} \sum_{i=1}^n [A_i - \langle A \rangle]^2\}^{1/2}$  donde  $n$  es el número de HBs,  $A_i$  es el ángulo en el  $i$ -ésimo HB y  $\langle A \rangle$  es  $n^{-1} \sum_{i=1}^n A_i$ . Como se esperaba, las moléculas estructuradas poseen una menor desviación que aquellas no estructuradas. Todos los resultados son consistentes con aquellos obtenidos para  $\langle r_{\text{HB}} \rangle$  en HBs [21].

#### 2.2.4. Dependencia con la temperatura de la fracción de moléculas estructuradas $\xi$ : testeando las predicciones del modelo de doble estado

En esta subsección vamos a cuantificar la fracción de moléculas estructuradas  $\xi$  a las diferentes temperaturas y densidades estudiadas. Estos datos luego se compararan con las predicciones de un modelo de doble estado [142, 134, 135, 21].

Entre los diferentes modelos de doble estado de agua que han sido propuestos, parece interesante hacer una breve referencia a dos descripciones distintas recientes: *el modelo de mezcla* [150, 139, 155, 141, 112] y *el modelo de parámetro de doble orden* [142, 134, 135]. Mientras que en el primero de los casos el agua se supone que existe en uno de los dos motivos estructurales (como por ejemplo LDA y HDA, dos sólidos amorfos), el segundo se refiere al agua como una mezcla de unos pocos arreglos moleculares estructurados y de un estado desestructurado que puede tener muchas configuraciones locales diferentes. Ambos modelos consideran la existencia de un estado estructurado compuesto por uno, o unos pocos, arreglos moleculares. Así, la

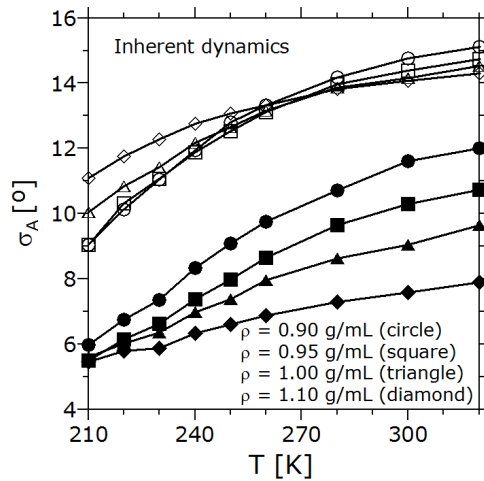


Figura 2.7: Desviación estándar  $\sigma_A$  de la distribución del menor ángulo  $A$  en HBs, para moléculas estructuradas (símbolos llenos) y desestructuradas (símbolos vacíos), en función de  $T$  para las cuatro densidades estudiadas.

principal diferencia consiste en que el *modelo de parámetro de doble orden* considera un estado desestructurado constituido por un set de diferentes configuraciones, lo que implica una mayor entropía, mientras que el *el modelo de mezcla* considera que este último estado es más bien único (no necesariamente desestructurado, pero sí con una estructura local diferente). El uso de la dinámica real para calcular el índice de estructura local, como se hizo hasta el momento, no ayuda mucho al respecto. Esto se debe a que en la dinámica real las vibraciones térmicas inducen una distorsión de la estructura local preferida, en consecuencia, haciendo más difícil distinguir entre ambos tipos de estados. Nuestros resultados basados en las dinámicas inherentes (donde el efecto aleatorio introducido por las fluctuaciones térmicas es explícitamente removido, llevando cada configuración a su estructura local energéticamente preferida, es decir, a un mínimo local de su cuenca de atracción) pueden darnos evidencia relevante. En primer lugar, la forma de los picos de las Figs. 2.2 y 2.5 parece mostrarnos que para todas las temperaturas y densidades estudiadas, el estado desestructurado, y también el estructurado, no es representado por una única configuración, sino más bien por diferentes configuraciones locales. También es interesante el hecho de que la posición de la abcisa en el máximo del pico estructurado y el mínimo que separa ambos picos ( $I_{\min}$ ), es completamente independiente de  $T$  y de  $\rho$ , mientras que la posición de la abcisa en el máximo del pico desestructurado, se mueve significativamente en el rango

de temperaturas estudiadas. Por lo tanto, estos hallazgos son en principio, consistentes con un estado estructurado que implica la contribución de unas pocas y bien definidas configuraciones locales, y un estado desestructurado representado por un estado de elevada entropía, derivada de las diferentes configuraciones preferenciales con una baja estructura local.

Otra predicción del *modelo de parámetro de doble orden* propuesto por el Prof. H. Tanaka, es que la fracción de moléculas estructuradas (para valores bajos de esta fracción) puede ser descrita por un factor de Boltzman (donde el peso de Boltzman expresa la estabilización, debida a la energía del enlace de hidrógeno de la estructura de HBs localmente favorecidos, y la desestabilización, proporcional a la presión, derivada del concurrente aumento del volumen [142, 134, 135]). En consecuencia, la fracción de moléculas estructuradas mostraría una dependencia exponencial con la recíproca de la temperatura absoluta. Para testear esta predicción, para las diferentes densidades estudiadas, calculamos la fracción de moléculas estructuradas  $\xi$ , integrando la distribución correspondiente de  $I(\mu, t)$  entre  $I_{\min}$  y  $+\infty$ . En la Fig. 2.8 graficamos, para cada  $\rho$ , la curva para el  $\log \xi$  vs.  $1/T$ . Podemos ver, que en todos los casos, los datos pueden ser ajustados por una dependencia lineal, como lo demanda el *modelo de parámetro de doble orden*.

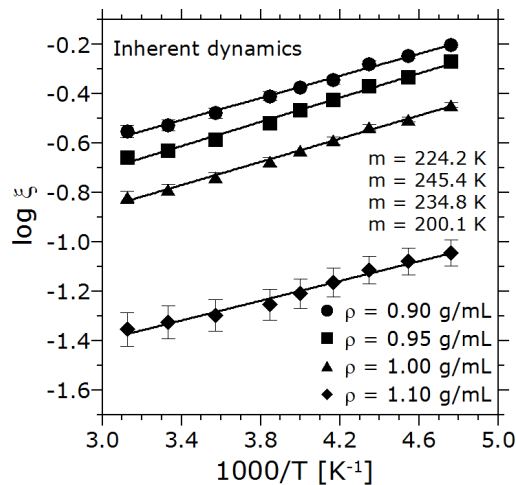


Figura 2.8:  $\log \xi$  en función de la temperatura absoluta recíproca para los distintos valores de  $\rho$  estudiados. Se incluye también el ajuste lineal (líneas sólidas) de los cuatro conjuntos de datos. Se muestran las pendientes  $m$  para cada uno de los ajustes, ordenadas de arriba hacia abajo respetando el mismo ordenamiento que el de las densidades. Se incluyen también las barras de error correspondientes a una cuantía de error positiva y negativa del 5%.

Comparando las pendientes  $m$ , de las tendencias lineales, para las densidades 0,90, 0,95 y 1,00 g/mL, vemos que el coeficiente de variación porcentual se encuentra dentro del 10 %, aunque incluyendo el valor para  $\rho = 1,10$  g/mL los valores se encuentran dentro del 20 %. Vale la pena señalar que cada uno de los ajustes lineales se encuentra dentro del  $\pm 5\%$  con respecto a sus correspondientes datos.

En la Fig. 2.9 mostramos los mismos datos reportados en la figura anterior, pero ahora los datos se ajustan con una tendencia exponencial. Nuevamente, la predicción del modelo de doble estado es consistente con estos datos. Sin embargo, estos datos también pueden ser ajustados por dos tendencias lineales (una para elevados y otra para bajos regímenes de  $T$ ) y donde cada par de líneas, para cada una de las densidades, se cruzan a  $T \sim 250$  K.

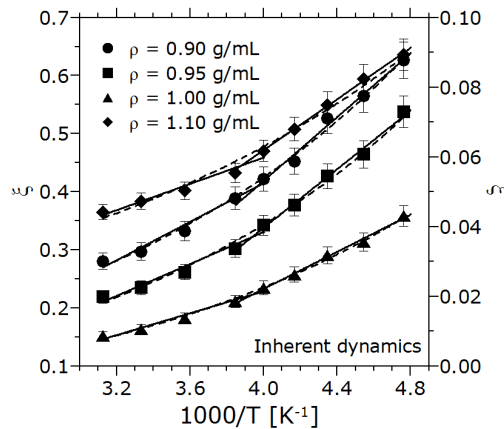


Figura 2.9: Fracción de área  $\xi$  debajo de la porción estructurada de la distribución de  $I(\mu, t)$  en función de la recíproca de la temperatura absoluta para las cuatro densidades estudiadas. Las barras de error corresponden a una cuantía de error de  $\pm 5\%$ . Eje izquierdo:  $\rho \leq 1,00$  g/mL; eje derecho:  $\rho = 1,10$  g/mL. Se incluye un ajuste exponencial de los datos (curvas segmentadas). Para cada una de las cuatro densidades también se muestran dos ajustes lineales (líneas sólidas), uno para valores bajos de  $T$  y otro para valores altos de  $T$ . Para cada  $\rho$  estas líneas se cruzan a  $T \sim 250$  K.

Previamente mostramos en la subsección 2.2.2 que cambios moderados en el radio de corte  $r_{th}$  en el cálculo de  $I(\mu, t)$  no afectan la bimodalidad de la distribución de  $I(\mu, t)$  (Fig. 2.4), pero sí producen un desplazamiento de la posición del mínimo ( $I_{min}$ ) entre los dos máximos. Por lo tanto resulta interesante estudiar la robustez de la dependencia de  $\xi$  con  $T$ , con respecto

a este último parámetro. Como anticipamos en la subsección 2.2.2, podemos afirmar que la dependencia funcional de  $\xi$  es la misma para los tres valores de  $r_{\text{th}}$  estudiados. Esto se pone en evidencia en la Fig. 2.10, donde allí graficamos nuevamente el  $\log\xi$  contra  $1/T$  para  $\rho = 1,0$  g/mL y los tres radios de corte mencionados anteriormente.

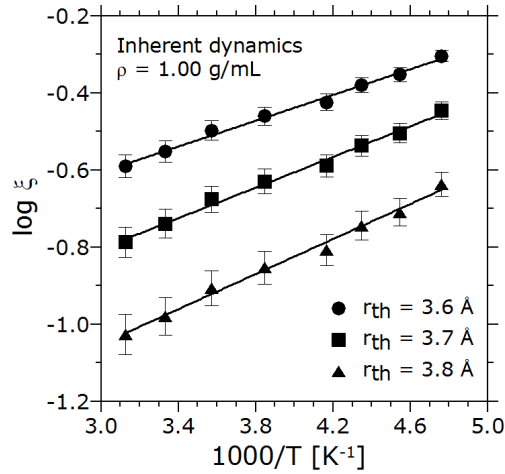


Figura 2.10:  $\log\xi$  en función de  $1/T$  para  $\rho = 1,0$  g/mL y los tres valores de  $r_{\text{th}}$  estudiados. Las barras de error corresponden a una cuantía de error positiva y negativa del 5%. Se incluye el ajuste lineal de cada set de datos (líneas sólidas).

## 2.3. Conclusiones

En este Capítulo exploramos la robustez y validez (para diferentes densidades y temperaturas) de un método para determinar moléculas de agua estructuradas y desestructuradas y determinamos la calidad (geométrica) de la estructura local de ambos tipos de moléculas. Observamos que la dependencia con la temperatura de la fracción de moléculas estructuradas puede ajustarse con una exponencial negativa, lo cual brinda un sustento a modelos *two state* que postulan que la misma puede ser gobernada por un peso de Boltzmann debido a la ganancia energética al perfeccionar las interacciones locales al estructurarse y al costo dado en el incremento local del volumen. El trabajo presentado en este capítulo fue realizado en colaboración con el Prof. Francesco Sciortino, del Departamento de Física de la Universidad de Roma, La Sapienza. El trabajo que realizamos con el índice de estructura local sujeto a la dinámica inhe-

rente ha sido continuado por el grupo de A. Nilsson [44] y nuestros resultados han sido también utilizados con posterioridad por el Prof. H. Tanaka en la extensiones del modelo de parámetro de doble orden [34].



## Capítulo 3

# Agua de hidratación y la transición vítreo en proteínas

En este capítulo estudiamos la estructura y orientación, en función de la temperatura, de moléculas de agua que forman parte de las primeras capas de hidratación de la proteína lisozima. Comparamos los resultados obtenidos con los de las primeras capas de hidratación de una superficie modelo, una placa de grafeno, cuyos resultados fueron obtenidos previamente por nuestro grupo de investigación [1, pág.78] [22, 24]. También ponemos en evidencia, a través de diferentes parámetros estructurales, la existencia de una transición líquido-líquido (cruce de la línea de Widom, cruce del hipotético punto crítico líquido-líquido del agua) a la temperatura de la transición dinámica incorrectamente asociada con una “transición vítreo de proteínas”.

### 3.1. Agua de hidratación o biológica

Las proteínas son las biomoléculas más versátiles y diversas, y tienen un inmenso interés científico dado el rol preponderante que tienen en los mecanismos biológicos. Las mismas están presentes en una infinidad de sistemas y realizan funciones que se destacan en los siguientes aspectos: estructurales, reguladores, transportadores, inmunológicos, enzimáticos, contráctiles, homeostáticos, transducción de señales y protectores o defensivos [1, pág. 95]. Por otro lado, las proteínas carecen de actividad en ausencia del agua de hidratación y además ésta representa el

solvente natural tanto *in vivo* como *in vitro* [1, pág. 95]. Las moléculas de agua forman una parte integral en la mayoría de las interacciones proteína-proteína, proteína-ADN y proteína-ligando, ayudando en el reconocimiento mutuo y participando tanto de la termodinámica como de la cinética en la unión. La fuerza impulsora de la unión no solo depende de la interacción de una unidad biológica con la otra, sino que también del costo energético necesario para remover el agua de hidratación y la ganancia energética del subsiguiente arreglo molecular de estas mismas moléculas de agua [9]. Por lo tanto, la comprensión de la estructura y el comportamiento del agua de hidratación es fundamental para diversos campos de investigación en proteínas (estabilidad, plegado, unión y función biológica), sin embargo dicha comprensión todavía no llega a ser completa. El agua *bulk* experimenta un cambio, en su arreglo molecular, cerca de la superficie de una proteína. Dicha superficie contiene tanto residuos hidrofílicos como hidrofóbicos. Mientras que los residuos hidrofílicos contienen principalmente átomos polares y con carga, los residuos hidrofóbicos contienen átomos no polares. El agua puede formar interacciones por puente de hidrógeno (HB) con los residuos polares mientras que los residuos hidrofóbicos no son capaces de establecer dicha interacción con el agua. Otro factor importante sobre la superficie de las proteínas es su compleja y heterogénea topografía, es decir que la geometría local de la superficie puede ser diferente en diferentes regiones de la proteína. Estos factores (heterogeneidad superficial de carga y topografía) se espera que tengan un impacto crucial en la estructura local de las moléculas de agua en la superficie de una proteína. Con el fin de comprender mejor estos aspectos, las simulaciones de dinámica molecular juegan un papel importante, y tanto la hidratación de superficies modelo, como de proteínas ha sido estudiada [17, 20]. Dichos estudios muestran que las primeras capas de hidratación de agua tienden a ser más estructuradas que el *bulk* y (dependiendo del tipo de superficie) posiblemente asuman una orientación preferencial con respecto al vector normal a la superficie.

### 3.2. Estudio del agua de hidratación de la lisozima

A continuación estudiamos la dependencia con la temperatura, de la estructura de las primeras capas de hidratación de agua de la proteína lisozima. También analizamos la orientación

y tendencia a formar puentes de hidrógeno (HBs) de dichas moléculas de agua. La lisozima es una enzima presente en las lágrimas y la saliva en donde actúa como una barrera frente a las infecciones. También es muy abundante en la clara de huevo, de donde se extrae para su uso industrial, en particular para el control de las bacterias lácticas en los vinos. La lisozima fue descubierta por A. Fleming en 1922 y caracterizada por primera vez por D. C. Phillips en 1965, quién consiguió una imagen de 2 Å de resolución. La lisozima también está presente en el bazo, los pulmones, los leucocitos, el plasma, la leche y el cartílago. La deficiencia en lisozima, debida a mutaciones en el gen *LYZ* situado en el cromosoma 12, ha sido asociada a displasias esqueléticas y a un aumento de la propensión a las infecciones. La proteína simulada fue la lisozima en la forma ortorómbica, [165] cuya conformación estructural fue extraída del Protein Data Bank (PDB: 2LZH).

### 3.2.1. Modelado y equilibración de los sistemas lisozima-agua y grafeno-agua

El modelo de agua utilizado en las simulaciones lisozima-agua y grafeno-agua, cuyos resultados presentamos a continuación, es TIP5P [131] (ver detalles del modelo en la subsección 1.2.1). Todas las simulaciones, tanto para la lisozima como para la placa de grafeno, se llevaron a cabo con el paquete de simulación AMBER 10 [83] en el ensamble  $NpT$  y utilizando el termostato de Langevin para controlar la temperatura de los sistemas [156]. Se trabajó con un time step de 2 fs y se usó un cutoff de 8 Å para calcular las interacciones de largo alcance. Se llevaron a cabo simulaciones a diferentes temperaturas (dentro del régimen de líquido normal y líquido sobreenfriado), 1 bar de presión promedio y una densidad media de 1,0 kg/dm<sup>3</sup>. La forma ortorómbica de la lisozima de la clara de huevo de gallina (PDB:2LZH) fue solvatada en una caja de simulación hexaédrica de 12600 moléculas de agua TIP5P, utilizando condiciones periódicas de contorno en las tres dimensiones y el algoritmo de simplificación SHAKE [157]. Se utilizó el campo de fuerza ff99SB de AMBER, el cual fue especialmente diseñado para simular moléculas biológicas en solventes acuosos [84]. Para las simulaciones de la placa de grafeno se utilizó el campo de fuerza GAFF (*Generalized Amber Force Field*) de AMBER y simultáneamente el ff99SB para llevar a cabo las simulaciones en agua [98]. Los átomos de carbono de la placa se modelaron como partículas de Lennard-Jones sin carga (las densidades de carga no son nulas

solo en los bordes). Los parámetros  $\sigma_{CC}$  y  $\epsilon_{CC}$  corresponden a carbonos  $sp^2$  en el campo de fuerza GAFF. La placa se fijó por medio de un potencial armónico del tipo  $k\Delta x$ , donde la constante  $k$  vale  $500 \text{ kcal/mol } \text{Å}^2$  y  $\Delta x$  corresponde al desplazamiento respecto a la posición inicial o de referencia. La superficie fue centrada en la caja de simulación paralela al plano  $XY$ . La placa de grafeno consta de 10 por 10 anillos de benceno (aproximadamente  $400 \text{ Å}^2$ ), la misma fue solvatada con 10788 moléculas de agua TIP5P, utilizando condiciones periódicas de contorno en las tres dimensiones y SHAKE. Con el fin de evitar a las moléculas de agua cercanas a los bordes de la placa, para el análisis consideramos solamente a las moléculas que se encuentran dentro de un cilindro de radio  $r = 8,4 \text{ Å}$ , normal al plano  $XY$ , cuyo centro coincide con en el centro de la placa de grafeno [22]. En ambos sistemas las equilibraciones se llevaron a cabo en dos pasos, utilizando en ambos el termostato de Langevin, un time step de 2 fs y el algoritmo SHAKE: en la primera etapa en el ensamble canónico  $NVE$  se llevó el sistema a la temperatura de trabajo; luego se realizó una equilibración más larga en el ensamble  $NpT$  hasta obtener una densidad media de  $1,0 \text{ kg/dm}^3$ .

### 3.2.2. Estructura, orientación y puentes de hidrógeno del agua de hidratación

En este capítulo no utilizamos la dinámica inherente, como en el capítulo 2 subsección 2.2.1, para caracterizar la estructura local de las moléculas de agua, ya que en este caso nos basta con los resultados arrojados por la dinámica real y no hay necesidad de recurrir a técnicas computacionales más costosas. Sin embargo, mientras que el índice de estructura  $I$  [145] puede ser empleado directamente para agua en interfases (como vimos en el capítulo 2 subsección 2.2.1,  $I$  mide la calidad de las interacciones locales independientemente de cuantas sean), el índice  $q$  [125] no puede ser utilizado con su formulación original (la cual requiere de la evaluación de las posiciones de los primeros cuatro vecinos, ver capítulo 2 subsección 2.2.1) y debe ser reformulado de manera de poder adaptarlo a una interfase [25]. Dado a que las moléculas de agua de la primera capa de hidratación poseen preferentemente tres moléculas de agua vecinas, definimos el índice  $q_3$  como sigue:

$$q_3(\mu, t) = 1 - \frac{3}{4} \sum_{j=1}^2 \sum_{k=j+1}^3 \left( \frac{1}{3} + \cos \psi_{jk}(\mu, t) \right)^2 ,$$

El uso de este tipo de índices (como  $I$  o  $q_3$ ) para caracterizar la estructura local de las moléculas de agua en una interfase (en vez de estudiar perfiles de densidad y otros parámetros ya usados que proveen menos información) también es relevante ya que pueden establecer un punto de conexión con ciertas propiedades dinámicas (por ejemplo han sido útiles en la determinación, por parte de nuestro grupo de investigación, de la existencia de un nexo causal entre la estructura y la propensión dinámica del agua vítrea, un tema de gran interés en la física de sistemas vítreos [28, 30, 31, 3]).

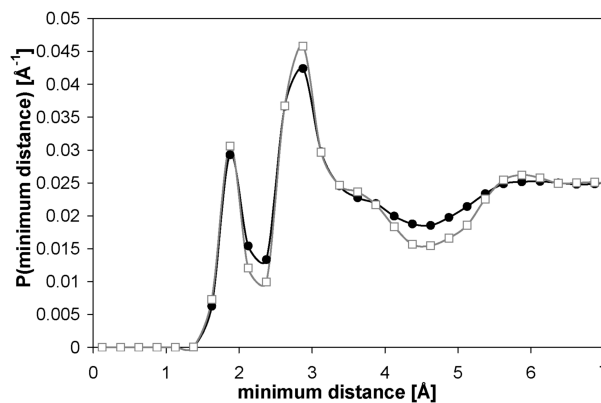


Figura 3.1: Distribución de distancias mínimas de las moléculas de agua a la proteína (al átomo más cercano de la proteína) a  $T = 300$  K (línea negra, círculos) y  $T = 240$  K (línea gris, cuadrados).

En la Fig. 3.1 se muestra la distribución de distancias mínimas de diferentes moléculas de agua a la proteína lisozima para  $T = 300$  K y  $T = 240$  K. Como se esperaba, a la temperatura más baja los picos se definen mejor, lo cual nos indica que la estructura de las moléculas de agua se ve favorecida con el descenso de la temperatura. El primero de los picos, ubicado por debajo de los  $2 \text{ \AA}$ , es típico en proteínas y se atribuye a la presencia de puentes de hidrógeno agua-proteína. Luego podemos observar un pico más grande, cuyo máximo se encuentra justo antes de los  $3 \text{ \AA}$  seguido de un hombro, un mínimo, y finalmente un pico mucho más ancho. El mismo análisis efectuado a una placa de grafeno (perfiles de densidad normales al plano  $XY$  de la placa) muestran el comienzo de un primer pico por encima de los  $2,5 \text{ \AA}$ , extendiéndose hasta los  $4,25 \text{ \AA}$ . A este pico le sigue un mínimo profundo que se extiende hasta alrededor de los  $5,5 \text{ \AA} - 6 \text{ \AA}$  y luego un pico pequeño [22]. Para poder caracterizar cuantitativamente la

estructura del agua de hidratación de la lisozima, calculamos la densidad de probabilidad del índice de estructura local  $I(\mu, t)$ ,  $P(I)$ , para las moléculas de agua ubicadas entre 2,125 Å y 4,25 Å de la superficie, y para las moléculas del *bulk* (aquellas cuya distancia a la superficie es mayor que 10 Å y por lo tanto se encuentran lo suficientemente lejos como para no sentir ningún efecto de la superficie) [25]. Con esta elección evitamos el pico pequeño correspondiente a las interacciones de puente de hidrógeno proteína-agua, ya que la estructura de este pequeño número de moléculas podría ser consecuencia de esta particular interacción (estudiaremos este grupo de moléculas más adelante). El cálculo de  $P(I)$  para las moléculas de agua en una región determinada, se lleva a cabo calculando el índice en muchas configuraciones que ocurren a diferentes tiempos a lo largo de una extensa trayectoria (así siempre utilizamos configuraciones instantáneas acumulando estadística de muchas configuraciones diferentes, que garantizan que todas las moléculas de agua consideradas residen dentro de la región de interés).

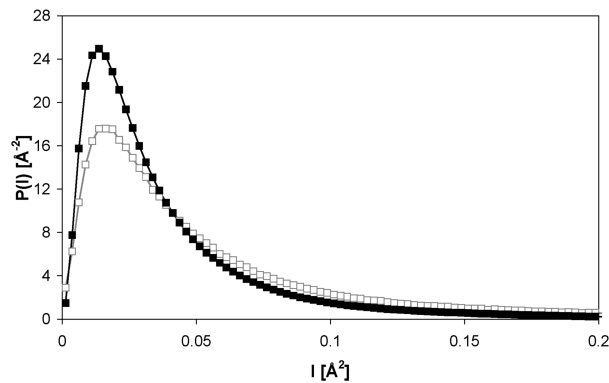


Figura 3.2: Densidad de probabilidad  $P(I)$  de encontrar una molécula de agua con un índice de estructura local  $I$  para la región localizada entre 2,125 Å y 4,25 Å de la superficie de la proteína lisozima (línea gris) y para las moléculas de agua del *bulk* (moléculas de agua a más de 10 Å de distancia de la superficie; línea negra).  $T = 300$  K en ambos casos.

Los resultados se muestran en la Fig. 3.2. Vemos que las moléculas de agua cercanas a la superficie de la proteína son más estructuradas que las moléculas del *bulk* [145]. Este resultado es similar al encontrado por nuestro grupo para el caso de una placa de grafeno [22]. Es importante aclarar que un valor más elevado de  $I$  para las moléculas de la superficie que para las del *bulk*, no es debido a una mejor coordinación de HBs de las moléculas de la superficie (de hecho, hemos

chequeado que la geometría de los HBs es similar). El punto más importante aquí es que la primera capa de coordinación de una dada molécula está bien separada de su segunda capa de coordinación por un *gap* de distancia considerable (el índice  $I$  es sensible a la presencia de dicho *gap* y, consecuentemente, al decrecimiento de la densidad local), este hecho sera aclarado más adelante.

A continuación, en la Fig. 3.3 se muestran las distribuciones normalizadas de la distancia  $d$  a los primeros seis vecinos de las moléculas de agua localizadas entre 2,125 Å y 4,25 Å de la superficie de la lisozima (líneas grises) y para las moléculas del *bulk* (líneas negras)

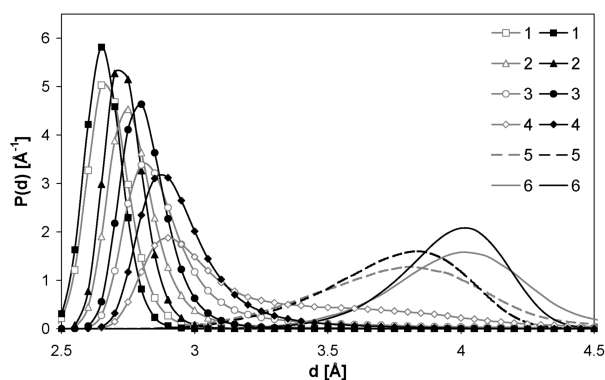


Figura 3.3: Densidad de probabilidad, para  $T = 240$  K, de encontrar distancias  $d$  de los primeros seis vecinos para moléculas cercanas a la superficie de la lisozima (entre 2,125 Å y 4,25 Å, líneas grises) y para las moléculas del *bulk* (líneas negras). Los números en la leyenda (1, 2, 3, ...) indican el número de vecino considerado (primero, segundo, tercero, ...).

Mostramos solo los resultados para  $T = 240$  K ya que las moléculas de agua se ven mejor estructuradas a temperaturas bajas. Sin embargo para temperaturas superiores los resultados son similares, presentando distribuciones más anchas con máximos más pequeños (estos datos no se muestran). También obtuvimos resultados similares para la placa de grafeno (estos datos no se muestran). Un resultado interesante que nos muestra la Fig. 3.3 es que las moléculas cercanas a la superficie tienden a estar rodeadas solo por tres moléculas de agua más cercanas, en vez de cuatro moléculas, como es el caso del agua *bulk*. Se pueden distinguir dos regiones de picos de esta figura, una región alrededor de los 3 Å y una segunda región aproximadamente a unos 4 Å. Para las moléculas del *bulk*, los primeros cuatro vecinos están localizados en la primera de estas regiones (primera capa de coordinación), donde estas moléculas pueden formar HBs

con la molécula central, mientras que los vecinos cinco y seis pertenecen a la segunda región (segunda capa de coordinación). Sin embargo, para las moléculas cercanas a la superficie, la distribución para el cuarto vecino posee una extensa cola (insinuando una bimodalidad), lo cual nos indica que algunas moléculas no pueden formar HBs con su cuarto vecino (se consideran distancias razonables para enlaces por puente de hidrógeno en el agua hasta 3,5 Å). Así estas moléculas de agua poseen su cuarto vecino más cerca de la segunda capa de coordinación.

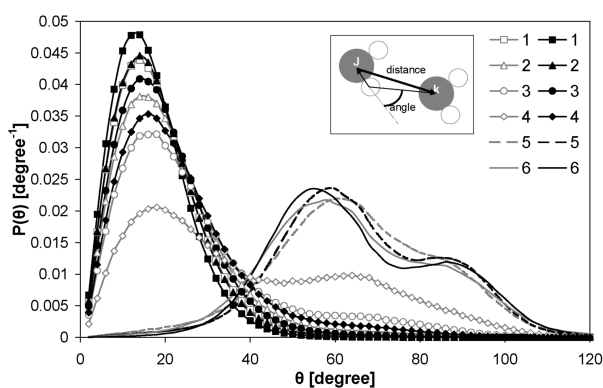


Figura 3.4: Densidad de probabilidad, para  $T = 240$  K, del ángulo suplementario O-H...O (definido en el inset,  $\theta$ ) de los primeros seis vecinos para moléculas cercanas a la superficie de la lisozima (entre 2,125 Å y 4,25 Å, líneas grises) y para las moléculas del *bulk* (líneas negras). Los números en la leyenda (1, 2, 3, ...) indican el número de vecino considerado (primero, segundo, tercero, ...).

Del mismo modo en la Fig. 3.4 se pueden observar las distribuciones normalizadas del ángulo mínimo  $\theta$  entre el enlace O-H y el H...O (ver inset en la figura) para una molécula de agua central y sus seis primeras vecinas. Nuevamente podemos ver que las moléculas del *bulk* (líneas negras) tienden a formar puentes de hidrógeno con sus primeros cuatro vecinos (ángulos cercanos a  $0^\circ$ ). Estos resultados son también para la temperatura  $T = 240$  K, dado a que el sistema posee mejor estructura que a temperaturas más elevadas. A mayores temperaturas algunos cuartos vecinos no poseen ángulos apropiados para la formación de enlaces por puente de hidrógeno con la molécula central (estos datos no se muestran). Por otro lado las moléculas de la superficie (líneas grises) muestran una notable bimodalidad en la distribución correspondiente al cuarto vecino. El segundo pico ancho que se puede observar arroja valores de  $\theta$  superiores a  $40^\circ$ , lo cual es inconsistente con la formación de puente de hidrógeno. En conclusión, podemos afirmar que



gran parte de las moléculas cercanas a la superficie de la proteína tienden a estar rodeadas por tres moléculas, en vez de cuatro moléculas de agua vecinas como es típico en el *bulk*, mientras que el cuarto vecino es excluido hacia la segunda capa de coordinación.

A continuación examinaremos los puentes de hidrógeno agua-agua y la orientación de las moléculas de agua cercanas a la superficie de la lisozima y en algunos casos particulares compararemos los resultados con aquellos obtenidos previamente para una placa de grafeno. Para identificar los HBs utilizamos un criterio geométrico (que en estudios anteriores [22, 29, 27, 28] ha demostrado rendir resultados razonables y consistentes con criterios más elaborados): Un HB implica que la distancia O-O sea menor que 3,5 Å y el ángulo O-H...O sea menor a 40° (el ángulo óptimo para un HB sería de 0°). En la Fig. 3.5 a) podemos observar la distribución de mínimas distancias a la proteína, para las moléculas que se encuentran entre 2,125 Å y 4,25 Å, discriminando entre aquellas que participan en cuatro o más HBs agua-agua (curva gris) de aquellas que forman solo tres o menos HBs (curva negra). Cuando se realizan estos mismos estudios para la placa de grafeno [22], es evidente que la región correspondiente al primer pico del perfil de densidad del agua, normal a la placa (entre 2,5 Å y 4,25 Å), es básicamente el producto de dos regiones distintas (la primera y segunda capa de hidratación), que pueden ser llamadas *1ra capa* y *2da capa* de hidratación [22]. La *1ra capa* de hidratación, que se extiende desde 2,5 Å hasta 3,5 Å, esta poblada principalmente de moléculas con tres HBs, mientras que la *2da capa*, que se extiende desde 3,5 Å hasta 4,25 Å, contiene preferentemente moléculas de agua con cuatro HBs [22]. Si bien menos neta que para el caso del grafeno, la situación para la lisozima es similar, lo cual lo denota la Fig. 3.5 a). El cruce de las dos curvas ocurre alrededor de los 3 Å, que de ahora en más lo consideraremos el límite entre las dos capas de hidratación más allá del pequeño primer pico de la Fig. 3.1 (que de ahora en más lo consideraremos como la *1ra capa* de hidratación): la *2da capa* (extendiéndose desde 2,125 Å hasta 3 Å) y la *3ra capa* (entre 3 Å y 4,25 Å). La Fig. 3.5 b) nos muestra que en la *2da capa* (curva gris) predominan las moléculas con tres HBs, mientras que la *3ra capa* (curva segmentada) muestra un comportamiento más típico del *bulk* (cuatro HBs agua-agua).

Un hecho interesante surge de analizar la Fig. 3.5 c), donde discriminamos a las moléculas de agua enlazadas por puente de hidrógeno a la proteína, entre moléculas aceptoras (donde el HB

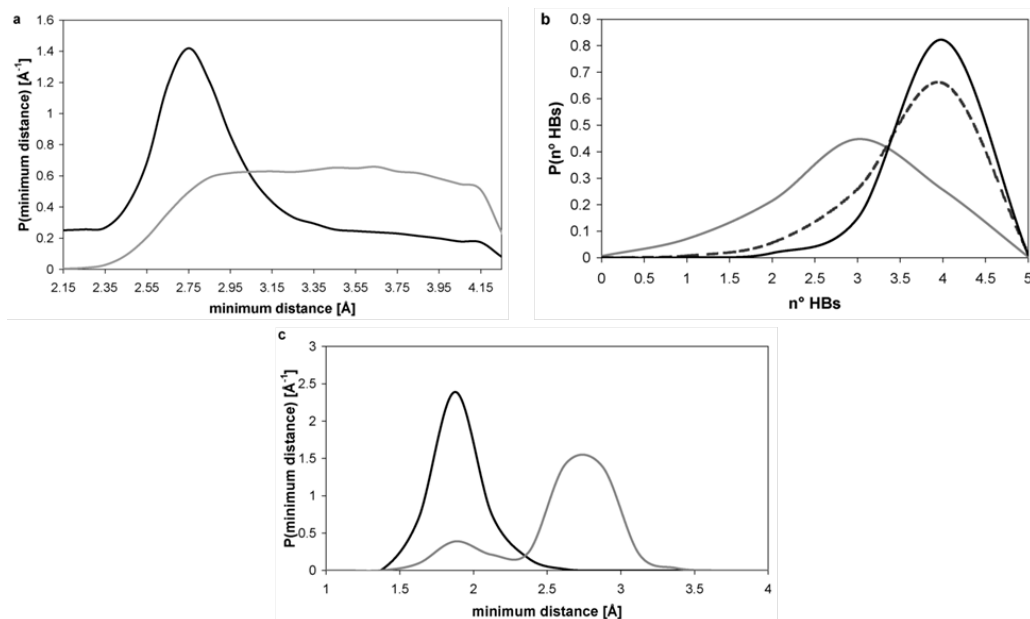


Figura 3.5: a) Distribución normalizada de distancia a la proteína lisozima a  $T = 240$  K para moléculas con tres o menos HBs (línea negra) y moléculas con cuatro o más HBs (línea gris). Vale decir que los resultados son idénticos si consideramos moléculas con estrictamente tres y moléculas con estrictamente cuatro HBs agua-agua. b) Distribución del número de enlaces HB agua-agua formados por moléculas localizadas en la *2da capa* (entre 2,125 Å y 3 Å de la superficie de la proteína, curva gris), *3ra capa* (entre 3 Å y 4,25 Å, curva segmentada), y para las moléculas del *bulk* (a más de 10 Å de la superficie, curva negra). c) Distribución normalizada de mínimas distancias, a cualquier átomo de la proteína, para moléculas enlazadas por puente de hidrógeno a la proteína, tanto como donores (curva gris) o aceptores (curva negra).

con la lisozima es formado via un átomo de hidrógeno de la proteína, curva negra) y donoras (el HB se forma via un átomo de hidrógeno de la molécula de agua, curva gris). Encontramos dos distribuciones de distancias mínimas diferentes con picos en 1,875 Å y 2,75 Å respectivamente. Por ende, podemos afirmar que la *1ra capa* posee mayoritariamente moléculas de agua actuando como aceptoras en el enlace por puente de hidrógeno proteína-agua, mientras que aquellas que se enlazan como donoras en el HB se encuentran preferentemente en la *2da capa* de hidratación.

Para poder comprender mejor como es la estructura de las primeras capas de hidratación estudiaremos a continuación la orientación de las moléculas de agua que forman parte de cada una de estas capas.

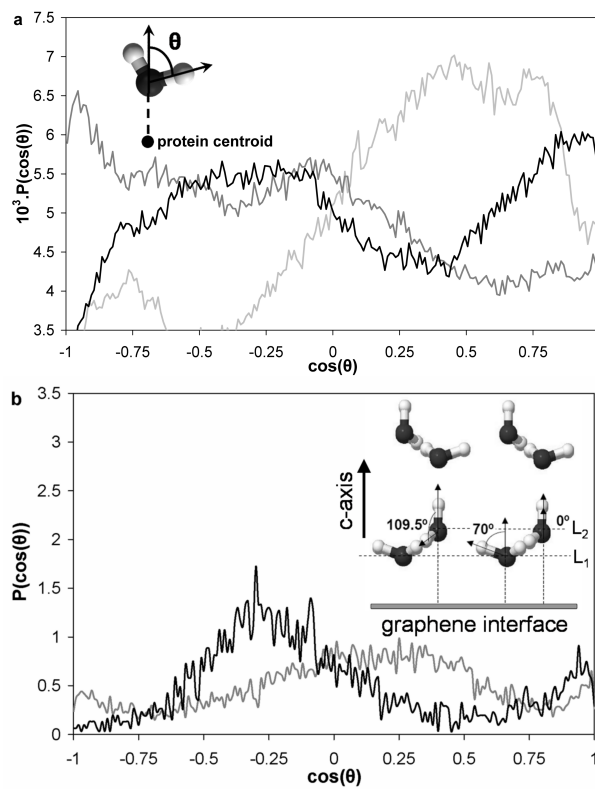


Figura 3.6: a) Distribución normalizada del  $\cos(\theta)$  para la *1ra capa* (gris claro), *2da capa* (gris oscuro), y *3ra capa* (negro) de hidratación alrededor de la proteína lisozima. b) Ídem para la placa de grafeno, donde el comportamiento de la *1ra capa* corresponde a la línea gris y el de la *2da capa* a la línea negra. En ambos sistemas el modelo de agua utilizado es TIP5P. El inset muestra la orientación de la *1ra capa* y *2da capa* en la placa de grafeno si se satisface un arreglo perfecto tipo hielo Ih.

En la Fig. 3.6 a) presentamos este estudio para la *1ra*, *2da* y *3ra capa* de hidratación de la lisozima (para comparar, la figura Fig. 3.6 b) muestra este mismo estudio sobre la placa de grafeno). Graficamos la densidad de probabilidad del  $\cos(\theta)$  donde  $\theta$  es ahora el ángulo formado entre los enlaces O-H de las moléculas de agua y el vector que une el centroide de la lisozima con el correspondiente átomo de O de la molécula de agua. Esto obviamente representa una aproximación ya que funcionaría satisfactoriamente si la forma de la molécula fuera completamente esférica y cuyo centroide coincidiera con el centro de la esfera. Este parámetro es interesante dado a que es fácil de comparar con datos experimentales de espectroscopía vibracional [86].

En los estudios previos de nuestro grupo sobre el grafeno [22], encontramos que el agua de

hidratación presenta un notable ordenamiento preferencial (en este caso medido respecto a la normal a la superficie), a diferencia de la situación que ocurre en el *bulk*. La capa de hidratación del grafeno, entre 2,5 Å y 3,5 Å (*1ra capa*), muestra un pico correspondiente a ángulos de 70°, mientras que la capa entre 3,5 Å y 4,25 Å (*2da capa*) presenta picos para ángulos de 110° y 0° [22] (ver Fig. 3.6 b)). Las moléculas de la *1ra capa* tienden a formar un HB con aquellas moléculas de la *2da capa*, solo formando tres HBs debido a la restricción geométrica que impone la superficie (imposibilidad de completar el tetrahedro en la dirección de la placa de grafeno) [22]. Por otro lado las moléculas de la *2da capa* también forman un HB con las moléculas bajo el segundo pico del perfil de densidad (localizado a alrededor de los 6 Å - 7 Å de la superficie) con este HB alineado con el vector normal a la superficie [22]. Este arreglo es típico del hielo Ih. La situación es similar a aquella en la interfase agua-aire, pero a diferencia de ese caso, donde las moléculas de agua en la primera capa de hidratación orientan un enlace O-H hacia el aire [86], aquí la superficie de grafeno impide dicha orientación. Así, alrededor del grafeno, la estructura de hielo Ih local crece en la dirección contraria a lo largo del eje *c* del plano basal [22, 86] (ver inset Fig. 3.6 b)).

Este impedimento para las moléculas de agua en la vecindad del grafeno, de poder orientar un O-H contra la superficie del grafeno ( $\theta = 180^\circ$ ) y la imposibilidad de formar cuatro HBs provocan la exclusión de sus cuartos vecinos hacia el segundo pico en el perfil de densidad (a 6 Å - 7 Å de la superficie) [22]. Esta brecha en el perfil de densidad, que promueve una baja densidad local, es necesaria para la formación de una estructura local tipo hielo Ih [22]. El inset en la Fig. 3.6 b) nos muestra un esquema del plano basal (a lo largo del eje *c*) del hielo Ih en contacto con una placa de grafeno, indicándonos la orientación de la *1ra capa* y *2da capa* de hidratación en lo que sería un arreglo perfecto de dicha estructura. A manera de resumen, en el cuadro 3.1 se muestran nuevamente los ángulos  $\theta$  preferenciales de las dos primeras capas en las interfases grafeno-agua y aire-agua, asumiendo que se satisface una orientación perfecta.

La Fig. 3.6 a) muestra la distribución normalizada del ángulo  $\theta$  en las tres regiones ya mencionadas [*1ra capa* (gris claro), *2da capa* (gris oscuro), y *3ra capa* (negro)] alrededor de la proteína lisozima: 0 Å - 2,125 Å, 2,125 Å - 3 Å, y 3 Å - 4,25 Å, respectivamente. La primera región incluye al primer pico en la distribución de distancias mínimas a la proteína (Fig. 3.1)

$\theta$	Grafeno-Agua	Aire-Agua
1ra capa	70°	180° y 70°
2da capa	110° y 0°	110° y 0°

Cuadro 3.1: Orientaciones preferenciales (ángulos  $\theta$  en grados sexagesimales) en las interfases Grafeno-Agua y Aire-Agua suponiendo una orientación perfecta para una estructura tipo hielo Ih.

poblada mayoritariamente por moléculas de agua que forman un HB, en el rol de aceptores, con la proteína (se enlazan a través de un par de electrones libres del oxígeno a un átomo de hidrógeno de la proteína). Por otro lado, las moléculas de agua enlazadas por un HB, en el rol de donores, se encuentran principalmente dentro de la *2da capa* de hidratación. Las moléculas de agua que forman un HB, como aceptoras con la proteína, tienden a orientar los pares de electrones libres del átomo de oxígeno hacia la superficie de la misma (como las moléculas de la *1ra capa* en el grafeno) y por lo tanto presentan ángulos  $\theta$  cercanos a los 70°, lo cual puede apreciarse en la Fig. 3.6 a) para la *1ra capa* (línea gris claro, entre 0 Å - 2,125 Å de la superficie de la proteína). Adicionalmente, un pequeño pico aparece alrededor de los 160°, el cual es atribuible a un grupo reducido de moléculas de agua enlazadas por un HB, como donoras, vía un átomo de hidrógeno orientado hacia la superficie de la proteína. El resto de las moléculas de agua enlazadas por un HB, como donoras a la proteína, se encuentran en la *2da capa* de hidratación. El caso de la *2da capa* (entre 2,125 Å - 3 Å) es más complicado, y como se puede observar en la Fig. 3.6 a) (línea gris oscuro), los picos son más anchos y bajos (en este sentido, la imperfecta naturaleza esférica de la proteína y la rugosidad de la superficie de la misma conspiran contra el desarrollo de una distribución con buenos picos). Sin embargo podemos ver picos a alrededor de los 160° y 90°. Esta orientación es compatible con la presencia de aquellas moléculas de agua enlazadas como donoras de HB con la proteína (esta orientación es similar a aquella lograda en la interfase aire-agua, donde las moléculas de agua orientan un átomo de hidrógeno hacia la interfase). Adicionalmente, cuando consideramos solo a las moléculas de agua que no forman un HB con la proteína, los valores de  $\theta$  obtenidos son cercanos a aquellos valores obtenidos en la *1ra capa* de las interfases con grafeno y aire, contribuyendo de esta manera a los mismos picos (vale decir que la *1ra capa* en las interfases con grafeno y aire está localizada básicamente a una distancia similar

a la correspondiente a la *2da capa* en la interfase con la proteína). Finalmente la *3ra capa* (entre 3 Å - 4,25 Å) presenta picos para ángulos alrededor de 110° y 0°, lo cual coincide con los ángulos observados para la *2da capa* de hidratación en las interfases con grafeno y aire. Por lo tanto algunas regiones en las primeras capas de hidratación de la lisozima presentan cierta tendencia a un ordenamiento tipo hielo Ih como el observado en la interfase con grafeno; mientras que otras regiones presentan una estructura local tipo hielo Ih pero como aquella observada en la interfase con aire. El ordenamiento local en dichas regiones es compatible con la formación de puente de hidrógeno con la proteína, ya sea como donador o aceptor.

Con el fin de descartar la posibilidad de que la orientación preferencial de las primeras capas de hidratación sea consecuencia del modelo de agua utilizado, en la Fig. 3.7 se muestran las distribuciones del  $\cos(\theta)$  para el agua de hidratación del grafeno cuando se utiliza el modelo de agua TIP3P [Fig. 3.7 a)] y el modelo SPC/E [Fig. 3.7 b)]. El comportamiento del modelo TIP3P es muy similar al del TIP5P. En el caso del modelo SPC/E el comportamiento de la *2da capa* se asemeja mucho a lo observado en los otros dos modelos, sin embargo la *1ra capa* presenta una desviación del pico en 70° hacia valores más cercanos a los 90°.

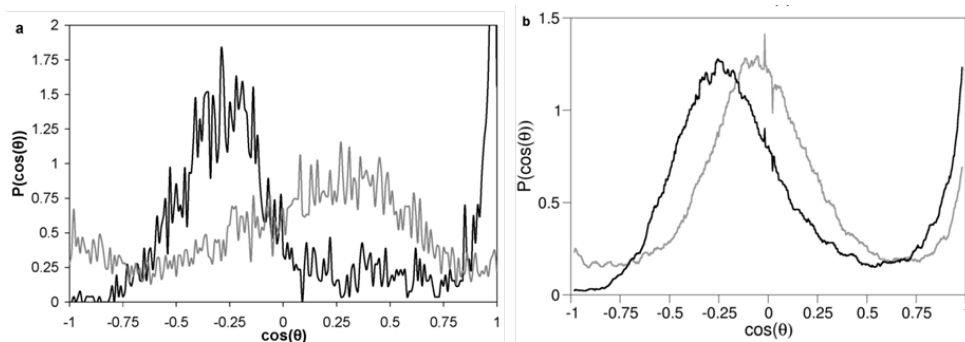


Figura 3.7: a) Ídem a la Fig. 3.6 b) (es decir, para la interfase con el grafeno), pero cuando el modelo de agua utilizado es TIP3P. b) Ídem, pero para el modelo de agua SPC/E. En ambos casos, el gris corresponde a la *1ra capa* y el negro a la *2da capa*.

### 3.2.3. El cambio en la estructura del agua de hidratación con la temperatura y la transición líquido-líquido del agua

A bajas temperaturas, las proteínas experimentan un fenómeno que durante años recibió un considerable interés y se conoció como la “transición vítrea de proteínas” [75, 56, 87]. La “transición vítrea de proteínas” es un fenómeno observado a una  $T \sim 200$  K en las proteínas hidratadas, donde por encima de esta temperatura la proteína presenta movimiento difusivo y por debajo de esta temperatura queda atrapada en modos armónicos. Distintos experimentos y simulaciones por computadora [99, 121] sugieren que esta transición depende fuertemente del solvente, lo que nos lleva a pensar que la transición vítrea en proteínas hidratadas está directamente relacionada con una transición dinámica del solvente que la rodea [1]. Por lo tanto, tomando la hipótesis planteada en el capítulo 2, que propone la existencia de una transición líquido-líquido entre agua líquida de alta densidad (HDL) y agua líquida de baja densidad (LDL), y que dicha transición posee un punto crítico  $C_2$  a  $T \sim 200 \pm 20$  K y  $P \sim 250 \pm 100$  MPa; podemos asociar la transición vítrea de proteínas hidratadas a esta transición líquido-líquido del agua sobreenfriada. Un trabajo reciente [87] investigó la relación existente entre la transición dinámica de biomoléculas (lisozima y ADN) y las propiedades dinámicas y termodinámicas del agua de hidratación. Este relevante trabajo del grupo del Prof. de H. E. Stanley demostró que la llamada transición vítrea de proteínas está mal definida (en dos aspectos): no es una transición vítrea, y además no se debe al comportamiento de la proteína. Al contrario, este fenómeno parece ocurrir en el cruce dinámico en la difusividad del agua de hidratación como resultado de cruzar el hipotético punto crítico líquido-líquido del agua (cruce de la línea de Widom) [87]. Una evidencia presentada a favor de este escenario fue la demostración de que la velocidad de cambio con la temperatura del parámetro ampliamente usado  $q$  (parámetro de orden tetraédrico local),  $dq/dT$ , del agua de hidratación para la lisozima y el ADN presenta un máximo a la temperatura  $T$  de dicho cruce dinámico [87]. Sin embargo, existen dos problemas en relación a estos resultados que nos pueden conducir a conclusiones erróneas [20]: en primer lugar como explicamos al comienzo de este capítulo (ver subsección 3.2.2), el uso del parámetro  $q$ , originalmente definido para agua *bulk*, no es adecuado para el agua interfacial, aún cuando de hecho se lo haya utilizado amplia-

mente en superficies; y en segundo lugar hay un problema con el número de moléculas de agua de hidratación consideradas. El primero de los puntos es el de principal preocupación: como mostramos a lo largo de la subsección 3.2.2 tanto para el grafeno como para la lisozima, la red de HBs del agua se encuentra interrumpida en las cercanías de la superficie, y las moléculas de la primera capa de hidratación tienden a estar enlazadas por HBs con tres moléculas en vez de cuatro como lo es en el caso del agua *bulk*. Este también es el caso de la interfase agua-aire, donde aquí nuevamente las moléculas de agua están coordinadas por HBs con solo tres de sus vecinas [86]. Si calculamos el índice  $q$  para el agua interfacial tal cual es empleado para el agua *bulk* estaríamos forzando una coordinación tetraédrica completa, ya que este índice demanda la presencia de los cuatro primeros vecinos. En este caso, el cuarto vecino posee una orientación angular no tetraédrica y se encuentra a una distancia mayor en comparación con los primeros tres vecinos; de esta manera el cálculo de  $q$  nos rendirá un valor erróneo, más bajo del esperado. El segundo punto se refiere al número de moléculas de agua consideradas para el cálculo. En la Ref. [87] un número total de moléculas de agua  $N = 1242$  TIP5P fueron utilizadas. Este número de moléculas de agua implica considerar alrededor de tres capas de hidratación. Tanto para la placa de grafeno, [25] como también para la lisozima, solo las moléculas de agua bajo el primer pico del perfil de densidad muestran una mejor estructura local con respecto al *bulk* como nos indica el índice de estructura local  $I$  en la Fig. 3.2. Más aun, la dependencia con la temperatura de la estructura local de esta primera capa de hidratación puede ser diferente de aquella correspondiente al resto de las capas de hidratación [22, 25]. En el caso del grafeno y la lisozima, la estructura de las primeras capas de hidratación (*1ra capa* y *2da capa*) es menos dependiente de  $T$  que el resto de las moléculas de agua, cuya estructura disminuye con el aumento de  $T$  [22, 25]. Así en las simulaciones realizadas en Ref. [87] para la lisozima solo alrededor de la mitad de las moléculas de agua consideradas pertenecen a la capa de hidratación biológica, por lo tanto efectos secundarios provenientes de las capas sucesivas de hidratación consideradas pueden estar afectando los resultados presentados en este importante trabajo. Estos problemas fueron solucionados por nuestro grupo en la Ref. [20], donde teniendo en cuenta todo lo expuesto anteriormente, reproducimos las simulaciones para la lisozima realizadas en la Ref. [87] a diferentes temperaturas. En este caso utilizamos una caja de agua de mayor tamaño con  $N =$



12600 moléculas de agua TIP5P.

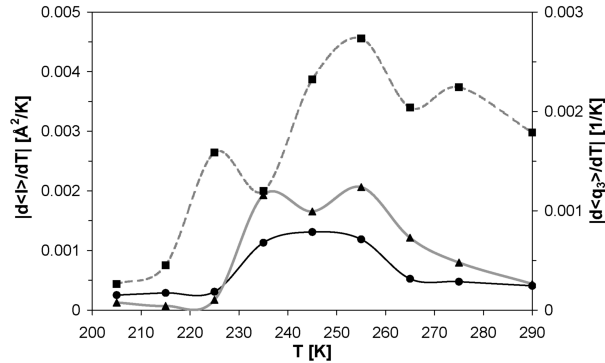


Figura 3.8: Derivada con respecto a la temperatura del índice de estructura local  $I$  para el agua de hidratación de la lisozima (TIP5P): Agua localizada entre 2,125 Å y 4,25 Å de la superficie de la proteína (línea llena negra con círculos) y agua *bulk* (línea llena gris con triángulos). Un máximo en la curva a la temperatura del cruce de la línea de Widom nos indica que la derivada del índice de estructura local  $I$  posee un máximo sobre la línea de Widom. También presentamos la curva correspondiente para la reformulación del parámetro de orden tetraédrico local  $q_3$  para las moléculas de agua superficiales (línea gris segmentada con cuadrados).

En la Fig. 3.8 se muestran los resultados de la derivada de  $I$  con respecto a  $T$ , para moléculas cuya distancia a la proteína se encuentra entre 2,125 Å y 4,25 Å y para las moléculas del *bulk*. Podemos ver que en ambos casos las curvas presentan un máximo alrededor de  $T = 250$  K, el valor de temperatura indicado en Ref. [87] para el cruce de la línea de Widom. Vemos también que ambos tipos de moléculas de agua manifiestan la transición crítica líquido-líquido, dando soporte a la hipótesis propuesta por el grupo de Stanley [75, 87, 158]. Incluimos también en la Fig. 3.8 la derivada con respecto a  $T$  del índice  $q_3$ , es decir el índice  $q$  redefinido de modo de ser apto para tratar agua interfacial. Podemos ver que este índice también presenta un máximo a la temperatura del cruce de la línea de Widom.

A su vez, ahora en la Fig. 3.9 mostramos el valor medio de  $I$  ( $\langle I \rangle$ , línea sólida) y de  $q_3$  ( $\langle q_3 \rangle$ , línea discontinua) en función de  $T$  para las primeras capas de hidratación del grafeno y la lisozima. El cambio de comportamiento en la línea de Widom (alrededor de  $T = 250$  K) es evidente.

Con el fin de tener una mejor idea en cuanto a los cambios estructurales producidos por  $T$

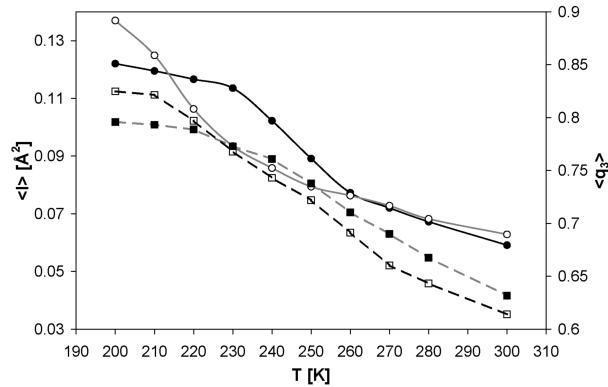


Figura 3.9:  $\langle I \rangle$  (línea negra sólida, círculos llenos) y  $\langle q \rangle$  (línea gris segmentada, cuadrados llenos) en función de  $T$  para moléculas de agua entre 2,125 Å y 4,25 Å de la superficie de la lisozima. La línea gris sólida con círculos vacíos y la línea negra segmentada con cuadrados vacíos muestran el comportamiento correspondiente al agua alrededor del grafeno.

y también evidenciar la presencia de la línea de Widom, mostramos en la Fig. 3.10 la distancia media  $\langle d \rangle$  de los primeros vecinos de las moléculas localizadas entre 2,125 Å y 4,25 Å y para las moléculas del *bulk* alrededor de la lisozima en función de  $T$ . En el caso del agua *bulk*, se observa el comportamiento esperado: a medida que  $T$  disminuye, la primera capa de coordinación (que comprende los primeros cuatro vecinos) se contrae un poco alrededor de la molécula central permitiendo una mejor estructura local con mejores HBs. Sucesivamente, la segunda capa de coordinación se desplaza a mayores distancias de la molécula central, dejando así un vacío o *gap* con respecto a la primera capa de coordinación. Este hecho, responsable de la característica disminución de la densidad durante el enfriamiento, permite que la primera capa de hidratación logre una mejor estructura (todo lo contrario a lo que ocurre cuando aumenta  $T$ , provocando que la segunda capa se aproxime a la primera, perturbando la estructura local de la molécula central). El índice  $I$  es precisamente sensible a la presencia de dicho *gap*. Sin embargo, para la capa superficial la situación es un poco diferente. Incluso cuando la primera capa de hidratación se comprime y se acerca a la molécula central y la segunda capa se aleja, la primera capa solo contiene tres vecinos (a distancias compatibles con la formación de un HB). Sin embargo, el cuarto vecino es expulsado hacia mayores distancias de la superficie, dificultando la formación de un HB con la molécula central. Este efecto geométrico impuesto por la superficie es el resultado

de la imposibilidad de la molécula central de completar su estructura tetrahédrica (dado a que uno de los vértices, aquel que apunta directo a la superficie, no puede ser ocupado). Esta exclusión del cuarto vecino (el cual se encuentra localizado bajo el segundo pico del perfil de densidad, alrededor de  $6 \text{ \AA} - 7 \text{ \AA}$  de la superficie) es necesaria para el crecimiento local de una estructura tipo hielo Ih, como se mostró previamente para el grafeno y para la lisozima. Dicho cuarto vecino generalmente forma un HB con las moléculas de la capa previa. Esto explica también el hecho de que la posición del mismo con respecto a la molécula central y a la superficie es menos sensible con  $T$ . La mayor estructuración del agua superficial respecto del *bulk*, con índices de estructura más típicos de agua de menor densidad resulta consistente también con estudios previos de nuestro grupo que mostraron una menor movilidad de dicha agua de hidratación [24] y con los resultados de Bizarri y Cannistraro [121, 132] anteriormente descriptos.

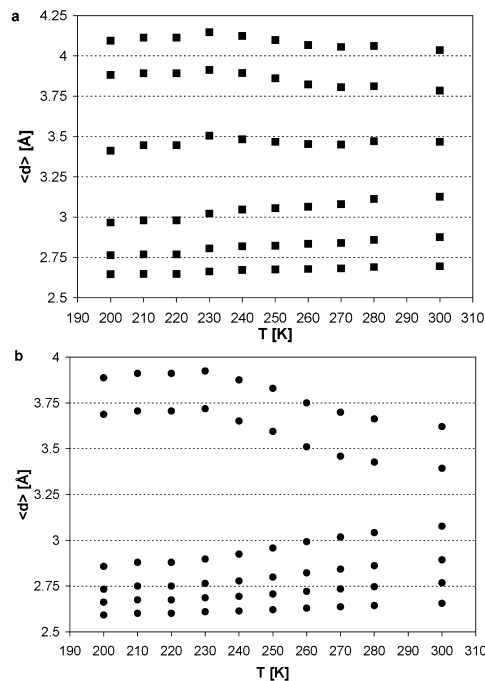


Figura 3.10: Comportamiento con la temperatura de la distancia media  $\langle d \rangle$  a los primeros seis vecinos, de moléculas de agua, a) localizadas entre  $2,125 \text{ \AA}$  y  $4,25 \text{ \AA}$  y b) moléculas del *bulk*. Ambos casos corresponden a las capas de hidratación de la lisozima. En ambos casos el número de vecino puede determinarse por orden de distancia (de abajo hacia arriba: primero, segundo, tercero, y así sucesivamente).

### 3.3. Conclusiones

En este Capítulo estudiamos la estructura y orientación, en función de la temperatura, de moléculas de agua que forman parte de las primeras capas de hidratación de la proteína lisozima y de una superficie modelo como una placa de grafeno. Mostramos, en los dos casos, que estas capas poseen mejor estructura que el *bulk*. La interfase introduce una restricción geométrica que reduce, tanto la coordinación tetraédrica típica del agua como la densidad local de las moléculas de agua adyacentes a la superficie, las cuales tienden a minimizar la pérdida de coordinación formando tres HBs con sus moléculas de agua vecinas. Observamos que el ordenamiento local tetraédrico mejora con el descenso de la temperatura y, en el caso de la placa de grafeno, muestra una tendencia a asemejarse al hielo hexagonal Ih (similar al comportamiento en la interfase agua-aire pero en el sentido contrario a lo largo del eje-c del plano basal). En cambio, el caso de las primeras capas de hidratación de la lisozima parece ser más complicado, pero aun así, encontramos indicios de ambos tipos de comportamientos antes mencionados, junto con una tendencia de algunas de las moléculas de agua más próximas a la superficie a formar puentes de hidrógeno con la proteína. Adicionalmente se puso en evidencia, a través de diferentes parámetros estructurales, la existencia de una transición dinámica líquido-líquido (atribuible al cruce de la línea de Widom, cruce del hipotético punto crítico líquido-líquido del agua) a la temperatura de la transición dinámica incorrectamente asociada con una “transición vítrea de proteínas”.

## Capítulo 4

# Agua de hidratación en los defectos de empaquetamiento de proteínas

En este capítulo establecemos una correlación entre el grado de “arropamiento” o protección de los puentes de hidrógeno de diferentes proteínas y la facilidad de remoción del agua de hidratación circundante a dichas interacciones. Con este fin realizamos cálculos dinámicos, midiendo simplemente la difusividad traslacional para las moléculas de agua que se encuentran dentro de la primera capa de hidratación de la proteína. Específicamente calculamos el desplazamiento cuadrático medio (MSD) de las moléculas de agua que se encuentran en la vecindad, o dominio de desolvatación, de los puentes de hidrógeno del *backbone* de diferentes proteínas [*backbone hydrogen bonds* (BHBs)]. Mostramos que la exposición del *backbone* o cadena principal de las proteínas está correlacionada con la pérdida de hidratación de la superficie, y así estas interacciones intramoleculares expuestas al solvente (dehidrones) están rodeadas por moléculas de agua fácilmente removibles. De esta manera, promoviendo su deshidratación local, estos motivos proporcionan un recurso por el cual un ligando podría desplazar a las moléculas de agua durante el proceso de unión (*binding*).

## 4.1. La asociación de proteínas y el papel del agua de hidratación

El agua de hidratación en los sitios de *binding* de proteínas parece jugar un papel preponderante en el proceso de unión de ligandos y en la asociación de proteínas [72, 81, 51, 118, 114, 2, 113, 82, 79, 106, 48, 26, 18, 19]. Por un lado, los ligandos desplazan el agua de hidratación de los sitios de *binding* [72, 81, 51], y además reemplazan a estas moléculas de agua, supuestamente “desfavorables”, con grupos del mismo ligando complementarios a la superficie de la proteína. Por lo tanto, se espera que estos hechos constituyan la principal fuerza impulsora del proceso de *binding* [81, 51]. De hecho, esta descripción ha mostrado tener validez para una fracción significativa de receptores de interés farmacéutico [81, 51, 92]. Incluso en algunos casos, una porción del sitio activo o de unión del receptor es tan desfavorable para las moléculas de agua que tiende a permanecer prácticamente seco [51]. Así, la estimación de la contribución a la energía libre, involucrada en el desplazamiento de las moléculas de agua cuasi-localizadas con energía libre desfavorable en el sitio activo del receptor, constituye un problema de gran interés en el diseño computacional de drogas [81, 51]. Con la misma filosofía, un estudio reciente [46] de agrupación de fragmentos (*fragment clustering*), de diversos fragmentos orgánicos de prueba sobre la lisozima, mostró que varias regiones de la proteína podían ser blanco de *clusters* de estos fragmentos. Sin embargo, combinando esta estrategia con la exclusión del agua de hidratación (por ejemplo superponiendo estas regiones de *clustering* con el mapa de moléculas de agua fuertemente unidas a la proteína), el sitio de *binding* experimental podría ser predicho satisfactoriamente, dado a que el resto de los sitios de *clustering* coincidirán con sitios que alojan moléculas de agua fuertemente unidas. Esto se debe a que estas últimas moléculas de agua excluirían a los ligandos de la superficie de la proteína, bloqueando así aquellos posibles sitios de *binding*.

Por su parte, para ser estables en agua, las proteínas solubles protegen (o “arropan”) sus puentes de hidrógeno del *backbone* (BHBs) por medio de los grupos carbonáceos de sus cadenas laterales (tal como lo ha determinado el Prof. Ariel Fernández [2, 48, 113, 114, 79]). El efecto de protección indicado es debido a que el grupo carbonáceo como hidrófobo, organiza al agua a su alrededor de manera que esta se agrupa de modo de excluirlo, maximizando el número de

puentes de hidrógeno agua-agua, tal como vimos en el capítulo 3 (formando tres HBs). Debido a ello, el agua se ve sustraída de la posibilidad de hidratar al BHB de la proteína, lo cual, de otro modo, resultaría en un efecto disruptivo.

A su vez, estudios previos del Prof. Ariel Fernández [48, 113] han demostrado la existencia de sitios “pegajosos” en proteínas, sitios cuyos defectos de empaquetamiento provocan la exposición de puentes de hidrógeno desprotegidos de la cadena principal o *backbone* (dehidrones) al solvente. Estas interacciones expuestas al agua llamadas dehidrones veremos que promueven la existencia de regiones con agua con propensión a la deshidratación. De esta manera, está emergiendo un escenario de *binding* de proteínas donde las regiones de moléculas de agua fácilmente removibles, regiones con defectos de empaquetamiento en proteínas, serían las principales protagonistas de la unión de pequeños ligandos, como así también de la unión de proteínas.

## 4.2. Cuantificando las interacciones cooperativas: definición del parámetro $\rho$

Para implementar el concepto de cooperativismo invocamos a las correlaciones de tres cuerpos intramoleculares, donde básicamente un tercer cuerpo no polar “protege” a una interacción electrostática, contribuyendo a la deshidratación o exclusión del agua de la misma [2, 48, 113, 114, 79, 18]. En ese sentido seguiremos los relevantes conceptos desarrollados por el Prof. Ariel Fernández en las Ref. [2, 48, 113, 114, 79]. Para implementar dichos conceptos necesitamos definir un parámetro que nos permita evaluar el grado de protección o de “arropamiento” de un puente de hidrógeno de la cadena principal de una proteína (*backbone hydrogen bond*, BHB) mediante las coordenadas cartesianas de la estructura de la proteína. Este parámetro, denotado como  $\rho$ , se define como el número de átomos de carbono ( $\text{CH}_n$ ,  $n = 0, 1, 2, 3$ ) contenidos dentro de un dominio de desolvatación que representa el microambiente del BHB [2]. Estos átomos de carbono (grupos carbonáceos) deben poseer las siguientes características: a) Deben ser originarios de las cadenas laterales de los aminoácidos que conforman la proteína. b) No deben estar enlazados covalentemente a otros átomos más electronegativos que el átomo de carbono mismo (de modo de preservar un carácter hidrofóbico). Finalmente el microambiente del BHB se define como la

reunión de dos esferas intersectadas cuyo radio  $r = 6 \text{ \AA}$  (espesor de aproximadamente tres capas de agua), cada una de ellas centrada en el átomo de carbono  $\alpha$  de los residuos vinculados por un BHB [2]. La Fig. 4.1 esquematiza la metodología descrita recientemente para calcular  $\rho$ .

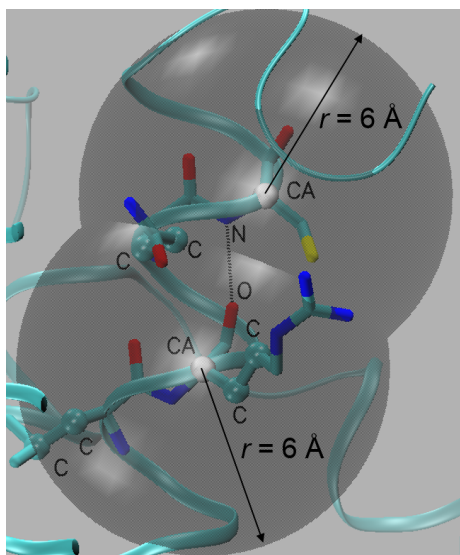


Figura 4.1: Esquematización del cálculo de  $\rho$ . Las esferas blancas representan los carbonos  $\alpha$  (CA) de ambos residuos vinculados por un BHB a través de los átomos N y O respectivamente (se representa el BHB con una línea discontinua entre los átomos N y O). Ambos CA son centro de las dos esferas intersectadas cuyo radio es  $r = 6 \text{ \AA}$ . Las mismas delimitan el microambiente del BHB. Las esferas celestes representan los átomos de carbono C de las cadenas laterales que contribuyen y dan origen, en este caso, a un valor de  $\rho = 6$  (solo átomos de carbono no enlazados covalentemente a otros átomos más electronegativos que el átomo de carbono mismo). Vale decir que solo se representan en la figura unos pocos residuos alrededor del BHB (los valores de  $\rho$  suelen ser mayores), de manera que sea más clara la representación gráfica del concepto que se quiere exponer.

En las estructuras de proteínas solubles reportadas en el *Protein Data Bank* (PDB), el grado de protección promedio ( $\langle \rho \rangle$ ) de los BHBs es  $\rho = 26,6 \pm 7,5$  para esferas de desolvatación de  $6 \text{ \AA}$  [2]. Por lo tanto, definimos como deficiencias estructurales a aquellas que yacen en la cola de la distribución de  $\rho$ , es decir que el microambiente de estos BHBs contiene 19 o menos grupos no polares, por lo que sus respectivos valores de  $\rho$  están por debajo de la media (26,6) menos una desviación estandar (7,5). A este tipo particular de BHBs los llamaremos de ahora en más dehidrones. Mientras que la estadística de  $\rho$  para los BHBs varía con el radio de las



esferas de desolvatación, las colas de la distribución permanecen invariantes, asegurándonos una identificación robusta de las deficiencias estructurales (dehidrones) [113, 114, 79, 26].

### 4.3. El desplazamiento cuadrático medio como medida de la movilidad del agua de hidratación

Con el fin de comprender la influencia de la superficie de una proteína sobre la dinámica del agua de hidratación, mostramos en la Fig.4.2 el desplazamiento cuadrático medio,  $\text{MSD} = \langle r^2(t) \rangle = \langle [\mathbf{r}_i(t) - \mathbf{r}_i(0)]^2 \rangle$ , de las moléculas de agua cercanas a las proteínas que estudiaremos, donde  $\mathbf{r}_i(t)$  es la posición del átomo de oxígeno de la molécula de agua  $i$  al tiempo  $t$  y  $\langle \dots \rangle$  es el promedio sobre todas las  $i$  moléculas de agua. Basándonos en resultados previos, que indican que solamente las moléculas de agua bajo el primer pico de la distribución radial agua-proteína exhiben una dinámica significativamente diferente a aquella correspondiente a las moléculas de agua del *bulk* [24, 20, 17, 121], hemos decidido mostrar solo el comportamiento de las moléculas que se encuentran hasta 4 Å de la superficie de la proteína (ver distribución de distancias mínimas a la lisozima en Fig. 3.1, subsección 3.2.2) en el instante inicial (tiempo cero). Para comparar, también incluimos el comportamiento de las moléculas que en el instante inicial se encuentran alejadas de la superficie de la proteína, con un comportamiento tipo *bulk*.

La inspección directa de la Fig. 4.2 revela que las moléculas de agua cercanas a la proteína presentan una dinámica más “lenta” que aquellas moléculas pertenecientes al *bulk* (tener en cuenta la escala logarítmica de los ejes). Mientras que la curva del MSD para las moléculas alejadas de la superficie de la proteína presenta un comportamiento típico del agua *bulk* con un régimen difusivo [región con pendiente unitaria ( $\langle r^2(t) \rangle \propto t$ ) después del usual régimen balístico o inercial ( $\langle r^2(t) \rangle \propto t^2$ )], las moléculas de agua cercanas a la proteína evidencian un comportamiento subdifusivo, más lento, y solo tienden al comportamiento del *bulk* (difusivo) a tiempos más largos. Este hecho es consistente con resultados previos [24, 20, 121, 132, 93], e inclusive se mantiene si se discrimina entre la componente paralela y perpendicular del MSD a la superficie [121] (en particular, este resultado es consistente con nuestros estudios del agua de hidratación de superficies gráficas [24] y con los estudios de la plastocianina de Bizarri y

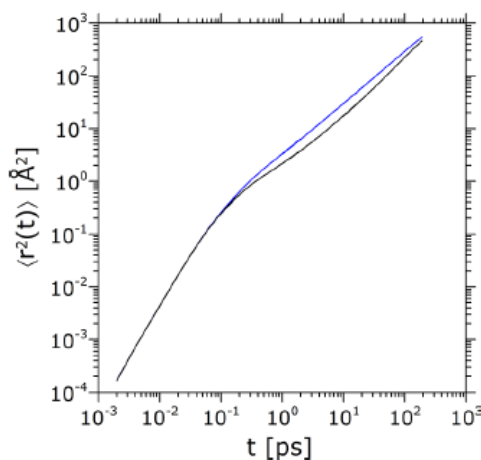


Figura 4.2: El desplazamiento cuadrático medio para las moléculas de agua cercanas a diferentes proteínas (moléculas cuya distancia mínima a la proteína es menor a 4 Å, curva negra) y para moléculas de agua distantes de la proteína (a más de 14 Å, curva azul).

Cannistraro [121, 132]). Este comportamiento de las moléculas cercanas a la proteína habla de la existencia de un régimen superficial a tiempos cortos y distancias cortas y de un régimen tipo *bulk* el cual adoptan las moléculas cuando se alejan de la superficie de la proteína. Por lo tanto, el cálculo de la constante de difusión para una molécula superficial (tal como se obtiene por la relación de Einstein y la cual implica la extrapolación del MSD a tiempos largos) no sería válido. Para resolver este problema, decidimos calcular el desplazamiento cuadrático medio (MSD) a tiempos cortos. Consideramos una escala de tiempo de  $\phi = 4$  ps (una escala de tiempo suficientemente mayor que el tiempo de finalización del régimen balístico). Como se puede apreciar en la Fig. 4.2, a dicha escala de tiempo las moléculas de agua superficiales se mueven en promedio una distancia agua-agua, proporcionando así una medida razonable de las difusividades locales.

Resumiendo, para evaluar la movilidad de las moléculas de agua, dentro del dominio de desolvatación de los BHBs, calculamos los desplazamientos cuadráticos medios de estas últimas moléculas en un intervalo fijo de tiempo  $\phi$  como sigue:  $\langle r^2(t) \rangle = \langle [\mathbf{r}_i(t = \phi) - \mathbf{r}_i(t = 0)]^2 \rangle$ . Como indicamos, el intervalo de tiempo considerado es  $\phi = 4$  ps, el cual es lo suficientemente pequeño para ser sensible a la difusión local de las moléculas de agua abandonando las primeras capas de hidratación, y de esta manera proporcionando una buena medida de la difusividad

local traslacional. Aplicamos este estudio sobre las capas de hidratación de un gran número de proteínas (proteínas completas, sin residuos faltantes y sin ligandos), cuyos códigos en el *Protein Data Bank* son: 1AHO, 1AKI, 1B6D, 1BYI, 1CW6, 1d8v, 1DIV, 1DPT, 1DWU, 1EJG, 1GCN, 1GH5, 1IFB, 1L1I, 1M8L, 1N4I, 1TVM, 1UBI, 1UCS, 1UOY, 1VYC, 1WNJ, 2B4N, 2BZT, 2eyz, 2FDQ, 2GEG, 2jqx, 2JQY, 2JU6, 2K0P, 2K4Q, 2KJG, 2KV4, 2KWD, 2KWL, 2L3V, 2L4V, 2L5R, 2L7W, 2LA1, 2LAO, 2LCU, 2LFN, 2LHC, 2LHS, 2LJM, 2LKB, 2LKY, 2LOL, 2LPK, 2PNE, 2PPP, 2QHE, 2QZW, 2RN2, 2RN4, 2ROG, 3A7L, 3IZP, 3N0K, 4GCR. Estas 62 proteínas fueron elegidas aleatoriamente. Las moléculas de agua fueron modeladas con el modelo de agua explícito TIP3P [131] utilizando el *software* AMBER en las versiones 10 y 11 [83]. Se utilizaron condiciones periódicas de contorno y una caja de simulación de agua que se extiende hasta más de 14 Å de la superficie de la proteína. Las simulaciones de equilibración y de producción del sistema se llevaron a cabo de acuerdo al tutorial oficial de AMBER en la Ref. [11]. El tiempo de simulación en la etapa de producción es de  $t = \phi$  ( $\phi = 4$  ps) y la temperatura seleccionada fue  $T = 300$  K. Con el fin de lograr una buena estadística, para cada una de las 62 proteínas se generaron 400 trayectorias de producción diferentes, cada una de ellas con un tiempo de simulación de  $t = \phi$ .

#### 4.4. La movilidad del agua en los dominios de desolvatación de BHBs

Calculamos la movilidad de las moléculas de agua en los dominios de desolvatación de cada BHB del set de proteínas estudiadas. Como definimos anteriormente, este dominio de desolvatación representa la reunión de dos esferas que se intersectan, cada una de ellas con un radio  $r = 6$  Å y centradas en los carbonos  $\alpha$  de los residuos que participan del BHB [82, 79, 26, 18, 19]. Así, la mayoría de estas moléculas en esta región, se encuentran muy cerca de la cadena principal de la proteína (*backbone*), por lo que se espera que tengan valores de movilidad muy bajos.

La Fig. 4.3 muestra la movilidad de las moléculas de agua en los dominios de desolvatación de los BHBs en función del parámetro  $\rho$  (ver sección 4.2) del BHB, promediada sobre todas las proteínas y configuraciones estudiadas. Claramente se puede observar que existe una correlación

entre el parámetro  $\rho$  y la difusividad local, dado a que el valor de la movilidad promedio decrece de forma monótona con el aumento del valor  $\rho$ , conservando una dependencia bastante lineal. Esto significa que las moléculas de agua cercanas a los BHBs con un mayor nivel de protección o “arropamiento”, poseen valores de movilidad muy bajos (son prácticamente inmóviles), en otras palabras, estas moléculas de agua se encuentran fuertemente unidas a la superficie y así serán más difíciles de remover. Por otro lado, las moléculas de agua cercanas a los dominios de BHBs expuestos o poco “arropados” son claramente más lábiles. Cuando simplemente discriminamos entre dehidrones ( $\rho < 19$ ) y no dehidrones ( $\rho \geq 19$ ), encontramos que la movilidad promedio de las moléculas de agua alrededor de estos últimos es de  $\langle r^2(\phi) \rangle = 3,5 \text{ \AA}^2$ , mientras que para las moléculas de agua en los alrededores de dehidrones es de  $\langle r^2(\phi) \rangle = 5 \text{ \AA}^2$ . Estos resultados son consistentes con la noción de que los dehidrones son motivos estructurales que promueven su desolvatación local. Sucesivamente, como veremos en el siguiente capítulo, los dehidrones han sido determinados como una de las principales componentes de los sitios de unión de proteínas [118, 114, 2, 113, 82, 79, 106, 48, 26, 18, 19], este comportamiento podría ser entonces instrumental en la unión de proteínas o ligandos, los cuales deben ser capaces de remover las moléculas de agua de hidratación para asociarse.

El inset en la Fig. 4.3 prueba que esta tendencia se mantiene cuando utilizamos como radio de las esferas de desolvatación valores de  $r = 5,4 \text{ \AA}$  y  $r = 6,6 \text{ \AA}$ .

Resulta interesante también verificar si las moléculas de agua alrededor de los dehidrones ( $\rho < 19$ ) presentan diferentes propiedades estructurales en comparación con las moléculas de agua alrededor de los BHBs bien protegidos o “arropados” ( $\rho \geq 19$ ). Los dominios de desolvatación de los dehidrones generalmente presentan mayor cantidad de moléculas de agua que aquellos dominios pertenecientes al resto de los BHBs, lo cual podría resultar intuitivo, sin embargo no resulta tan obvio pensar que estas moléculas de agua de hecho presentan un entorno local menos estructurado que el resto de las moléculas de agua. Para evidenciar el incremento en el orden local de las moléculas de agua con el aumento del grado de exclusión del BHB, en la Fig. 4.4 podemos ver el valor medio del índice de estructura local  $I$  (ver subsección 2.2.1), para las moléculas de agua en los dominios de desolvatación de los BHBs, en función del parámetro  $\rho$ . Se puede observar una evidente relación lineal entre el parámetro  $\rho$  de los BHBs y el valor medio

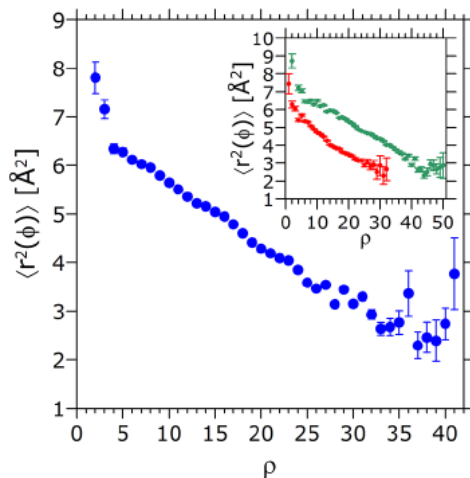


Figura 4.3:  $\langle r^2(\phi) \rangle$  vs.  $\rho$ , para las  $i$  moléculas de agua en los dominios de desolvatación de los BHBs. Las barras de error representan la desviación estandar de los correspondientes valores medios.  $\phi = 4$  ps. El inset muestra las curvas cuando, en vez de usar un radio de  $6,0 \text{ \AA}$  para las esferas de desolvatación, utilizamos los valores  $r = 5,4 \text{ \AA}$  (puntos rojos) o  $r = 6,6 \text{ \AA}$  (puntos verdes).

del índice de estructura local  $I$  para las moléculas de agua en los dominios de dichos BHBs.

El inset de esta última figura muestra la distribución del índice de estructura local  $I$  para las moléculas de agua que se encuentran en los dominios de desolvatación de los dehidrones (curva roja) y no-dehidrones (curva verde). Como era de esperar, la curva para los dehidrones (roja) se encuentra un poco desplazada hacia la izquierda (menores valores de  $I$ ), evidenciando una estructura local más pobre para estas moléculas de agua, con respecto a las moléculas de agua en los entornos del resto de los BHBs. En este caso, la diferencia entre ambas distribuciones es muy pequeña, dado a que hemos discriminado groseramente entre dehidrones y no-dehidrones, sin embargo los resultados siguen siendo consistentes.

Vale la pena mencionar que, en promedio, en los límites del espectro de  $\rho$ , los BHBs con  $\rho = 4$  tienen 21 moléculas de agua en sus dominios de desolvatación, mientras que BHBs con  $\rho = 33$  solo tienen 3. Por lo tanto el ambiente local de no-dehidrones y dehidrones difiere significativamente uno de otro. Sin embargo, la distribución de distancias de las moléculas de agua (en los dominios de desolvatación) al centro geométrico de los BHBs, es prácticamente idéntica para dehidrones y no-dehidrones. Incluso, aunque hay en promedio más moléculas de agua en

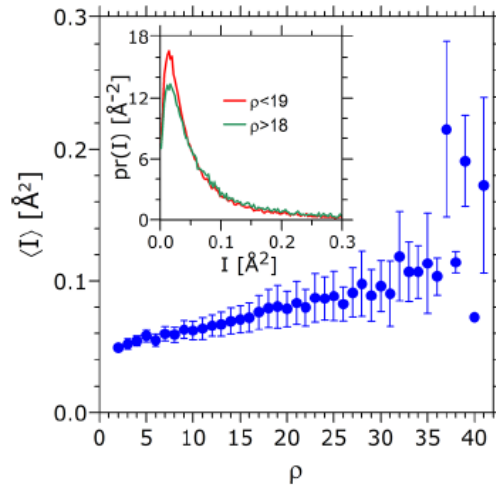


Figura 4.4: Valor medio del índice de estructura local  $I$  en función del grado de exclusión del BHB  $\rho$  (puntos azules). Inset: distribución del índice de estructura local  $I$  (donde  $pr(I)$  indica la probabilidad de encontrar un valor de  $I$  para el índice de estructura local) para las moléculas de agua en los dominios de desolvatación de dehidrones (curva roja) y no-dehidrones (curva verde). No se muestran los datos para aquellas moléculas de agua con solo un vecino hasta los 3, 7Å, las cuales rinden una valor de  $I = 0$ .

los dominios de desolvatación de dehidrones que en los de no-dehidrones, la probabilidad de encontrar una molécula de agua a una cierta distancia del centro geométrico de un BHB es prácticamente la misma para ambos casos. Adicionalmente, vale decir que los dominios de desolvatación de los BHBs con los que estamos tratando no implican una geometría fija. De hecho, algunas porciones de las proteínas (regiones sin estructura o *loops*, las cuales son usualmente ricas en dehidrones) presentan considerables fluctuaciones estructurales. Incluso, ciertos BHBs pueden estar sujetos al efecto disruptivo del agua de hidratación, promoviendo así movimientos en la cadena principal o *backbone*. Por lo tanto, la situación que enfrentamos, más allá de compartir ciertas propiedades con modelos simples de confinamiento hidrofóbico/hidrofílico, es más complicada dadas las particulares características de la estructura y dinámica de las proteínas y las distintivas interacciones proteína-agua.

Con el objetivo de ilustrar mejor el rol de los dehidrones como sitios propensos al “secado” y de esta manera promotores del proceso de *binding*, comentaremos explícitamente la situación de una de las proteínas estudiadas: la proteína p53 (PDB: 2GEQ). Se eligió este sistema en

particular dado a que más allá de la relevancia práctica de esta proteína (la proteína p53 posee una mutación en la mayoría de los casos de cancer en seres humanos, estudiaremos en detalle el caso particular de la mutación Y220C en el Capítulo 6, Sección 6.2), el sitio de *binding* de esta molécula representa uno de los mayores *clusters* de dehidrones en el *Protein Data Bank* [79]. A continuación estudiamos en detalle el sitio de *binding* del p53 al ADN. La naturaleza dehidrónica del sitio de reconocimiento del ADN es evidente: las tres argininas (Arg 245, Arg 270 y Arg 277) involucradas en los contactos p53-ADN forman parte de un *cluster* de 9 dehidrones en la cadena proteica (un *cluster* de dehidrones se define como el máximo set de dehidrones cuyos dominios de desolvatación se intersectan). Dicho *cluster* está formado por los dehidrones (132, 270), (237, 271), (244, 239), (245, 237), (278, 274), (280, 276), (281, 277), (282, 279) y (284, 280), donde indicamos entre paréntesis el par de residuos (de acuerdo a su numeración) involucrados en el dehidrón. La movilidad promedio de las moléculas de agua que se encuentran en la vecindad del *cluster* (esto es, dentro del dominio de desolvatación de todos estos dehidrones) es aproximadamente  $\langle r^2(\phi) \rangle = 5 \text{ \AA}^2$ . Este valor, consistente con el valor de la movilidad promedio alrededor de dehidrones calculado anteriormente para el set de proteínas estudiadas, es significativamente mayor que la movilidad promedio del agua calculada sobre todo el resto de los BHBs de la proteína p53 (no-dehidrones), cuyo valor  $\langle r^2(\phi) \rangle = 3,5 \text{ \AA}^2$ , es cercano al valor medio obtenido para los no-dehidrones del set de proteínas estudiadas. Estos resultados muestran que el sitio de reconocimiento del p53 está de hecho rodeado de moléculas de agua lábiles que serán más fácilmente desplazadas en la asociación con el ADN. Dado a el gran tamaño de este *cluster* (9 dehidrones), esta molécula presenta un gran parche de agua de hidratación fácilmente removible en su región de contacto, proporcionando así un recurso para la aproximación y el *binding* del ADN. Más aún, en este caso particular la tendencia hacia los dehidrones en el sitio de *binding* es incluso más relevante si consideramos también la naturaleza electrostática del proceso de *binding*. Esto se debe a que el proceso de reconocimiento basado únicamente en las interacciones electrostáticas no podría ser operativo en un entorno de agua *bulk*, donde la elevada constante dieléctrica del medio debería apantallar las cargas eléctricas. Sin embargo, al promover la desolvatación local del sitio de *binding* del p53 al ADN, este *cluster* de dehidrones estaría jugando un papel muy importante en el *quenching* del medio dieléctrico local al sitio de *binding* y así desencadenando

la interacción electrostática entre las argininas cargadas positivamente y los fosfatos del ADN cargados negativamente.

## 4.5. Conclusiones

En este Capítulo presentamos firme evidencia de la existencia de un *link* o conexión entre el grado de exclusión o “arropamiento” de los BHBs (el parámetro estructural  $\rho$ ) y la movilidad de las moléculas de agua de hidratación (ver Fig. 4.3). Primeramente verificamos que en promedio las moléculas de agua que se encuentran en la superficie de las proteínas resultan más lentas que las moléculas del *bulk*, lo cual es consistente con nuestros estudios con grafeno [24] y con estudios anteriores sobre la proteína plastocianina [121, 132]. Sin embargo, el comportamiento del agua en regiones expuestas (en los defectos de empaquetamiento o dehidrones) es claramente diferente. Mostramos que la exposición del *backbone* está de hecho correlacionada con la pérdida de hidratación de la superficie de la proteína y así, estas interacciones intramoleculares desprotegidas o expuestas al agua de hidratación (dehidrones) se encuentran rodeadas de moléculas de agua debilmente unidas y facilmente removibles. De esta manera, promoviendo su deshidratación local, estos motivos estructurales llamados dehidrones proporcionan un recurso por el cual un ligando podría desplazar a las moléculas de agua durante el proceso de *binding*. Este trabajo se realizó en colaboración con el Prof. Ariel Fernández.



## Capítulo 5

# La exclusión del agua de hidratación y las interacciones cooperativas en la asociación de proteínas

En este capítulo delineamos las bases moleculares de las interacciones cooperativas, un concepto biomolecular crucial para el estudio de la asociación de proteínas como también para el diseño racional de drogas. En este contexto, hemos determinado que las interacciones cooperativas o de tres cuerpos, basadas en la exclusión de agua, son esenciales para mantener la integridad de la interfase proteína-proteína al prevenir la hidratación interfacial disruptiva por parte de entidades polares. Como mencionamos anteriormente, este atributo molecular puede ser explotado en el diseño racional de drogas, dado a que dichas regiones pueden ser utilizadas como arquetipos para diseñar pequeñas moléculas disruptivas de interfases proteína-proteína.

### 5.1. La necesidad de un concepto de cooperatividad biomolecular

La asociación de proteínas es un proceso molecular básico en biología [48, 143, 128, 47, 95, 57, 122, 123, 65, 153, 78, 77, 58]. A pesar de su importancia, sus fundamentos biofísicos continúan

siendo un tema de intenso debate [48, 143, 128, 47, 95, 57, 122, 123, 65, 153, 78, 77, 58]. Diversos estudios cristalográficos han mostrado que las interfases de asociaciones entre proteínas pueden ir de los 600 Å<sup>2</sup> a los 1300 Å<sup>2</sup> e incluyen varios contactos intermoleculares, pudiendo involucrar de 10 a 30 cadenas laterales de cada una de las unidades proteicas [153]. Sin embargo los experimentos nos han demostrado que solo unos pocos residuos de la interfase (llamados *hot spots*) son los responsables de la unión [153]. Uno de los problemas actuales más desafiantes involucra la caracterización de estos sitios llamados *hot spots* [143, 128, 47, 95, 57, 122, 123, 65, 153, 78, 77, 58]. Como mencionamos anteriormente, estos sitios involucran una pequeña fracción de residuos de la región interfacial y contribuyen significativamente a la estabilidad de la interfase proteína-proteína. Estudios de primeros principios, a través de potenciales de *docking*, han sido relativamente satisfactorios en la predicción de estos sitios singulares [143, 128, 47, 95, 57, 122, 123, 65, 153, 78, 77, 58], correlacionando con resultados experimentales como la mutación de sitios específicos o *alanine scanning* [153]. Esta última técnica experimental evalúa el impacto, en la energía libre de la asociación, de truncar la cadena lateral de un residuo individual en el carbono  $\beta$  (se realiza una mutación en la que dicho residuo se reemplaza por alanina) [18]. Más allá del logro en la predicción de estos sitios, la naturaleza física de los *hot spots* sigue siendo difícil de racionalizar.

Por otro lado, la estructura (secundaria y terciaria) de las proteínas en solución es sostenida por fuerzas esencialmente electrostáticas [2], es decir, para prevalecer en ambientes acuosos, las proteínas solubles deben proteger sus propios puentes de hidrógeno del *backbone* (BHBs) (puentes de hidrógeno formados entre los grupos -NH- y -C=O de la cadena principal) del efecto disruptivo del ataque del agua. Vale decir que en condiciones fisiológicas, el proceso de plegamiento de una proteína globular está claramente favorecido termodinámicamente. Sin embargo la entropía conformacional del plegado realiza una contribución positiva al  $\Delta G$  del proceso. Es decir, el cambio de entropía conformacional se opone al plegado. Por lo tanto un  $\Delta G < 0$  implica que, o bien el valor de  $\Delta H$  es negativo y de elevado valor absoluto, o que debe existir también algún tipo de incremento entrópico con el plegado. De hecho ambas cosas ocurren, y en particular la fuente de un cambio entálpico negativo es el cúmulo de interacciones energéticamente favorables que se dan en el interior del glóbulo proteico, interacciones que suelen ser no

covalentes. Más aun, la exclusión del agua circundante fortalece a algunas de estas interacciones (las de naturaleza electrostática), disminuyendo la constante dieléctrica local a la interacción [2]. Por lo tanto los BHBs formados en el interior del glóbulo proteico se verán estabilizados por la agrupación de residuos no polares alrededor de ellos. Es decir que las interacciones cooperativas o de tres cuerpos, donde un tercer cuerpo no polar “protege” a una interacción electrostática, son inherentes al plegado de la cadena proteica [2]. Por otro lado, aquellos BHBs no excluidos son “pegajosos” y por consiguiente promotores de la asociación de proteínas, ya que su estabilidad inherente se verá incrementada por la aproximación de residuos no polares. De este modo, la carencia en la protección intramolecular se ve removida de modo intermolecular. Así, la integridad de la interfase proteína-proteína en complejos será dependiente de la cooperatividad intermolecular [2, 48, 113, 114, 79, 18]. Este enfoque, ampliamente desarrollado por el Prof. Ariel Fernández [2, 48, 113, 114, 79], constituye el punto de partida de nuestros estudios.

## 5.2. Identificación y caracterización de *hot spots* en interfases proteína-proteína

En esta sección estudiamos diferentes complejos entre proteínas y calculamos los valores de  $\rho$  que surgen de correlaciones intra e intermoleculares [18]. Específicamente, determinamos los valores de  $\rho$  para los BHBs intramoleculares y para los (menos frecuentes) BHBs intermoleculares. Para cada uno de estos BHBs consideramos el dominio (microambiente del BHB) definido por la intersección de dos esferas de radio 6 Å centradas en los carbonos  $\alpha$  de los correspondientes residuos vinculados por un BHB (ver Fig. 4.1). Luego calculamos el grado de protección  $\rho$  como el número de grupos no polares  $\text{CH}_n$ ,  $n = 0, 1, 2, 3$  (átomos de carbono que no se encuentran enlazados a átomos electrofílicos o grupos polarizables) que residen dentro de dicho dominio de desolvatación y que además provienen de las cadenas laterales de cualquiera de los residuos de ambas proteínas involucradas en el complejo. Por lo tanto, un BHB puede recibir protección, o ser excluido del agua, por grupos no polares de las cadenas laterales de diferentes residuos, ya sea de una o ambas proteínas que conforman el complejo, contribuyendo así cada uno de ellos al valor total de  $\rho$  para dicho BHB [18]. La Fig. 5.1 a) muestra un ejemplo de protección inter-

molecular de un BHB intramolecular, mientras que la Fig. 5.1 b) muestra un caso de protección de un BHB intermolecular. Vale decir que cada uno de estos ejemplos ilustra la protección de uno o varios BHBs por solo un residuo. Usualmente, cada BHB se encuentra simultáneamente excluido por diferentes residuos, sin embargo en esta ocasión nuestra intención es ilustrar de forma clara el concepto de cooperatividad, por lo que suprimimos el resto de las interacciones que pueden existir en cada uno de estos ejemplos presentados en las Fig. 5.1 a) y b).

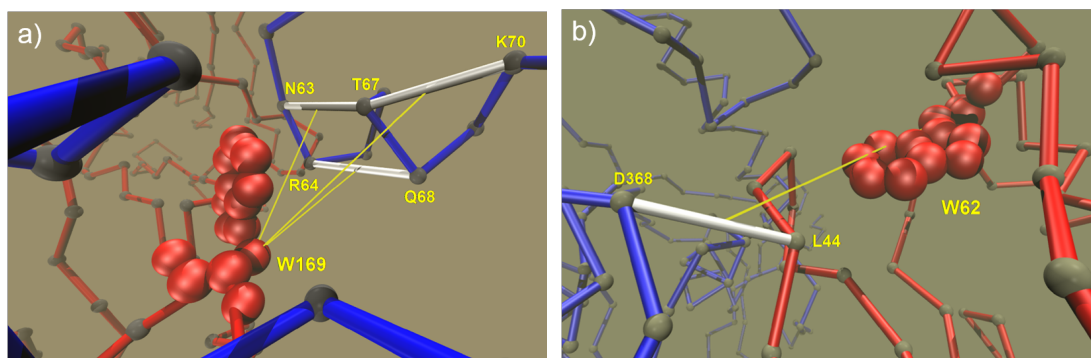


Figura 5.1: Ilustración de la protección o cooperatividad intermolecular: a) residuo W169 (esferas rojas) del receptor de la hormona de crecimiento humano hGHbp (cadena roja) protegiendo tres BHBs intramoleculares de la hormona de crecimiento humano hGH (cadena azul). Con el fin de simplificar la representación, los BHBs se representan como varillas blancas entre los correspondientes carbonos  $\alpha$ ; b) Similar a a) pero para el caso del complejo entre la glicoproteína del virus del HIV gp120 y el receptor CD4. En este caso el residuo W62 de la cadena del receptor CD4 protege un BHB intermolecular entre ambas proteínas.

### 5.2.1. *Alanine scanning* computacional basado en el concepto de cooperatividad

Con el fin de poder caracterizar y cuantificar los distintos *hot spots* de un set de complejos proteína-proteína, desarrollamos un algoritmo computacional para llevar a cabo un *alanine scanning* virtual [18], donde básicamente se reemplaza cada residuo de una cadena proteica (uno a la vez) por alanina dentro de la estructura 3D del complejo. Luego, para caracterizar estos sitios, se evalúa el impacto de la sustitución o mutación en la cooperatividad intermolecular (no se realiza una simulación de dinámica molecular, sino que trabajamos directamente con las

coordenadas del PDB reportado para cada complejo). Para la mayoría de los residuos (aquellos con cadenas laterales con más átomos que la alanina), esta mutación implica truncar la cadena lateral del residuo en el carbono  $\beta$ , por lo que la misma es reemplazada por un grupo  $-\text{CH}_3$ , de manera que se reduce significativamente el volumen de la cadena lateral de dicho residuo y por ende su capacidad de excluir el agua circundante [18]. En el caso especial de la glicina (la cual carece de carbono  $\beta$ ), incluimos un grupo metilo en la correspondiente posición (reemplazando un átomo de H del carbono  $\alpha$ ), incrementando la capacidad de protección de dicho residuo.

Los pasos a seguir por nuestro algoritmo se enumeran a continuación: En un primer estadio, calculamos el valor de  $\rho_i$  para cada uno de los BHBs a partir de la estructura 3D del complejo, obteniendo de esta manera un set de valores de  $\rho_i$  para el complejo en su estado natural o *wild type*. Luego para cada residuo que es sustituido por alanina (uno a la vez y sin modificar la estructura del resto del complejo), calculamos un nuevo set completo de valores de  $\rho$ , a los que llamaremos  $\rho_f$ . Por lo tanto si el complejo posee  $n$  aminoácidos y  $j$  BHBs, se llevarán a cabo  $n$  mutaciones individuales, y se obtendrán  $n \times j$  valores de  $\rho_f$ . La Fig. 5.2 esquematiza el proceso de *alanine scanning in silico* desarrollado por nuestro grupo.

Finalmente, de acuerdo a nuestra premisa de cooperatividad intermolecular, los *hot spots* pronosticados por nuestro *alanine scanning* computacional serán caracterizados, en bases a su rol como protectores intermoleculares, de acuerdo a las siguientes clases:

a) La sustitución por alanina de un residuo de una de las cadenas del complejo disminuye el valor de  $\rho$  de un BHB [un BHB intramolecular de la otra cadena del complejo (*partner*) o un BHB intermolecular] y donde  $\rho_f < \langle \rho \rangle$ . Este tipo de *hot spots* serán considerados de clase **A**. En el caso que  $\rho_f < 19$  (creación de un dehidrón) este tipo particular de *hot spots* clase **A** serán rotulados con un asterisco como superíndice (**A\***). *La mutación de estos aminoácidos disminuye significativamente la protección de los BHBs cuyos dominios de desolvatación alojan las cadenas laterales de este tipo de residuos. Incluso en algunos casos se generan nuevos dehidrones, desestabilizando la estructura de los complejos.*

b) La sustitución por alanina de un residuo como glicina, serina, cisteína, ácido aspártico o asparagina, contenido dentro del dominio de desolvatación de un BHB de su propia cadena proteica, que recibe protección por parte del *partner* y que además su valor, solo intramolecular,

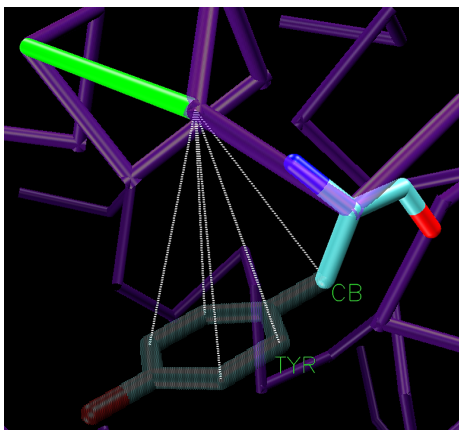


Figura 5.2: Esquematación de nuestro *alanine scanning* computacional. Primero calculamos el valor de  $\rho_i$  para cada uno de los  $j$  BHBs en la estructura 3D del complejo, generando un set de  $j$  valores de  $\rho_i$  para la estructura *wild type*. Por cada uno de los  $n$  residuos que es reemplazado por alanina (uno a la vez y sin modificar la estructura del resto del complejo), calculamos el valor de  $\rho_f$  para cada uno de los  $j$  BHBs, generando un set de  $j$  valores de  $\rho_f$  para cada una de las  $n$  estructuras mutadas (habrá tantas estructuras mutadas como residuos por mutar). Finalmente para cada uno de los  $j$  BHBs se obtiene un valor de  $\Delta\rho = \rho_i - \rho_f$  luego de cada una de las  $n$  mutaciones. Por lo tanto se obtienen  $n \times j$  valores de  $\Delta\rho$ . En la imagen se representa un BHB con una varilla de color verde conectando los carbonos  $\alpha$  de ambos residuos vinculados por el HB. Las líneas punteadas denotan que los átomos de la cadena lateral de la tirosina, representada en *licorice*, se encuentran en el dominio de desolvatación de dicho BHB, protegiendo al mismo y contribuyendo a su respectivo valor de  $\rho_i$ . Cuando la tirosina es reemplazada por alanina, la cadena lateral es truncada en el carbono  $\beta$  por lo que ahora se computa un valor de  $\rho_f$  para dicho BHB. Finalmente, esta mutación resulta en un  $\Delta\rho = \rho_i - \rho_f = 5$  para el BHB representado en esta imagen.

de  $\rho \leq 19$ . Esta sustitución incrementa intramolecularmente el valor de  $\rho$  para dicho BHB en  $\Delta\rho = 1$ . La predicción resultante de este tipo de *hot spots* se la considera de clase **B**. En el caso particular de que, intramolecularmente, el valor de  $\rho = 19$ , este tipo de *hot spots* clase **B** serán rotulados con un asterisco como superíndice (**B\***). *Estas mutaciones incrementan intramolecularmente el valor de  $\rho$  de un dado dehidrón pre-existente en la misma cadena. De esta manera la necesidad de cubrir esta deficiencia por parte del partner ahora es menor. En el caso de que la mutación implica elevar en una unidad un valor pre-existente de  $\rho = 19$  (remoción de un dehidrón), esta situación disminuye aún más la naturaleza adhesiva de dicho sitio.*

Decidimos dejar afuera del algoritmo los HBs entre cadenas laterales. Este tipo de interacciones en la superficie poseen un costo entrópico asociado a la formación de la misma, lo cual las hace marginalmente estables [48]. Además la distribución de  $\rho$  para los HBs de cadenas laterales no presenta una clara distinción de las colas de la distribución, ya que esta última es esencialmente “chata”. Esto es consecuencia de la riqueza conformacional que poseen las cadenas laterales. Por otro lado, la mutación de un residuo que posee una interacción por puente salino con el *partner* (puente salino intermolecular) será desestabilizante, ya que se perderá una interacción iónica, la cual se encontraba excluida del agua en la interfase del complejo, de manera que solo una de las cargas quedará “enterrada”, lo cual es una situación desfavorable [144]. Esta clase trivial de *hot spots* forman parte de aproximadamente el 15% de todos los *hot spots* comprendidos en los complejos estudiados, y obviamente están fuera del alcance de nuestro análisis basado en interacciones cooperativas [18].

Estudiamos las siguientes interfases proteína-proteína de complejos reportados en el PDB, para los cuales previamente ya existen resultados experimentales de *alanine scanning* (en cada caso, la primera proteína del complejo es sobre la cual se realizaron las mutaciones; para cada complejo indicamos el código en el PDB y la referencia al análisis experimental de *alanine scanning*): *Human growth hormone receptor/Human growth hormone* [153] (**3HHR**), *Trypsin inhibitor/Beta-Trypsin* [151] (**2PTC**), *p53/MDM2* [146] (**1YCR**), *CD4/gp120* [160] (**1GC1**), *Ribonuclease inhibitor/Ribonuclease A* [147] (**1DFJ**), *Colicin E9 immunity protein/Colicin E9 DNase domain* [136] (**1BXI**), *Barnase/Barstar* [154] (**1BRS**), *Barstar/Barnase* [154] (**1BRS**), *Ribonuclease inhibitor/Angiogenin* [147] (**1A4Y**).

La Fig. 5.3 muestra nuestras predicciones. La sustitución experimental por alanina sobre una proteína nativa involucra un cambio en la energía libre de la asociación del complejo ( $\Delta G$ ), dicho cambio se calcula como  $\Delta\Delta G = \Delta G_{\text{mut}} - \Delta G_{\text{wt}}$ , (mut = mutada, wt = nativa) a cuyo valor se le asocia en la Fig. 5.3 un determinado color de acuerdo a su valor experimental. Nuestras predicciones, basadas en el concepto de cooperatividad, se indican por debajo de los correspondientes residuos, caracterizando dichos *hot spots* de acuerdo a su rol como protectores intermoleculares con el código de letra definido previamente (**A**, **A\***, **B**, **B\***). La letra “S” indentifica los *hot spots* involucrados en puentes salinos intermoleculares. Como mencionamos

anteriormente, estos *hot spots* son triviales y por lo tanto no son considerados en la comparación con nuestras predicciones.

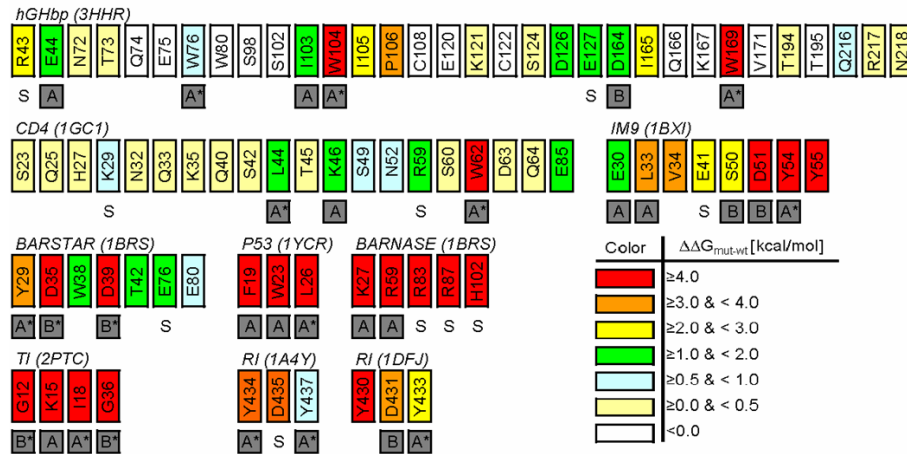


Figura 5.3: Resultados experimentales para el cambio en la energía libre de la asociación del complejo ( $\Delta\Delta G$ ) luego de reemplazar individualmente cada uno de los aminoácidos listados por alanina. Esta información se contrasta con los resultados arrojados por nuestro *alanine scanning* computacional basado en el concepto de cooperatividad. Para cada uno de los casos se muestra la porción de residuos para los cuales existen datos experimentales. La escala de colores que se muestra a la derecha indica el valor de  $\Delta\Delta G$  determinado experimentalmente para cada uno de estos aminoácidos. Los cuadrados grises indican nuestras predicciones computacionales y el código de letra indica de que clase de *hot spot* se trata de acuerdo a nuestro concepto de cooperatividad. La letra “S” indica cual de estos *hot spots* participa de un puente salino intermolecular.

En la la tabla 5.1 mostramos los *hot spots* clase **A** y **A\*** que se obtienen luego de realizar nuestro alanine scanning computacional sobre el receptor de la hormona de crecimiento humano (hGHbp) en el complejo (hGHbp-hGH).

Para cuantificar la capacidad predictiva de nuestro método, en la tabla 5.2 mostramos nuestras predicciones globales sobre el set completo de complejos estudiados. Esta comparación entre nuestro método teórico y los experimentos, revela que nuestro procedimiento computacional localiza la mayoría de los *hot spots* encontrados experimentalmente por *alanine scanning*. En el caso particular de aquellos *hot spots* con  $\Delta\Delta G \geq 4$  kcal/mol el porcentaje de acierto de nuestro algoritmo es del 89 %. La mayor contribución a este porcentaje es por parte de las mutaciones



Residuo mutado en hGHbp	BHB en hGH	$\rho_i$	$\rho_f$
Glu 44	Gln 68 - Arg 64	20	19
Trp 76	Gln 46 - Tyr 42	21	18
Trp 76	Asn 47 - Ser 43	10	9
Ile 103	Thr 67 - Asn 63	27	25
Trp 104	Thr 67 - Asn 63	27	26
Trp 104	Thr 175 - Asp 171	19	17
Trp 104	Ile 179 - Thr 175	21	19
Trp 169	Thr 67 - Asn 63	27	23
Trp 169	Gln 68 - Arg 64	20	14
Trp 169	Lys 70 - Thr 67	20	18

Cuadro 5.1: Predicciones detalladas de *hot spots* de clase **A** y **A\*** para el complejo entre la hormona de crecimiento humano y su receptor. La primera columna indica los residuos mutados por alanina en el receptor (hGHbp), la segunda columna indica los BHBs en la hormona de crecimiento humano (hGH) cuyo valor de  $\rho$  es afectado por la mutación.  $\rho_i$  y  $\rho_f$  son los valores de  $\rho$  del BHB correspondiente en el complejo proteína-proteína antes y después de la mutación, respectivamente.

de clase **A** (**A** y **A\***) (61 %), mientras que el restante corresponde a las mutaciones de clase **B** (**B** y **B\***) (28 %). La última columna de la tabla indica los porcentaje de acierto en nuestras predicciones considerando solo la creación de un dehidrón, **A\***, y la remoción de un dehidrón, **B\***. En concordancia con nuestra premisa, se espera que este tipo de *hot spots* sean de gran importancia, y de hecho lo son, ya que estas mutaciones representan el 56 % de las mutaciones con  $\Delta\Delta G \geq 4$  kcal/mol determinadas experimentalmente.

<i>hot spots</i> experimentales (valor de $\Delta\Delta G$ )	Porcentajes de acierto (%)		
	A+A*+B+B*	A+A*	A*+B*
$\geq 4$ kcal/mol	89	61	56
$\geq 3$ kcal/mol	83	58	50
$\geq 2$ kcal/mol	79	54	46
$\geq 1$ kcal/mol	74	53	37

Cuadro 5.2: Predicciones obtenidas para los diferentes complejos de proteínas estudiados. La segunda columna muestra el desempeño global de nuestro método.

A esta altura vale decir que nuestro método ignora las interacciones convencionales de dos cuerpos a menos que estén involucradas en una correlación de tres cuerpos. Este nuevo enfoque emerge como consecuencia de que ninguna interfase proteína-proteína ha demostrado ser trivial

al nivel de interacciones de a pares, para desarrollar un procedimiento satisfactorio para la predicción de *hot spots* [143, 128, 47, 95, 57, 122, 123, 65, 153, 78, 77, 58]. Este último punto también dificulta establecer un control para nuestros resultados, no obstante para ello definimos un análisis elemental basado en la complementaridad entre entidades polares e hidrofóbicas. Con este fin, hemos caracterizado a los aminoácidos en hidrofóbicos (cadenas laterales aromáticas o alifáticas) o polares (cadenas laterales polares o con carga) y construimos una matriz de contacto para la interfase del complejo. Para cada residuo calculamos la distancia mínima entre su carbono  $\alpha$  y los carbonos  $\alpha$  de los aminoácidos del *partner* como también calculamos la distancia mínima entre el centroide de la cadena lateral del aminoácido en cuestión y los centroides de las cadenas laterales de los residuos del *partner*. Cuando alguna de estas distancias es menor a 8 Å, consideramos a ese par de residuos involucrados en la medida, como interfaciales. Así, habiendo identificado la interfase de ambas cadenas proteicas, un contacto entre residuos de ambas cadenas ocurre cuando la distancia entre carbonos  $\alpha$  o la distancia entre centroides de las cadenas laterales es menor a 6 Å. Los resultados son robustos a moderados cambios en el parámetro de contacto y se ajustan a un criterio adoptado previamente para interfases proteína-proteína [153].

La Fig. 5.4 muestra la matriz de contacto para la interfase del complejo hGH/hGHbp, donde los puntos rojos indican un contacto hidrofóbico/polar (*HP*), mientras que los puntos azules indican un contacto polar/polar o hidrofóbico/hidrofóbico (*PP* o *HH*). En esta figura podemos apreciar que la interfase presenta un número significativo de contactos *HP* (alrededor de un 37 %), por lo tanto es evidente que la asociación de proteínas no puede ser simplemente racionalizada con la búsqueda de interacciones de a pares del tipo *PP* y *HH*. Es interesante recalcar, que cuando acotamos el análisis solo a los *hot spots* determinados experimentalmente, el porcentaje de contactos *HP* asciende ligeramente (42 %). Más aún, si solo analizamos los dos *hot spots* más importantes de esta interfase (TRP 104 y TRP 169, los únicos residuos con  $\Delta\Delta G \geq 4$  kcal/mol), estos aminoácidos están involucrados en 8 contactos *HP* y solo 1 contacto *HH*. Esto es razonable si pensamos que gran parte del contenido polar en la superficie de la proteína termina siendo “enterrado” una vez formado el complejo. Sin embargo, cuando nos focalizamos en las interacciones de tres cuerpos, descubrimos que gran parte de los residuos hidrofóbicos en

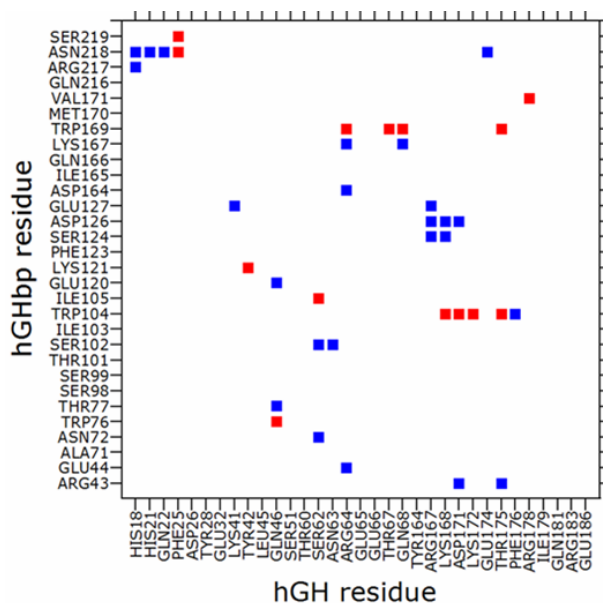


Figura 5.4: Matriz de contacto para la interfase del complejo hGH/hGHbp. La distancia máxima aceptada para un contacto intermolecular es 6 Å (ver texto). Los puntos rojos indican un contacto hidrofóbico-polar (*HP*) entre los correspondientes residuos, mientras que los puntos azules implican un contacto polar-polar o hidrofóbico-hidrofóbico (*PP* o *HH*).

la interfase del complejo se aproximan a residuos polares con el fin de excluir el agua de los BHBs en los cuales estos últimos residuos están involucrados, y así favorecer la formación del complejo [18].

### 5.3. Conclusiones

En este Capítulo mostramos que las interfases proteína-proteína eluden el análisis fisicoquímico estándar. La racionalización en términos de la aditividad de interacciones complementarias de a pares en la región de contacto es insatisfactoria, especialmente en lo que respecta al rol de los *hot spots* como determinantes en la asociación de proteínas. En contraposición a esta situación, este simple motivo molecular presentado en este capítulo, y edificado sobre los conceptos desarrollados por el Prof. Ariel Fernández, demuestra ser ubicuo en la determinación de asociaciones entre proteínas. Dicho motivo es un indicador de la naturaleza cooperativa de las interacciones de tres cuerpos. En esencia, este concepto surge de la estabilización proporcionada

por un grupo hidrofóbico, de una cadena proteica, sobre un BHB de la proteína *partner* o un BHB intermolecular. De esta manera la interacción se da entre tres cuerpos.

Como veremos en el siguiente capítulo, estos resultados son una herramienta a tener en cuenta para el diseño racional de pequeñas moléculas disruptivas de interfases proteína-proteína, las cuales, en base a un diseño racional, pueden imitar la capacidad “protectora” de los *hot spots*. Este nuevo concepto de reconocimiento biomolecular nos permite caracterizar y racionalizar las interacciones en las interfases proteína-proteína, un problema que desde hace mucho tiempo no ha podido resolverse dado que solo se han considerado enfoques basados en la aditividad de las interacciones de a pares [18].

Este trabajo se realizó en colaboración con el Prof. Ariel Fernández.

## Capítulo 6

# Interacciones droga-proteína: hacia el diseño racional y la reingeniería de drogas

En este capítulo estudiamos algunos casos exitosos de complejos ya reportados en la literatura, entre pequeñas moléculas disruptivas de interfases proteína-proteína y sus respectivas proteínas. Verificamos la importancia de las interacciones de tres cuerpos (descriptas en el capítulo anterior) en la interacción entre drogas y proteínas al verificar que este tipo de drogas tienden a mimetizar las interacciones intermoleculares cooperativas de la proteína que reemplazan. Esta comprensión es de gran relevancia para el diseño racional de drogas destinadas a bloquear interfases de proteínas relevantes para inhibir ciertas enfermedades. También es importante para rediseñar racionalmente compuestos guía (*lead compounds*) y como filtro de selectividad para evitar efectos colaterales indeseables. Sobre la mitad de este capítulo estudiamos un caso particular de mutación oncogénica, la mutación Y220C de la proteína supresora de tumores p53. En base a nuestro enfoque de cooperativismo, racionalizamos la naturaleza de la mutación oncogénica Y220C, explicamos el *binding* de una molécula pequeña ya diseñada para reestablecer la función del p53 y proporcionamos pistas que ayuden a mejorar la función de rescate de dicha molécula.

## 6.1. Drogas disruptivas de interfases proteína-proteína y la mimetización de interacciones intermoleculares cooperativas

Como vimos anteriormente en el Capítulo 5, la comprensión de las interacciones entre proteínas es de enorme importancia, ya que a nivel molecular la mayoría de los procesos biológicos implican la asociación de proteínas [18]. Estos estudios nos brindan algunas de las herramientas necesarias para poder desarrollar pequeñas moléculas, con fines terapéuticos, capaces de unirse a una determinada proteína en estudio. Sin embargo, el descubrimiento y diseño de pequeñas moléculas disruptivas de interfases proteína-proteína (moléculas que se unen a la proteína objetivo, reemplazando así a la otra cadena proteica natural), con una elevada especificidad y afinidad, constituye un enorme desafío [76]. Diversos estudios muestran que pequeñas moléculas disruptivas (ligandos) son capaces de unirse potencialmente a su correspondiente proteína objetivo (receptor o *target*), estableciendo interacciones con los mismos residuos (*hot spots*) del receptor con los que interaccionaba también la proteína desplazada [76]. En estos casos, estas moléculas muestran una similitud tanto espacial como química (hidrofobicidad) con los residuos de la proteína desplazada, los cuales se encontraban involucrados en la interfase proteína-proteína [76]. El análisis de complejos entre exitosas moléculas disruptivas y proteínas, sugiere la posibilidad de que estos ligandos de alguna manera mimeticen el comportamiento de la cadena proteica que reemplazan (*binding partner*) [76]. Sin embargo, la idea de una mimetización, por parte de la pequeña molécula, no puede terminar de ser precisa a menos que se comprenda la naturaleza de las interacciones involucradas en la unión, lo cual también conlleva a la racionalización previa de las interacciones que ocurren en complejos proteína-proteína. En ese sentido, el mérito de nuestro enfoque basado en la cooperatividad intermolecular, presentado en el Capítulo 5, fue el de elucidar la naturaleza física de los llamados *hot spots* [18]. Vale aclarar que también otro tipo de interacciones (como puentes salinos entre otras) juegan un papel importante también en el *binding*. Sin embargo el concepto de cooperatividad nos permite racionalizar la naturaleza de los *hot spots*, algo hasta el momento inaccesible aún cuando se pudieran predecir satisfactoriamente. Este conocimiento, como ya habíamos anticipado, puede ser de gran ayuda en la ingeniería de proteínas y el diseño de drogas. Por lo tanto en este contexto aplicamos este conocimiento al

estudio de las correlaciones de tres cuerpos en complejos entre pequeñas moléculas disruptivas y su correspondiente receptor, y compararemos estos resultados con aquellos correspondientes al complejo proteína-proteína natural. Así demostraremos que, al nivel de las interacciones cooperativas, las pequeñas moléculas disruptivas mimetizan sustancialmente las interacciones efectuadas con la proteína desplazada [19].

### 6.1.1. El concepto de cooperativismo intermolecular en complejos proteína-droga

Para cada uno de los sistemas entre drogas disruptivas y proteínas que analizamos, en primer lugar debemos considerar el complejo natural proteína-proteína. Por lo tanto de acuerdo a lo expuesto en el Capítulo 5, sección 5.2, las interacciones intermoleculares cooperativas relevantes, en el complejo proteína-proteína, se darán en los siguientes casos: Supongamos que un BHB intramolecular en la proteína A posee un determinado valor de  $\rho_A$  cuando la proteína no se encuentra formando complejo con su *binding partner*, al cual llamaremos proteína B (a este valor de  $\rho$  solo contribuye la proteína A; ver capítulo 5 sección 5.2). Luego, al formarse el complejo AB, dicho BHB posee al menos una interacción cooperativa intermolecular. Por lo tanto en el complejo, dicho BHB posee un determinado valor de  $\rho_{AB}$  (a este valor de  $\rho$  contribuye tanto la proteína A como la B). Así, las interacciones intermoleculares cooperativas relevantes serán solo aquellas en las que inicialmente  $\rho_A$  sea menor al valor de  $\langle\rho\rangle$  [19] (ver capítulo 4 sección 4.2). De este modo  $\Delta\rho_{AB} = \rho_{AB} - \rho_A$  será la contribución adicional al valor de  $\rho$  de dicho BHB por parte del *binding partner* o proteína B. La situación será aún más significativa si un BHB, que posee al menos una interacción cooperativa intermolecular, tiene un valor de  $\rho_A \leq 19$ . En esta última situación, al considerar solo el aporte al valor de  $\rho$  de la proteína A, dicho BHB es considerado un dehidrón ya que  $\rho_A \leq \langle\rho\rangle - \sigma$  (ver capítulo 4, sección 4.2). Por otro lado, la proteína *target* (proteína A), presentará también BHBs en la interfase del complejo cuyo valor de  $\rho_A \geq \langle\rho\rangle$ . En este último caso, si estos BHBs poseen interacciones intermoleculares cooperativas con la proteína B, estas contribuciones serán consideradas irrelevantes, ya que estos BHBs están protegidos intramolecularmente del ataque del agua. Finalmente, a fin de comparar con el complejo entre la pequeña molécula disruptiva o droga y su correspondiente

proteína *target*, calculamos, para el set de BHBs interfaciales, el valor de  $\rho_{AD}$ . Este último valor incluye la contribución al valor de  $\rho$  de dicho BHB tanto por parte de la proteína A como de la tentativa droga D. Así, a fin de comparar el grado de cooperativismo intermolecular entre el complejo proteína-proteína y el complejo proteína-droga, calculamos  $\Delta\rho_{AD} = \rho_{AD} - \rho_A$  y comparamos este valor con  $\Delta\rho_{AB}$ . Vale decir que el valor  $\rho_A$  siempre se calcula a partir de la proteína A originaria del complejo AB, es decir que se calcula a partir del PDB originario del complejo proteína-proteína. También, para cada BHB, calculamos  $\rho_D$ , cuyo valor corresponde solo a la contribución, por parte de la pequeña molécula, al valor de  $\rho$  total de dicho puente de hidrógeno en el complejo AD.

El cálculo de  $\rho$  se lleva a cabo tal cual ya fue descrito en el Capítulo 4, sección 4.2. Sin embargo en el caso de los complejos entre las pequeñas moléculas disruptivas y proteínas, vale decir que los átomos que contribuyen al valor de  $\rho_D$  (la contribución por parte de la pequeña molécula) son los átomos de carbono que no están enlazados covalentemente a otros átomos más electronegativos que el átomo de carbono mismo. En el caso particular de anillos aromáticos que poseen grupos halógenos (en las pequeñas moléculas) podemos decir que en estos casos particulares todos los átomos de carbono del anillo, como así también los grupos halógenos, son considerados como hidrófobos y por ende contribuyen al valor de  $\rho_D$  (esto se basa en el hecho de que los efectos de resonancia que ocurren en los anillos aromáticos decrecen considerablemente la polaridad de los halógenos enlazados covalentemente al anillo). Estos casos particulares son pocos, y verificamos que nuestros resultados no varían cualitativamente si no consideramos como grupos protectores a estos átomos de halógenos y cuantificamos  $\rho_D$  con las mismas restricciones para grupos carbonaceos que usamos en los complejos proteína-proteína.

### 6.1.2. Resultados del cooperativismo molecular en la disrupción de complejos proteína-proteína por parte de pequeñas moléculas

Analizamos cuatro casos satisfactorios de pequeñas moléculas disruptivas ,ya estudiadas experimentalmente en la Ref. [76], basando nuestro estudio en las interacciones cooperativas de tres cuerpos. Los complejos proteína-droga estudiados son los siguientes: HDM2/Nutlin-3 [101], IL2/SP4206 [102], BCL-XL/ABT-737 [80] y ZipA/Compuesto-1 [96]. Estos sistemas poseen una



significativa relevancia terapéutica y las estructuras 3D, tanto de los sistemas proteína-droga como de los correspondientes sistemas proteína-proteína, ya fueron determinados y depositados en el *Protein Data Bank*. La tabla 6.1 resume los sistemas estudiados.

Ligando	PDB	Afinidad( $\mu\text{M}$ )
<b>HDM2</b>		
p53(aminoácidos 15-29)	1YCR	0,6
Nutlin-3	1RV1	0,09
<b>IL-2</b>		
Receptor de IL-2 cadena- $\alpha$	1Z92	0,0105
SP4206	1PY2	0,06
<b>BCL-X<sub>L</sub></b>		
BAD(aminoácidos 100-126)	2BZW	0,0006
ABT-737	2YXJ	0,0006
<b>ZipA</b>		
FtsZ (aminoácidos 367-383)	1F47	21,6
Compuesto 1	1Y2F	12

Cuadro 6.1: Indicamos el código en el PDB y la afinidad para cada uno de los sistemas estudiados.

En todos los casos llamamos a la proteína *target* como proteína A, el complejo proteína-proteína como AB, donde B corresponde al *binding partner* natural, el cual la pequeña molécula tiende a reemplazar. Finalmente la tentativa droga la llamamos D, y así AD corresponde al complejo entre esta última y la proteína A. Adicionalmente BHB (como venimos nombrando a las puentes de hidrógeno del *backbone*) corresponde a un puente de hidrógeno intramolecular en la proteína A que está involucrado en por lo menos una interacción cooperativa de tres cuerpos con la proteína B y cuyo  $\rho_A < \langle \rho \rangle$  ( $\langle \rho \rangle = 26,6$ , ver Capítulo 4 sección 4.2). Si tenemos en cuenta nuestro análisis previo en complejos proteína-proteína (ver Capítulo 5), estos últimos sitios, en la proteína B, involucrados en interacciones de tres cuerpos con BHBs en la proteína A, representarían predicciones teóricas de supuestos *hot spots* [18].

En la tabla 6.2 se muestra el análisis de las interacciones intermoleculares de tres cuerpos para los complejos formados por la proteína HDM2. Solo se muestran las interacciones relevantes de acuerdo a lo expuesto anteriormente.

De la tabla 6.2 podemos concluir que la molécula Nutlin-3 mimetiza significativamente las interacciones intermoleculares cooperativas por parte del *binding partner* (p53). Los siete BHBs

BHB	$\rho_A$	$\Delta\rho_{AB}$	$\Delta\rho_{AD}$	$\rho_D$
L54-M50	19	3	7	7
F55-K51	24	1	0	2
M62-G58	14	12	10	8
V93-H73	21	8	9	10
I99-E95	18	1	6	1
Y100-H96	20	2	2	3
I103-I99	26	1	2	1

Cuadro 6.2: Interacciones intermoleculares de tres cuerpos para los complejos proteína-proteína y proteína-droga formados por la proteína HDM2.

en la proteína HDM2 son parte de interacciones cooperativas tanto con la proteína p53 como con la pequeña molécula Nutlin-3, de los cuales tres de ellos son dehidrones ( $\rho_A \leq 19$ ). Vale decir que Nutlin-3, adicionalmente forma otras interacciones intermoleculares cooperativas con la proteína *target*, las cuales no están presentes en el complejo natural con la proteína p53, una de ellas con  $\rho_{AD} = 18$  y un valor de  $\Delta\rho_{AD} = 1$  (estas últimas interacciones, en las cuales solo participa Nutlin-3 y no la proteína p53, no se muestran en la tabla 6.2). La Fig. 6.1 a) muestra a la molécula Nutlin-3 (en complejo con la proteína HDM2) junto con los dehidrones (varillas verdes entre el par de carbonos  $\alpha$  de los residuos participantes del BHB) en HDM2 que interactúan con la molécula. En la Fig. 6.1 b) se puede observar nuevamente a la pequeña molécula superpuesta a las cadenas laterales de los aminoácidos de la proteína p53 que interactúan con estos mismos BHBs en el complejo proteína-proteína (las coordenadas de estos residuos se obtienen por superposición del *backbone* de la proteína HDM2 en ambos complejos). Vale decir que estudios experimentales de *alanine scanning* [148] lograron identificar a estos mismos aminoácidos de la proteína p53 como *hot spots* del complejo p53-HDM2, y como vimos en el Capítulo 5, estos mismos aminoácidos fueron también identificados como *hot spots* por nuestro análisis computacional basado en interacciones cooperativas de tres cuerpos [18]. En esta última imagen podemos observar claramente que las porciones hidrofóbicas de la molécula Nutlin-3 se encuentran ocupando la posición de las tres cadenas laterales pertenecientes a los tres *hot spots* de la proteína p53 (Trp 23, Phe 19 y Leu 26). Este caso particular fue notificado anteriormente por los autores de los trabajos [76, 89] cuando la mimetización en la unión de pequeñas moléculas disruptivas fue propuesta. Sin embargo, nuestros resultados explican la razón de esta

mimetización: la exclusión del agua, a través de la formación de interacciones intermoleculares cooperativas de tres cuerpos, que a su vez permiten reforzar los BHBs en la interfase del complejo. Vale la pena aclarar que también existen otros tipos de interacciones convencionales en el proceso de unión, sin embargo, como vimos en el Capítulo 5, este tipo de interacciones triviales no son suficientes para poder racionalizar el proceso de unión entre proteínas y entre pequeñas moléculas y proteínas. En el caso del complejo p53-HDM2 existen tres puentes de hidrógeno intermoleculares y un puente salino intermolecular en la interfase de dicho complejo, mientras que no se detectaron puentes de hidrógeno intermoleculares, ni puentes salinos intermoleculares entre Nutlin-3 y la proteína HDM2.

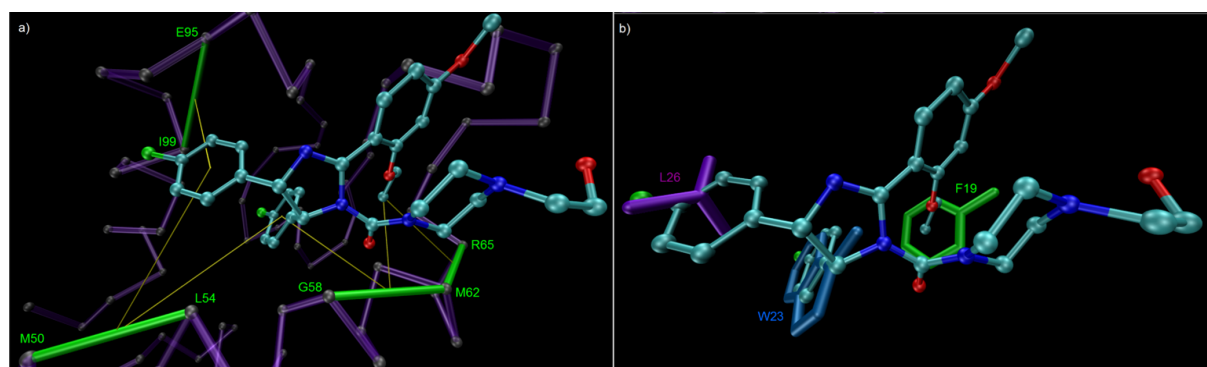


Figura 6.1: a) La proteína HDM2 (solo se representa el *backbone* con varillas violetas y los carbonos  $\alpha$  con esferas grises) en complejo con la molécula Nutlin-3. Se muestran los dehidrones como varillas de color verde que conectan los carbonos  $\alpha$  de ambos residuos que conforman el BHB (en estos casos también se indica en la figura el nombre y número de residuo). Las líneas finas amarillas representan las interacciones intermoleculares de tres cuerpos entre los átomos de la pequeña molécula y los dehidrones de la proteína HDM2. b) Superposición de Nutlin-3 con las cadenas laterales de los *hot spots* del *binding partner* (p53) que interactúan, mediante interacciones cooperativas, con los dehidrones descritos en a).

En la tabla 6.3 se muestran los resultados de nuestro análisis para los complejos formados por la proteína IL-2. Podemos ver que de los cuatro BHBs que interactúan con el *partner* natural (IL-2R $\alpha$ ), con dos de ellos la molécula SP4206 posee interacciones similares en lo que respecta a interacciones intermoleculares cooperativas. Una situación interesante es la del BHB entre Glu 68 y Lys 64 y por otro lado la del BHB entre Glu 61 y Gln 57: En el primero de estos casos la pequeña molécula no posee una interacción directa con el BHB, sino que ocurre lo que podríamos

llamar como “cooperatividad inducida” (ver Fig. 6.2 a)). En el complejo con la molécula SP4206, la proteína IL-2 adopta una conformación local diferente, en esta conformación el BHB Glu 68 - Lys 64 posee una contribución intramolecular, a su valor de  $\rho$ , superior. La cooperatividad inducida representa un mecanismo interesante aunque más sutil, y *a priori* no parece ser fácil incorporar dicho efecto en estrategias para el diseño de drogas. Sin embargo valdría la pena explorar como una pequeña molécula puede producir un pequeño cambio conformacional en la proteína *target* de manera de reforzar intramolecularmente la protección de los BHBs en la interfase. El segundo caso representa el efecto contrario, nuevamente la pequeña molécula no posee una interacción directa con el BHB ( $\rho_D = 0$ ), y además aquí  $\Delta\rho_{AD} < 0$ , por lo tanto la proteína IL-2 adopta una conformación local diferente donde este BHB se encuentra más expuesto.

BHB	$\rho_A$	$\Delta\rho_{AB}$	$\Delta\rho_{AD}$	$\rho_D$
T41-R38	14	2	2	7
F44-A112	24	1	2	1
E61-Q57	18	2	-4	0
E68-K64	18	5	5	0

Cuadro 6.3: Interacciones intermoleculares de tres cuerpos para los complejos proteína-proteína y proteína-droga formados por la proteína IL-2.

En resumen, la pequeña molécula SP4206, interacciona directamente con dos de cuatro BHBs de la proteína IL-2, los cuales interaccionan también con el *binding partner* natural (IL-2R $\alpha$ ). Adicionalmente, existe un par de BHBs que interaccionan con SP4206 y no con IL-2R $\alpha$ . Uno de ellos relevante ya que constituye un dehidrón si no consideramos la contribución al valor de  $\rho$  por parte de la pequeña molécula. Este sitio es el BHB entre los residuos Leu 72 y Glu 68, el cual recibe 7 unidades de  $\rho$  por parte de la molécula SP4206 (para alcanzar un valor total de  $\rho_{AD} = 24$  en el complejo). Este BHB no se encuentra formado en el complejo entre IL-2 y el receptor IL-2R $\alpha$ , probablemente dado a su inestabilidad debido al bajo contenido hidrofóbico ( $\rho$ ) en su entorno. La Fig. 6.3 a) muestra las interacciones intermoleculares cooperativas entre la pequeña molécula y los dehidrones de la proteína IL-2 y en la Fig. 6.3 b) podemos observar la superposición de SP4206 con los *hot spots* en el *binding partner* IL-2R $\alpha$ . Aquí, a diferencia del caso con Nutlin-3, no parece existir una mimetización espacial o física por parte de la

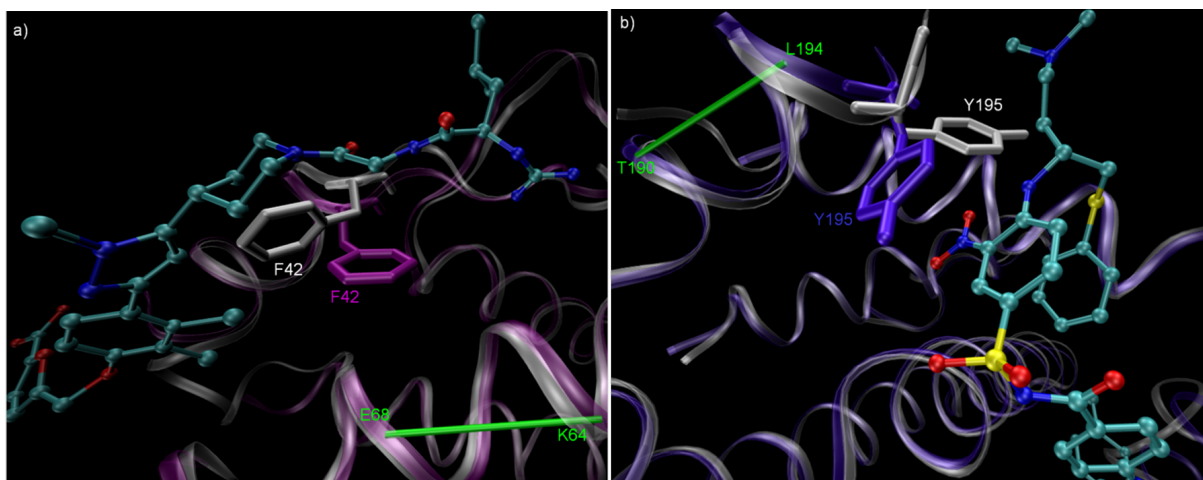


Figura 6.2: Ejemplos de cooperatividad inducida para a) IL-2 y b) BCL-XL; en el complejo con sus respectivas pequeñas moléculas. En ambos casos se representa el *backbone* de la proteína *target*, superpuesto para ambos complejos (proteína-proteína y proteína-droga), además de la pequeña molécula para cada uno de los sistemas (no se representa el *partner* natural). También se destaca el residuo cuya cadena lateral sufre un desplazamiento significativo, luego del *binding* con la pequeña molécula, así consumando una nueva interacción intramolecular (inducida) con un dehidrón de su propia cadena proteica. Una comparación directa de las conformaciones de la proteína *target* en el complejo proteína-proteína y proteína-droga, tanto en a) como en b), nos muestra que comparten básicamente la misma conformación del *backbone* en la interfase y solo la cadena lateral de un residuo (Phe 42 en IL-2, Tyr 195 en BCL-XL) sufre un desplazamiento considerable.

molécula SP4206, sin embargo, como mostramos anteriormente, la molécula SP4206 mimetiza parcialmente las interacciones intermoleculares cooperativas en las que participa IL-2R $\alpha$ .

También debemos mencionar que existen siete puentes de hidrógeno intermoleculares en el complejo IL-2/IL-2R $\alpha$ , ninguno de ellos del tipo *backbone-backbone*. Por otro lado la molécula SP4206 forma tres puentes de hidrógeno intermoleculares, de los cuales dos de ellos reemplazan a los mismos dos HBs intermoleculares con el *partner* IL-2R $\alpha$ . Finalmente, también se encuentran cinco puentes salinos intermoleculares en el complejo IL-2/IL-2R $\alpha$  y solo dos de ellos ocurren también con la pequeña molécula.

En la tabla 6.4 se muestran los resultados obtenidos para los complejos formados por la proteína BCL-XL. Nuevamente aquí podemos ver que la pequeña molécula (ABT-737) mime-

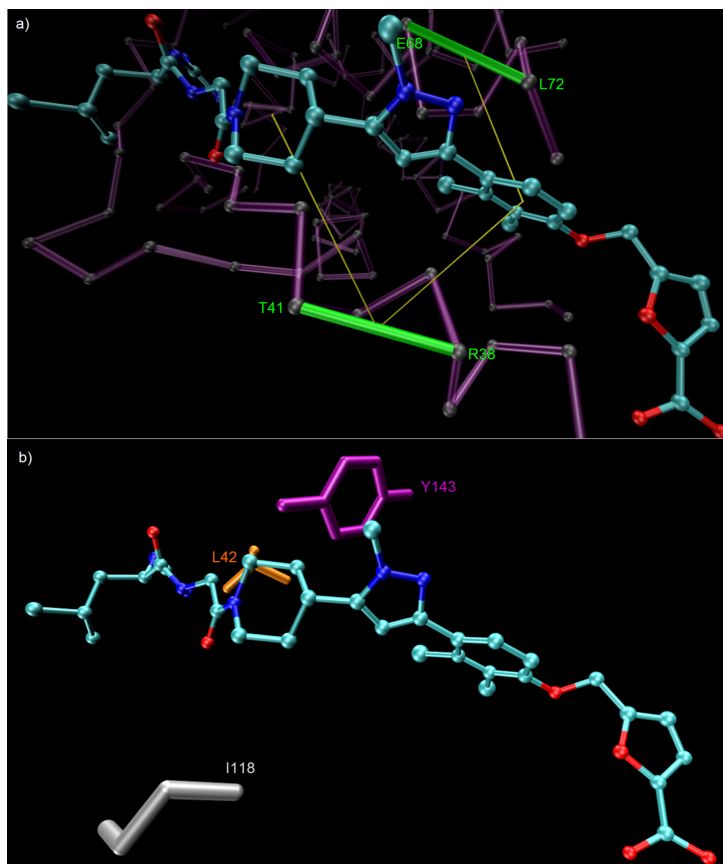


Figura 6.3: Ídem a la Fig. 6.1, pero para los complejos formados por la proteína IL-2.

tiza satisfactoriamente las interacciones cooperativas en las cuales participa el *partner* natural (BAD). En este caso la molécula ABT-737 interactúa cooperativamente con nueve de once BHBs de la proteína BCL-XL que también interactúan con BAD. Cuatro de estos BHBs constituyen dehidrones si solo consideramos el valor de  $\rho_A$ , de estos últimos, la molécula ABT-737 interactúa directamente con tres de ellos, mientras que la interacción con el dehidrón formado por los residuos Leu 194 y Thr 190 no es directa, sino que representa un caso de cooperativismo inducido (ver Fig. 6.2 b)). Además, la molécula ABT-737 interactúa cooperativamente con otros BHBs en BCL-XL, los cuales no interactúan con la proteína BAD, y dos de estos BHBs constituyen dehidrones ( $\rho_A \leq 19$ ), este es el caso de Leu 108 - Ala 104 y Arg 132 - Glu 129. El primero de ellos no se encuentra formado en el complejo proteína-proteína, presumiblemente dada su baja estabilidad debido a su pequeño valor de  $\rho$ ; mientras que el segundo de ellos existe también en el complejo proteína-proteína y posee un pequeño valor de  $\rho$ .

BHB	$\rho_A$	$\Delta\rho_{AB}$	$\Delta\rho_{AD}$	$\rho_D$
A93-R89	25	2	-2	3
E96-E92	18	2	-2	1
R100-E96	17	2	3	1
S122-T118	13	3	-1	0
Q125-Q121	17	5	-2	0
V126-S122	21	6	2	2
E129-Q125	12	3	2	2
L130-V126	23	9	11	6
A142-G138	24	13	14	16
L150-F146	23	1	3	5
L194-T190	18	5	5	0

Cuadro 6.4: Interacciones intermoleculares de tres cuerpos para los complejos proteína-proteína y proteína-droga formados por la proteína BCL-XL.

En la Fig. 6.4 a) podemos observar las interacciones intermoleculares cooperativas entre la molécula ABT-737 y los dehidrones en la proteína BCL-XL; mientras que en la Fig. 6.4 b) se puede observar la superposición de ABT-737 con los residuos de la proteína BAD que interaccionan cooperativamente con los dehidrones en BCL-XL. Adicionalmente en la Fig. 6.4 b) se muestra explícitamente uno de los dehidrones que interacciona tanto con la proteína BAD como con la molécula ABT-737. En este caso, las coordenadas de los aminoácidos de BAD que interaccionan con este dehidrón no se superponen con las coordenadas de la pequeña molécula, es decir que en esta región de la interfase no existe una mimetización espacial o física por parte de ABT-737, sin embargo como muestran los resultados se logra mimetizar satisfactoriamente las interacciones intermoleculares cooperativas. Vale decir que aquí también existen otros tipos de interacciones en los sistemas BCL-XL/BAD y BCL-XL/ABT-737. En el caso del complejo proteína-proteína existen tres HBs intermoleculares, ninguno de ellos del tipo *backbone-backbone*; y adicionalmente dos puentes salinos intermoleculares. En el complejo proteína-droga solo existen dos HBs intermoleculares, ninguno de ellos involucran residuos en BCL-XL que formen HBs intermoleculares con BAD.

Finalmente estudiamos los complejos formados por la proteína ZipA. La molécula disruptiva de la interfase ZipA/FtsZ estudiada se conoce como Compuesto 1. La afinidad de este compuesto ( $12 \mu\text{M}$ ) es mucho menor que en los otros casos estudiados (ver tabla 6.1). Como podemos ver en

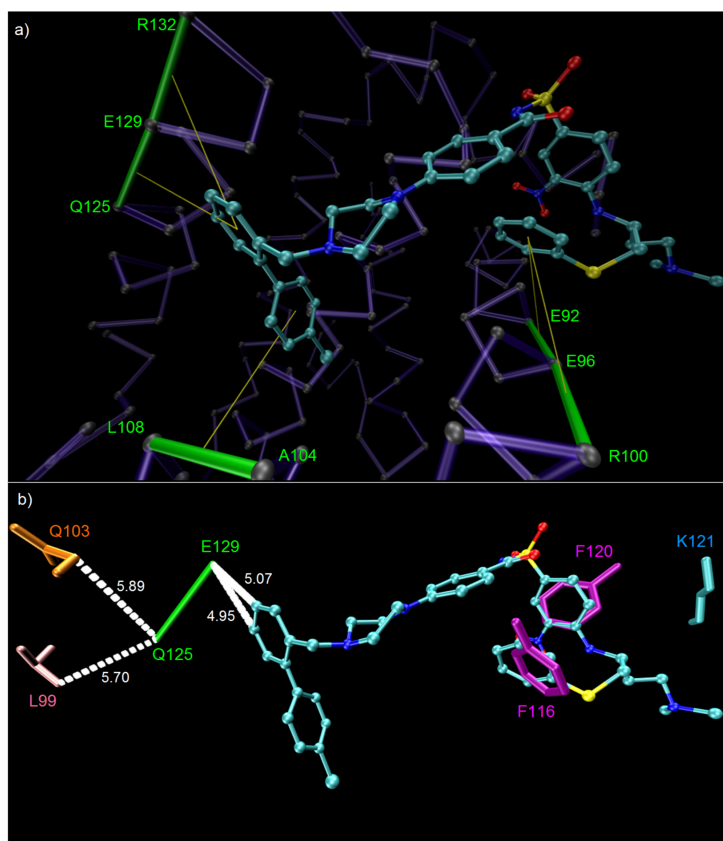


Figura 6.4: Ídem a la Fig. 6.1, pero para los complejos formados por la proteína BCL-XL.

la tabla 6.5, aquí el desempeño de la pequeña molécula es claramente subóptimo, interaccionando solo con uno (un dehidrón,  $\rho_A \leq 19$ ) de los tres BHBs expuestos ( $\rho < \langle \rho \rangle$ ) en la interfase del complejo. Por otro lado la molécula también interacciona con otros BHBs de la proteína ZipA que no interaccionan cooperativamente con el *partner* natural, sin embargo ninguno de ellos se encuentra por debajo del valor crítico de  $\rho_A \leq 19$ . También se muestra, como en los sistemas anteriores, la correspondiente representación 3D (ver Fig. 6.5)

BHB	$\rho_A$	$\Delta\rho_{AB}$	$\Delta\rho_{AD}$	$\rho_D$
A93-R89	23	1	-1	0
E96-E92	18	7	3	4
R100-E96	19	5	-2	0

Cuadro 6.5: Interacciones intermoleculares de tres cuerpos para los complejos proteína-proteína y proteína-droga formados por la proteína ZipA.

Adicionalmente, el complejo ZipA/FtsZ presenta dos BHBs intermoleculares (*backbone-backbone*),



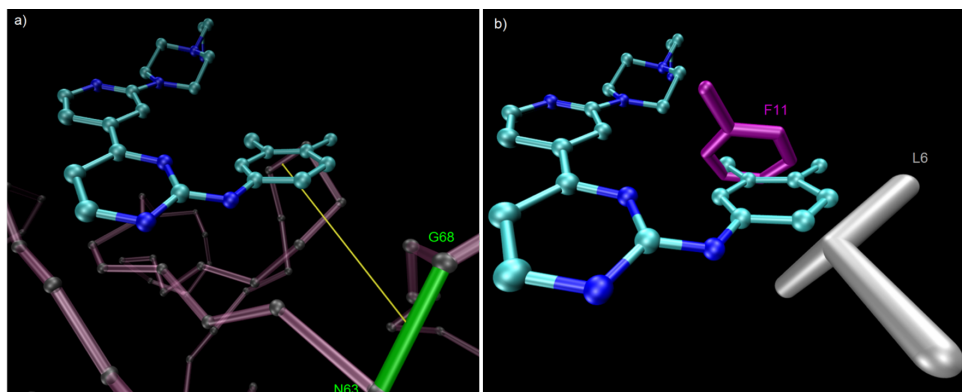


Figura 6.5: Ídem a la Fig. 6.1, pero para los complejos formados por la proteína ZipA.

sin embargo el complejo ZipA/Compuesto 1 no presenta ningún HB. Como vimos en la tabla 6.1, no solo el Compuesto 1 presenta una baja afinidad por la proteína ZipA, de hecho el *partner* natural, FtsZ, también presenta una afinidad pequeña. Así, dado que la afinidad por el *partner* natural es pequeña, no se debería esperar que una buena mimetización por parte de la pequeña molécula resulte en una afinidad elevada para el complejo proteína-droga.

También estudiamos otras moléculas alternativas (de menor afinidad por la proteína *target*), disruptivas de los complejos proteína-proteína para los sistemas en los que participan las proteínas IL-2 y BCL-XL [19]. Estos resultados no se presentan, aunque vale la pena mencionar que en estos casos la mimetización de las interacciones intermoleculares cooperativas, por parte de estas moléculas, es más pobre, correlacionando así con la menor afinidad de estas por su *target* [19]. Vale decir que nuestro método solo puede proporcionar resultados cualitativos. Sin embargo, la tendencia observada para todos los casos parece ser acorde con los datos experimentales de las afinidades de los complejos.

Finalmente, con el fin de tener una idea global de la mimetización de las interacciones cooperativas por parte de las pequeñas moléculas estudiadas anteriormente, calculamos para los cuatro sistemas proteína-proteína estudiados,  $\Delta\rho_{AB}$  para todos los BHBs que interaccionan cooperativamente con el *partner* natural. En este caso no consideramos la restricción de que  $\rho_A$  sea menor a  $\langle\rho\rangle$ , y además calculamos en todos los casos los correspondientes valores de  $\Delta\rho_{AD}$ . La Fig. 6.6 muestra los resultados para los cuatro sistemas, las barras negras representan los valores de  $\Delta\rho_{AB}$  mientras que las barras grises los de  $\Delta\rho_{AD}$  (el eje de las abcisas indica los

diferentes puentes de hidrógeno (HB), de manera genérica, en la proteína *target*). Podemos ver que existe una tendencia global, por parte de las pequeñas moléculas, de “copiar” o mimetizar, en terminos de interacciones intermoleculares cooperativas, el comportamiento de la proteína que desplazan.

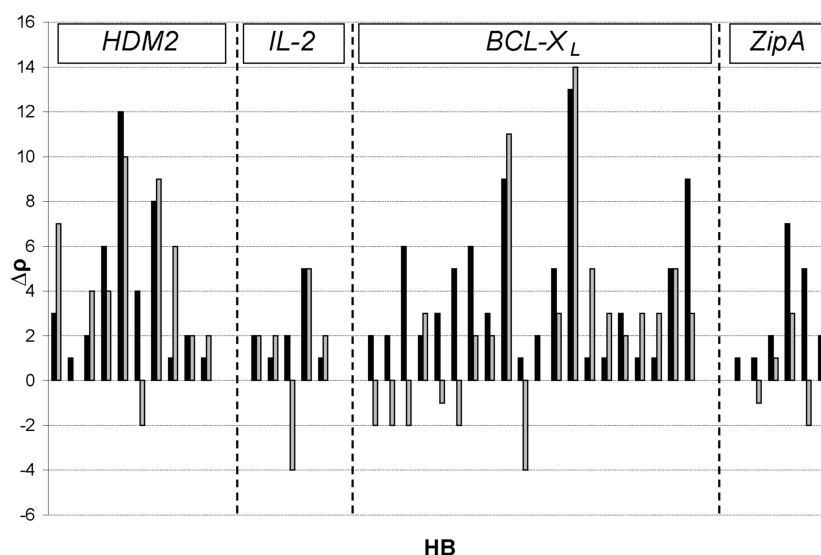


Figura 6.6: Comparación, entre el *partner* natural y la tentativa droga “D” (ambos considerados ligandos), de las interacciones intermoleculares cooperativas formadas entre estos últimos ligandos y los BHBs de su proteína *target* [ $\Delta\rho_{AB}$  (barras negras) y  $\Delta\rho_{AD}$  (barras grises) se indican como  $\Delta\rho$  en el eje de las ordenadas]. Los puentes de hidrógeno (HB en el eje de las abcisas) se indican solamente de manera genérica. Vale decir que este estudio incluye tanto las interacciones directas como indirectas (cooperatividad inducida).

Para cuantificar, en general, cuan buena es la mimetización de las interacciones intermoleculares cooperativas por parte de las cuatro tentativas drogas estudiadas, calculamos la función de correlación cruzada entre los parámetros  $\Delta\rho_{AB}$  y  $\Delta\rho_{AD}$  como  $R = \sum_{i=1}^k R_i$ , donde  $R_i$  es,

$$R_i = \frac{[X_i - \langle X \rangle]}{\left\{ \sum_{j=1}^k [X_j - \langle X \rangle]^2 \right\}^{1/2}} \cdot \frac{[Y_i - \langle Y \rangle]}{\left\{ \sum_{j=1}^k [Y_j - \langle Y \rangle]^2 \right\}^{1/2}},$$

donde  $\sum_{j=1}^k$  es la sumatoria sobre todas las condiciones  $j$  ( $1 \leq j \leq k, j \in \mathbb{N}$ ),  $\langle X \rangle = k^{-1} \sum_{j=1}^k X_j$  e  $\langle Y \rangle = k^{-1} \sum_{j=1}^k Y_j$  son los valores medios de las funciones correspondientes, en este caso  $\Delta\rho_{AB}$  y  $\Delta\rho_{AD}$  respectivamente. Además  $-1 \leq R \leq 1$ . Luego  $R$  es 1 si ambas funciones están

correlacionadas (son proporcionales  $\forall i$  en un factor constante y mayor a cero),  $R$  es -1 si están anticorrelacionadas (son proporcionales  $\forall i$  en un factor constante y menor a cero), ó  $R$  es cero si no hay correlación. Obtuvimos un valor de  $R = 0,6$ , el cual habla de una buena correlación entre ambas funciones [19].

## 6.2. El concepto de cooperativismo como estrategia de recuperación de la función de la proteína p53 con la mutación oncogénica Y220C

Esta sección está dedicada a estudiar la proteína p53 con la mutación oncogénica Y220C en base al concepto de cooperativismo. Nuestros resultados computacionales racionalizan la naturaleza de la mutación Y220C, explican el *binding* de una pequeña molécula ya diseñada para reestablecer la función de la proteína p53 y proveen indicios que ayudan a mejorar esta estrategia de recuperación y que también son de gran utilidad en el diseño de nuevas drogas o en técnicas de reingeniería.

### 6.2.1. El supresor tumoral p53: Un objetivo terapéutico clave en el tratamiento del cancer

El supresor tumoral p53 es conocido también, por el rol que desempeña, como el “guardian del genoma”. En respuesta a varias señales de estrés, la proteína p53 interacciona con una específica secuencia del ADN y regula así la transcripción de un importante grupo de genes involucrados en procesos celulares como: el **arresto del ciclo celular**, donde entendemos por *ciclo celular* a un conjunto ordenado de sucesos que conducen al crecimiento de la célula y la división en dos células hijas, así esta “pausa” en la progresión del ciclo celular da tiempo a reparar los daños producidos en el ADN; la **senescencia** (parada permanente en el ciclo celular); y la **apoptosis** (la destrucción o muerte celular programada provocada por ella misma)[4]. La inactivación del p53 es crítica para la formación y desarrollo de la mayoría de los tumores. En alrededor de la mitad de los casos de cancer en seres humanos, la proteína p53 presenta una mutación, y en

particular estos casos muestran una incrementada resistencia a los tratamientos de quimioterapia convencionales [4]. Es por esto que la restauración farmacológica de la función del p53 es una estrategia necesaria y prometedora para el desarrollo de nuevas terapias de tratamiento contra el cancer[4]. En el año 2011, de acuerdo a la Ref. [49], se estimó que aproximadamente 27 millones de personas padecen de algún tipo de cancer en el cual la actividad de la proteína p53 se ve disminuída [49]. En la mitad de estos casos, la proteína p53 no presenta una mutación, sin embargo su actividad se ve suprimida parcialmente. Uno de los mecanismos propuestos para esta disminuída actividad se basan en la sobreexpresión de los reguladores negativos del p53 como la proteína MDM2 [49]. En la otra mitad de los casos, donde la proteína p53 se encuentra inactiva, la misma presenta una mutación y en  $\sim 95\%$  de estos últimos casos las mutaciones yacen en el dominio de *binding* al ADN [49]. La investigación, con el fin de una restauración de la actividad del p53 vía una estrategia farmacológica, llevó al desarrollo de dos enfoques alternativos basados en la utilización de pequeñas moléculas: la estabilización del p53 mutante con el fin de recuperar su actividad, y la inhibición de la proteína MDM2 [49] (ver subsección 6.1.2).

### 6.2.2. Reactivación del p53 mutante

Se sabe que aproximadamente en un tercio de las mutaciones, ocurridas en la proteína p53, la pérdida de actividad de la misma es consecuencia de una disminución en la temperatura de *melting* ( $T_m$ ) del dominio de *binding* al ADN [59, 149, 137], produciendo así, una rápida desnaturalización de la proteína en la célula [115] (de hecho el p53 *wild-type*, esto es, la proteína sin mutaciones, por sí misma posee una baja estabilidad [59, 149, 137, 115, 90, 69, 50]). La mutación Y220C disminuye significativamente la estabilidad termodinámica de la proteína y además es una de las mutaciones oncogénicas más frecuentes en la proteína p53. El sitio donde ocurre esta mutación se encuentra en la superficie de la proteína y distante del sitio de *binding* con el ADN, por lo tanto no afecta la estructura de esta última región [90]. Además, dicha mutación de una tirosina a cisteína, promueve la formación de una cavidad en la superficie de la proteína, hecho que facilita el tratamiento farmacológico de esta mutación. Por lo tanto, el *binding* de una pequeña molécula en esta cavidad podría estabilizar al p53 mutante, sin afectar

la función del p53 y sin interferir con las moléculas de p53 que no presentan esta mutación (p53 *wild-type*). Este enfoque ha sido explotado por el grupo del Prof. Alan Fersht en una serie de excelentes trabajos [59, 149, 137, 115, 90, 69]. El grupo de Fersht encontró una tentativa droga, llamada PhiKan083, un derivado del carbazol que se une a la cavidad donde yace la mutación Y220C [69], con una afinidad moderada ( $K_d \sim 150\mu\text{M}$ ), elevando la  $T_m$  y disminuyendo la velocidad con que el mutante p53-Y220C se desnaturaliza.

En base al concepto de cooperativismo ya descrito, estudiamos el impacto producido por la mutación oncogénica Y220C. Analizamos computacionalmente las interacciones cooperativas en la proteína T-p53C sin y con la mutación Y220C (PDB: 1UOL y PDB: 2J1X respectivamente), la cual ya fue estudiada experimentalmente por el grupo de Alan Fersht. La proteína T-p53C es una versión superestable de la proteína p53 *wild-type*, y su estructura es virtualmente idéntica a esta última más allá de algunas diferencias en la secuencia de su cadena principal [59]. También analizamos las interacciones cooperativas intermoleculares formadas en el complejo entre la proteína T-p53C, con la mutación Y220C (T-p53C-Y220C), y la pequeña molécula PhiKan083 (PDB: 2VUK). La tabla 6.6 resume estos resultados. En este caso llamaremos a la proteína T-p53C sin la mutación Y220C como proteína **A** *wild-type* (**A-wt**); a la proteína T-p53C con la mutación oncogénica Y220C simplemente la llamaremos proteína **A** y a la pequeña molécula PhiKan083 la llamaremos **D**. Así  $\rho_{A-wt}$  serán las interacciones cooperativas intramoleculares para un determinado BHB en la proteína T-p53C;  $\rho_A$  serán las mismas interacciones pero para la proteína con la mutación Y220C (T-p53C-Y220C) y  $\rho_{AD}$  serán las interacciones cooperativas tanto intramoleculares como intermoleculares para un dado BHB en el complejo formado entre la proteína **A** (T-p53C-Y220C) y la pequeña molécula **D** (PhiKan083).

BHB	$\rho_{A-wt}$	$\rho_A$	$\rho_{AD}$
R110-W146	24	24	28
L145-T230	18	17	19
Y/C220-T155	23	18	20

Cuadro 6.6: Interacciones cooperativas de tres cuerpos de los BHBs ubicados en el sitio de ocurrencia de la mutación Y220C y que también interaccionan cooperativamente con la molécula PhiKan083, para los sistemas T-p53C, T-p53C-Y220C y T-p53C-Y220C/PhiKan083,  $\rho_{A-wt}$ ,  $\rho_A$  y  $\rho_{AD}$  respectivamente.

Los resultados en la tabla 6.6 nos muestran que la mutación Y220C, que provoca la formación de una cavidad en la superficie de la proteína, también conduce a la creación de un dehidrón. El puente de hidrógeno entre los aminoácidos Tyr 220 y Thr 155, inicialmente con un valor de  $\rho_{A-wt} = 23$ , luego de la mutación por Cys posee un valor de  $\rho_A = 18$ . Dicho valor es menor al valor medio de  $\rho$  menos  $\sigma$  para la distribución de BHBs [ $\rho_A < \langle \rho \rangle - \sigma$  (ver Capítulo 4, sección 4.2)], así la mutación oncogénica Y220C implica la creación de un dehidrón en la cadena proteica. Más aún, podemos ver también que el BHB entre Leu 145 y Thr 230 (dehidrón vecinal al sitio de la mutación Y220C) también presenta un decaimiento en su valor de  $\rho$  luego de la mutación Y220C (de  $\rho_{A-wt} = 18$  a  $\rho_A = 17$ ). Así la creación de una vulnerabilidad (un dehidrón), como la disminución del nivel de exclusión de agua en el sitio de la mutación, implican una desestabilización [2, 114] y un incremento del desorden local [79] de la proteína.

Por otro lado, la molécula PhiKan083 reestablece la función del p53 estabilizando la estructura de la proteína. De la tabla 6.6 podemos deducir que la molécula se une a la proteína precisamente en el sitio donde se encuentran los dehidrones Cys 220 - Thr 155 y Leu 145 - Thr230 interaccionando cooperativamente con estos BHBs. Además PhiKan083 eleva el valor de  $\rho$  para el primero de estos dehidrones por encima del valor de corte  $\rho = 19$ , así de esta manera remueve el dehidrón creado tras la mutación oncogénica Y220C [15]. Adicionalmente, también estudiamos otros complejos entre T-p53-Y220C con otros fragmentos o pequeñas moléculas, estudiadas experimentalmente por el grupo de Fersht, nombradas como “Fragmento 15” y “Fragmento 2”. Se encontró que estos fragmentos se unen en una posición cercana a la que se une la molécula PhiKan083 y que también interaccionan cooperativamente con el dehidrón creado tras la mutación oncogénica (estos resultados no se muestran).

En la Fig. 6.7 se muestra el complejo entre la molécula PhiKan083 y T-p53C-Y220C en el sitio de la mutación Y220C. Indicamos las interacciones cooperativas intermoleculares con líneas entre los átomos de la molécula y el correspondiente BHB de la proteína. Podemos ver que PhiKan083 interacciona con el dehidrón Cys 220-Thr 155 anisotrópicamente, es decir que hacia un lado de la interacción el agua tiene acceso, por lo que todavía el sitio presenta cierta accesibilidad al agua. En los estudios del grupo de Alan Fersht [59], ellos encontraron que algunos fragmentos (como el Fragmento 15) interaccionaban también desde el otro lado del dehidrón

Cys 220 - Thr 155 penetrando más en la cavidad de la proteína T-p53C-Y220C. Así, a pesar de que PhiKan083 eleva el valor de  $\rho$  del dehidrón en cuestión por encima del valor de corte  $\rho > 19$ , el mismo presenta todavía cierta accesibilidad al ataque del agua en la cavidad creada tras la mutación Y220C.

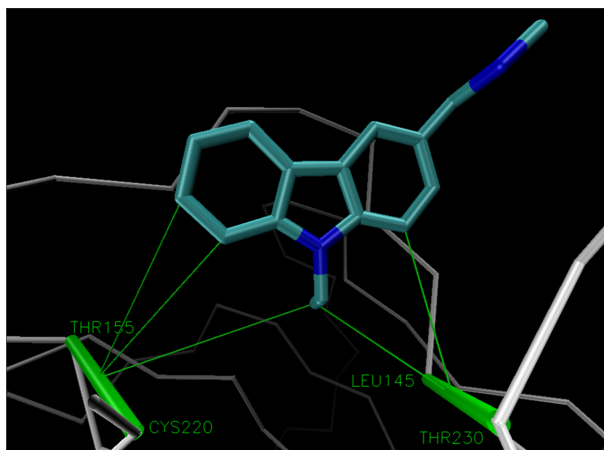


Figura 6.7: Interacciones cooperativas intermoleculares en el complejo entre PhiKan083 y la proteína T-p53C-Y220C (PDB: 2VUK)

De manera de ampliar nuestro análisis básico de cooperativismo y comprender mejor la naturaleza desestabilizante de la mutación oncogénica Y220C, realizamos simulaciones de dinámica molecular para los sistemas T-p53C-Y220C (PDB: 2J1X) y T-p53C-Y220C/PhiKan083 (PDB: 2VUK). Para ello utilizamos el paquete de simulación AMBER 11 [83]. Utilizamos el campo de fuerza ff99SB [84] para proteínas y para el caso del complejo con la pequeña molécula PhiKan083 combinamos dicho campo de fuerza con el GAFF (*Generalized Amber Force Field*) [98]. Los parámetros de entrada para la molécula orgánica se generaron con el subprograma ANTECHAMBER [85]. En todos los casos el modelo de agua utilizado es TIP5P ( $\sim 10000$  moléculas de agua) [161], la temperatura de trabajo es  $T = 300$  K, el valor medio de la presión es  $p = 1$  bar y la densidad media  $\sim 1.0$  kg/dm<sup>3</sup>. La equilibración en todos los sistemas se llevó a cabo en dos pasos siguiendo el mismo protocolo y utilizando los mismos parámetros que en el Capítulo 3, sección 3.2, subsección 3.2.1.

Por su parte, también realizamos una modificación de la molécula PhiKan083, en tal sentido partimos del PDB original 2VUK, correspondiente al complejo T-p53C-Y220/PhiKan083, mo-

dificamos la pequeña molécula y posteriormente generamos los parámetros de entrada para la simulación mediante ANTECHAMBER. A partir de aquí el protocolo que se siguió es el mismo que para los otros dos sistemas.

La Fig. 6.8 a) muestra la distribución de distancia  $d$  entre el átomo de oxígeno carbonílico del aminoácido Thr 155 y el átomo de oxígeno de la molécula de agua más cercana. Recordemos que el residuo Thr 155 participa de un BHB con el residuo Cys 220, tanto en la proteína libre T-p53C-Y220C (PDB: 2J1X) como en la proteína ligada a PhiKan083 (PDB: 2VUK). En el primero de estos casos el BHB Cys 220 - Thr 155 es un dehidrón, mientras que en el segundo caso no ( $\rho \geq 19$  para dicho BHB) (ver  $\rho_{AD}$  en la tabla 6.6). Es evidente que las moléculas de agua se acercan más al dehidrón de la proteína T-p53C-Y220C (proteína libre, PDB: 2J1X) con el fin de atacar la interacción intramolecular. De hecho el pico de la distribución es compatible con la formación de un puente de hidrógeno agua-proteína, donde el agua actúa como donador [20] (ver Capítulo 3, subsección 3.2.2).

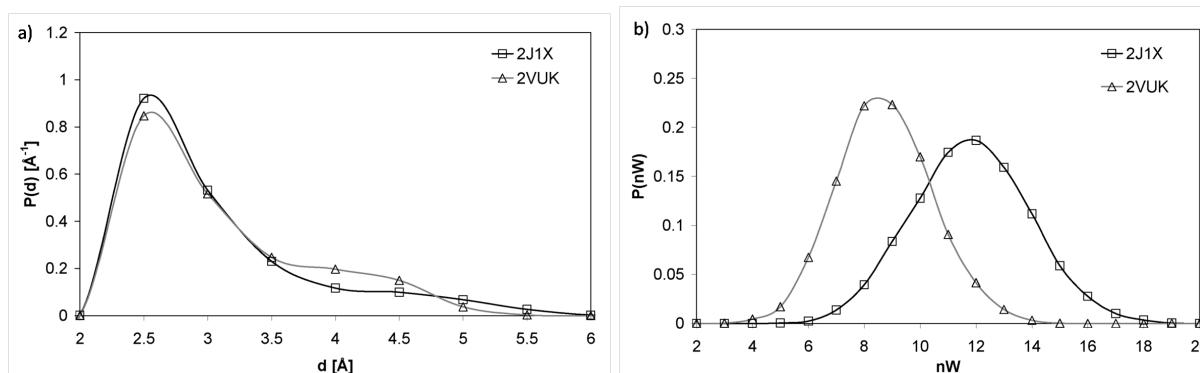


Figura 6.8: a) Distribución de distancia entre el átomo de oxígeno carbonílico del aminoácido Thr 155 y el oxígeno de la molécula de agua más cercana. Se muestra la curva para el complejo entre T-p53C-Y220C/PhiKan083 (PDB: 2VUK) y para la proteína libre T-p53C-Y220C (PDB: 2J1X). b) Para los mismos sistemas mencionados anteriormente se muestra la distribución del número de moléculas de agua en el dominio de desolvatación del BHB Cys 220 - Thr 155.

Más aún, la Fig. 6.8 b) muestra la distribución del número de moléculas de agua ( $nW$ ) en el dominio de desolvatación (sobre el dominio de desolvatación ver Capítulo 4, sección 4.2) del BHB Cys 220 - Thr 155. Como esperábamos, el dehidrón Cys 220 - Thr 155 en la proteína libre posee una población superior de moléculas de agua con respecto al mismo BHB en el complejo



con la pequeña molécula PhiKan083.

A su vez, la Fig. 6.9 a) b) y c) muestra la evolución en la distancia, entre el átomo de nitrógeno del grupo amino del aminoácido Cys 220 y el oxígeno carbonílico del aminoácido Thr 155 ( $dNO$ ), a lo largo de los 2 ns de trayectoria. Podemos ver que en el caso de la proteína libre (PDB: 2J1X) Fig. 6.9 a), las fluctuaciones son más significativas. De hecho, el BHB se rompe varias veces durante el intervalo de tiempo que va desde los 500 ps a los 1000 ps. La creación del dehidrón Cys 220 - Thr 155 luego de la mutación oncogénica Y220C promueve el desorden local debido a que una interacción intramolecular (el BHB en cuestión) se ve desestabilizada con una pérdida temporal de la estructura local del sitio (el BHB dehidrónico Cys 220 - Thr 155 representa un contacto intramolecular inestable que puede sufrir fácilmente una disrupción por parte del agua). Así, nuestros resultados proveen una posible explicación de la naturaleza oncogénica de la mutación Y220C, la cual desestabiliza a la proteína y disminuye su temperatura de *melting*  $T_m$ . La Fig. 6.10 muestra una configuración para la cual la distancia  $dNO$  supera el criterio de corte ( $r_{th} = 3,5\text{\AA}$ ) aceptado para un puente de hidrógeno. Vemos que el agua puede intercalarse entre ambos extremos del BHB Cys 220 - Thr 155 y formar un par de nuevos puentes de hidrógeno con la proteína,  $O \cdots H-N$  con Cys 220 y  $C=O \cdots H-O$  con Thr 155.

Por otro lado, en presencia de la molécula PhiKan083 (T-p53C-Y220C/PhiKan083 PDB: 2VUK), la cual interacciona cooperativamente con el BHB Cys 220 - Thr 155, la distancia  $dNO$  (ver Fig. 6.9 b)) para dicho BHB es más estable a lo largo de la trayectoria. Además las fluctuaciones alrededor del correspondiente valor medio son menores (ver valores de  $\sigma$  en la leyenda de la Fig. 6.9). Vale decir que el valor medio  $\langle r \rangle < r_{th}$ . Estos resultados racionalizan el comportamiento de la molécula PhiKan083, la cual estabiliza a la proteína T-p53C-Y220C y eleva su  $T_m$ , reestableciendo así su función [15].

También vale la pena destacar que, incluso en presencia de la molécula PhiKan083, encontramos moléculas de agua muy próximas al BHB Cys 220 - Thr 155, las cuales se encuentran en la región de la cavidad no ocupada por la pequeña molécula. Así, podríamos decir que existe una cierta anisotropía en la interacción de dicho BHB con PhiKan083. Por lo tanto, aunque el valor de  $\rho$  para dicho BHB supera el valor crítico de  $\rho = 19$  tras la interacción con PhiKan083, todavía sigue siendo accesible al ataque del agua. De hecho la Fig. 6.9 b) muestra que todavía,

en presencia de PhiKan083, la distancia  $dNO$  supera los 3,5 Å repetidas veces. Por lo tanto, a partir de esta observación podemos afirmar que no solo es importante la magnitud del parámetro  $\rho$  para un determinado BHB, sino también el grado de isotropismo en cuanto a las interacciones cooperativas en las que participa. Este hecho señala la posibilidad de rediseñar racionalmente la pequeña molécula PhiKan083. Por lo tanto, en base al concepto de cooperativismo, rediseñamos la tentativa droga con el objetivo de mejorar su desempeño y así reforzar la exclusión del agua circundante a la interacción Cys 220 - Thr 155. En consecuencia es esencial extender la molécula desde una posición que permita cubrir el resto de la cavidad generada tras la mutación Y220C. La Fig. 6.11 muestra dos modificaciones posibles (molécula A y molécula B). La Fig. 6.9 c) muestra la evolución de la distancia  $dNO$  para el complejo entre la proteína Tp53C-Y220C y la molécula B. Podemos observar que el desempeño de la nueva molécula es superior con respecto a PhiKan083. En este caso las fluctuaciones en el valor de  $dNO$  son menores (ver valores medios  $\mu$  y desviaciones estandar  $\sigma$  en la leyenda de la Fig. 6.9). Vale decir que la molécula A también presenta una mejoría respecto a PhiKan083, sin embargo el desempeño de la molécula B es superior.

La Fig. 6.12 a) muestra la distribución de probabilidad de la distancia  $dNO$  durante las simulaciones para el BHB Cys 220 - Thr 155 en la proteína libre (PDB: 2J1X) y en la proteína ligada a PhiKan083 (PDB: 2VUK). También se incluye el comportamiento promedio de los BHBs en la proteína libre (PDB: 2J1X) clasificándolos de acuerdo a su valor de  $\rho$  en dehidrones ( $\rho \leq 19$ ) y no dehidrones ( $\rho > 19$ ). De estas dos últimas distribuciones se puede ver claramente que los dehidrones presentan una cola más extensa por encima del corte  $r_{th} = 3,5$  Å, mientras que la distribución de probabilidad correspondiente a los BHBs con  $\rho > 19$  decae más rápidamente. Por otro lado también podemos ver que en el caso del complejo con PhiKan083 (PDB: 2VUK), el BHB Cys 220 - Thr 155 presenta una distribución de probabilidad cuya cola decae más rápidamente comparando con la distribución correspondiente al mismo BHB en la proteína libre (PDB: 2J1X). Finalmente la Fig. 6.12 b) muestra la mejora que produce la molécula B en la distancia N-O del BHB Cys 220 - Thr 155. Podemos ver que la nueva molécula, la cual interacciona cooperativamente con el BHB Cys 220 - Thr 155, produce una distribución similar a aquella para los BHB con  $\rho > 19$  en la proteína libre (PDB: 2J1X), incrementando así la

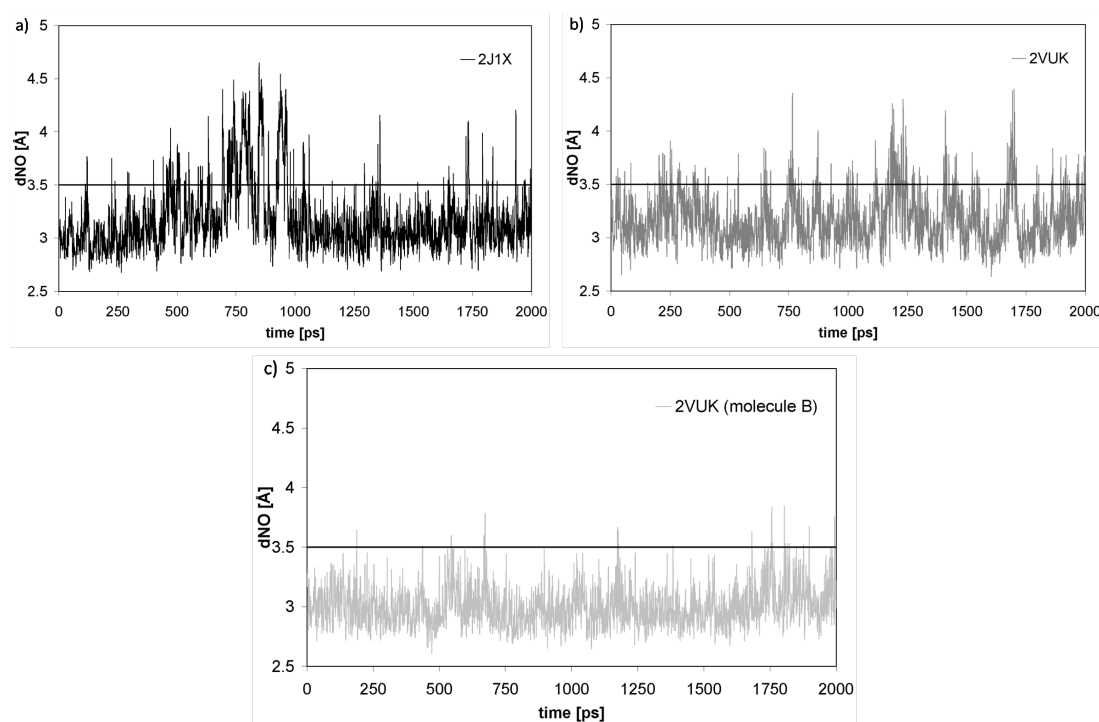


Figura 6.9: a), b), y c) Evolución temporal de la distancia ( $dNO$ ) entre el átomo de nitrógeno del residuo Cys 220 y el átomo de oxígeno del residuo Thr 155, ambos átomos de la cadena principal (*backbone*), para el BHB Cys 220 - Thr 155. a)  $\mu = 3,169$  Å y  $\sigma = 0,323$  Å; b)  $\mu = 3,165$  Å y  $\sigma = 0,254$  Å; c)  $\mu = 3,003$  Å y  $\sigma = 0,173$  Å. Se considera que un BHB está formado si su distancia  $dNO$  es menor a  $3,5$  Å.

estabilidad de dicho BHB.

Adicionalmente en la Fig. 6.13 a) se muestra la distribución del número de moléculas de agua  $nW$  en el dominio de desolvatación del BHB Cys 220 - Thr 155. Específicamente se muestran las distribuciones para dicho BHB en la proteína libre T-p53C-Y220C (PDB: 2J1X), la proteína en complejo con PhiKan083 (PDB: 2VUK) y finalmente el complejo entre la proteína y la nueva molécula B. Se puede observar que la exclusión del agua, en el dominio de desolvatación del BHB, por parte de la molécula B es más efectiva (la distribución se desplaza hacia menores valores de  $nW$ ) que en el caso de la molécula PhiKan083. La Fig. 6.13 b) muestra una configuración de la segunda modificación de PhiKan083 (molécula B) en la cavidad, generada tras la mutación Y220C. Se resaltan las interacciones intermoleculares cooperativas (líneas finas verdes) como así también los puentes de hidrógeno intermoleculares que forma la molécula con la proteína,

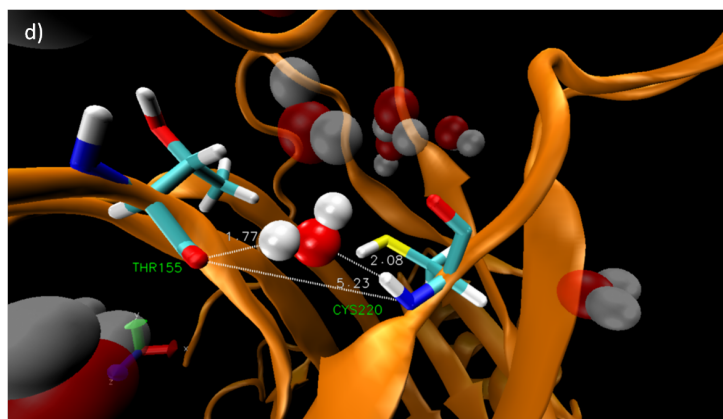


Figura 6.10: Una molécula de agua puede intercalarse entre ambos extremos del BHB Cys 220 - Thr 155 y formar un par de nuevos puentes de hidrógeno con la proteína,  $O \cdots H-N$  con Cys 220 y  $C=O \cdots H-O$  con Thr 155.

específicamente con los aminoácidos Pro 151 y Asp 228 (línea punteada). Es interesante recalcar que el grupo de Alan Fersht, en uno de sus trabajos, mostró que se encontraba una densidad superior de moléculas de isopropanol alrededor del aminoácido Pro 151 al realizar simulaciones de la proteína T-p53C-Y220C en una mezcla de agua e isopropanol [59].

Por último, con el objetivo de cuantificar el desempeño de la nueva molécula B con respecto a PhiKan083, estimamos mediante el programa de AMBER MM-GBSA (*Molecular Mechanics/Generalized Born Surface Area* [138, 66]), la energía libre de formación de cada uno de los complejos proteína-droga. Dentro de este formalismo, el valor de la energía libre de *binding* o formación de cada uno de los complejos ( $\Delta G_{\text{bind}}$ ) se calcula de acuerdo a la siguiente ecuación:

$$\Delta G_{\text{bind}} = G_{\text{com}} - G_{\text{rec}} - G_{\text{lig}}$$

donde “com”, “rec” y “lig” hacen referencia, al complejo, al receptor y al ligando respectivamente. La energía libre de cada uno de ellos se estima como la suma de los siguientes terminos:

$$G = \langle \text{EMM} \rangle + \langle G_{\text{psolv}} \rangle + \langle G_{\text{npsolv}} \rangle$$

$\langle \text{EMM} \rangle$  es el valor medio de la energía potencial de la molécula, la cual se calcula como:

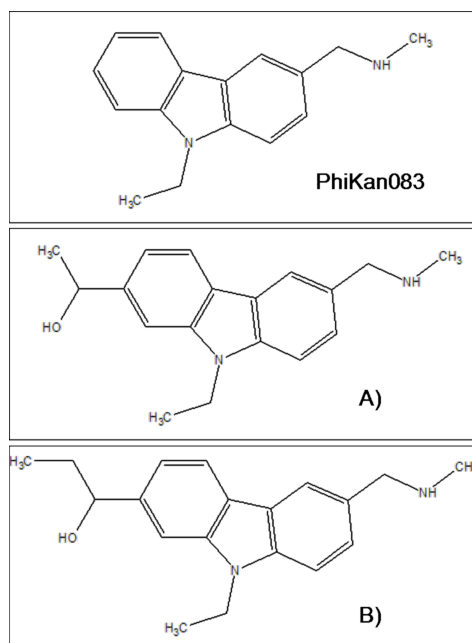


Figura 6.11: Estructura molecular de PhiKan083 y de las dos modificaciones sugeridas por nuestro grupo. Los valores medios calculados para la energía libre de unión de cada una de estas moléculas con la proteína T-p53C-Y220C son: PhiKan083 (PDB: 2VUK)  $\langle \Delta G_{\text{binding}} \rangle = -29,46$  kcal/mol (desviación estandar  $DES = 2,05$  kcal/mol); primera modificación de PhiKan083, molécula A  $\langle \Delta G_{\text{binding}} \rangle = -29,80$  kcal/mol (desviación estandar  $DES = 2,71$  kcal/mol); segunda modificación de PhiKan083, molécula B  $\langle \Delta G_{\text{binding}} \rangle = -35,33$  kcal/mol (desviación estandar  $DES = 2,81$  kcal/mol).

$EMM = E_{\text{cov}} + E_{\text{ncov}}$ , donde  $E_{\text{cov}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}}$  es la energía de cada uno de los enlaces covalentes de la molécula; y  $E_{\text{ncov}} = E_{\text{elec}} + E_{\text{VDW}}$  es la energía involucrada en cada una de las interacciones no covalentes en las que participa la molécula, ya sean interacciones electrostáticas o de van der Waals. Luego  $\langle G_{\text{psolv}} \rangle$  es el valor medio de la contribución polar a la energía de solvatación y  $\langle G_{\text{npsolv}} \rangle$  es el valor medio de la contribución no polar a dicha energía. Debemos tener presente que el valor obtenido de  $\Delta G_{\text{bind}}$  no es igual al valor real de la energía de formación del complejo, ya que no se estima en ningún momento la contribución (desfavorable) entrópica a dicha energía. Sin embargo este resultado es útil para comparar entre sistemas similares.

Para el análisis se tomaron configuraciones cada 10 ps de una trayectoria de dinámica molecular de 2 ns para el complejo proteína-droga, resultando así en 2000 configuraciones. Se

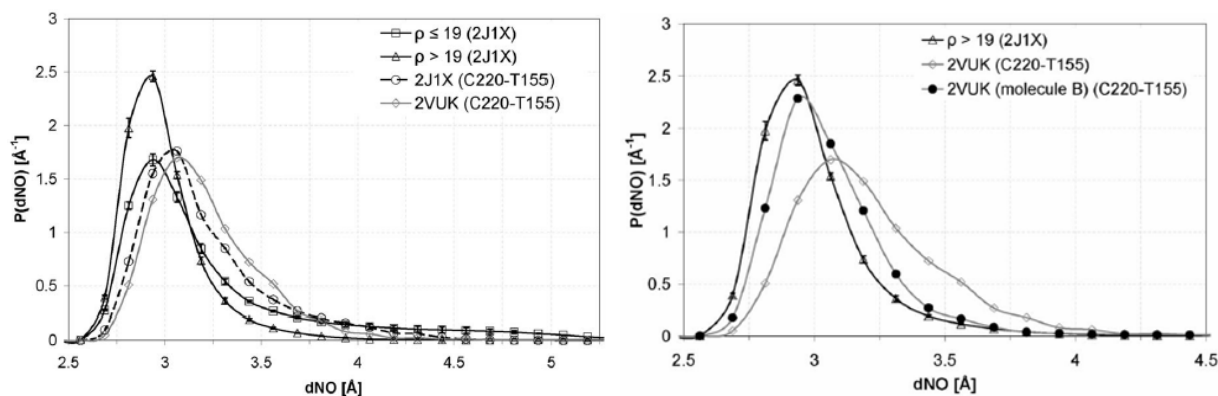


Figura 6.12: a) Distribución de distancia entre el átomo de N del grupo amino y el oxígeno del grupo carbonilo para los siguientes casos: BHBs con  $\rho > 19$  en la proteína T-p53C-Y220C (PDB: 2J1X) ( $\mu = 2,949 \text{ \AA}$  y  $\sigma = 0,276 \text{ \AA}$ ); BHBs con  $\rho \leq 19$  en la proteína T-p53C-Y220C (PDB: 2J1X) ( $\mu = 3,208 \text{ \AA}$  y  $\sigma = 0,564 \text{ \AA}$ ); BHB Cys 220 - Thr 155 en el complejo T-p53C-Y220C/PhiKan083 (PDB: 2VUK) ( $\mu = 3,150 \text{ \AA}$  y  $\sigma = 0,288 \text{ \AA}$ ) y el BHB Cys 220 - Thr 155 en la proteína T-p53C-Y220C (PDB: 2J1X) ( $\mu = 3,143 \text{ \AA}$  y  $\sigma = 0,339 \text{ \AA}$ ). b) ídem a) pero en este caso se incluye la distribución para el complejo entre la proteína T-p53C-Y220C y la molécula B ( $\mu = 3,000 \text{ \AA}$  y  $\sigma = 0,235 \text{ \AA}$ ).

consideró que no existe un cambio conformacional significativo, tanto en la proteína como en la pequeña molécula, entre los estados ligado y no ligado. Así, luego de hacer esta aproximación, solo es necesario una única trayectoria de dinámica molecular para poder estimar la energía de formación del complejo.

Las energías se obtienen utilizando el modulo MM-GBSA de AMBER 11 [12]. La energía potencial del complejo, el receptor y el ligando se obtiene mediante el modulo SANDER de AMBER 11. La contribución polar a la energía libre de solvatación del complejo, el receptor y el ligando se calcula mediante el método GB (*Generalized Born*) [103, 104] implementado en AMBER 11. Se utilizó el modelo GB de Onufriev para el cual el parámetro  $igb = 2$  [12]. Se considera que la constante dieléctrica interna al soluto es igual a 1 mientras que la constante dieléctrica en el exterior del soluto es igual a 80. Por último, el término correspondiente a la contribución no polar a la energía libre de solvatación del complejo, el receptor y el ligando se calcula a partir del área superficial accesible al solvente (SASA) [164], utilizando la siguiente ecuación:  $G_{\text{npolv}} = \delta \text{SASA} + b$ , donde el SASA se determina con el programa Molsurf [164]

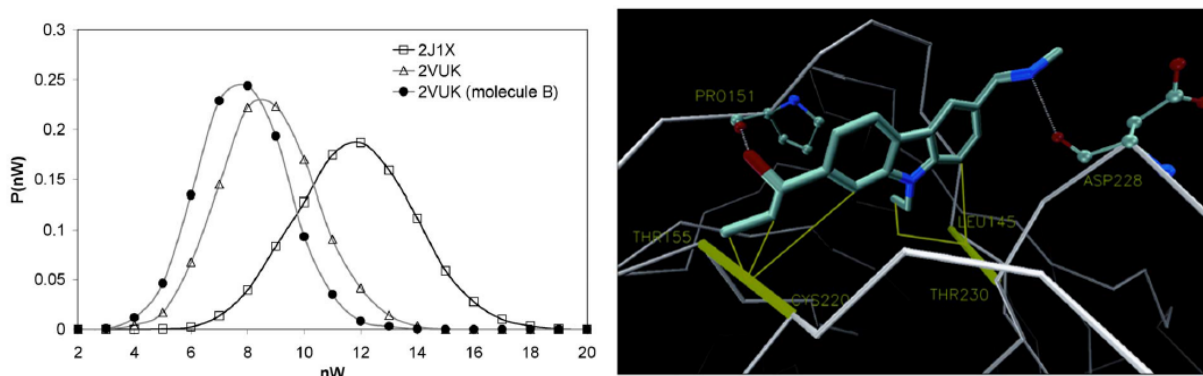


Figura 6.13: a) Distribución del número de moléculas de agua en el dominio de desolvatación del BHB Cys 220 - Thr 155. Se muestra la curva para el complejo T-p53C-Y220C/PhiKan083 (PDB: 2VUK), para la proteína libre T-p53C-Y220C (PDB: 2J1X) y ahora también para el complejo entre la proteína T-p53C-Y220C y la segunda modificación de PhiKan083 (molécula B). b) Interacciones cooperativas intermoleculares en el complejo entre la segunda modificación de PhiKan083 (molécula B) y la proteína T-p53C-Y220C. También se esquematizan, con líneas punteadas, los dos puentes de hidrógeno intermoleculares que forma la molécula B con los residuos Pro 151 y Asp 228.

incluido en AMBER 11, utilizando una esfera de prueba de radio  $r = 1,4 \text{ \AA}$ . Los parámetros  $\delta = 0,0072 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  y  $b = 0 \text{ kcal/mol}$  [12].

Nuestros resultados se muestran en la Fig. 6.14. Podemos ver que la nueva molécula B produce (en valor absoluto) un  $\Delta G$  de formación del complejo mayor. Mientras que el valor medio para el complejo con PhiKan083 es  $\sim -29 \text{ kcal/mol}$ , la molécula B presenta un valor medio de  $\sim -35 \text{ kcal/mol}$ . Así, estos resultados apuntan a la viabilidad de considerar los conceptos de cooperativismo en el rediseño de compuestos guía. En la modificación de PhiKan083 propuesta (molécula B), la estructura de la proteína se ve fortalecida por reforzar o estabilizar el BHB que se ve afectado (generando un dehidrón) tras la mutación oncogénica Y220C, donde la protección suministrada por la molécula PhiKan083 es incompleta.

Un trabajo reciente del grupo de Alan Fersht, publicó el desarrollo de nuevas moléculas las cuales extienden la región ocupada por PhiKan083 en ambos extremos de la cavidad [41]. Precisamente, solo la extensión del compuesto líder (PhiKan083) mediante un grupo acetileno, sobre la región que hemos indicado anteriormente, incrementó la afinidad del compuesto por

su proteína *target* y así logró estabilizar aún más a la proteína con la mutación Y220C. Estas moléculas, llamadas PhiKan5174 y PhiKan5196, se observó que también inducen la apoptosis en células humanas cancerígenas afectadas por la mutación Y220C.

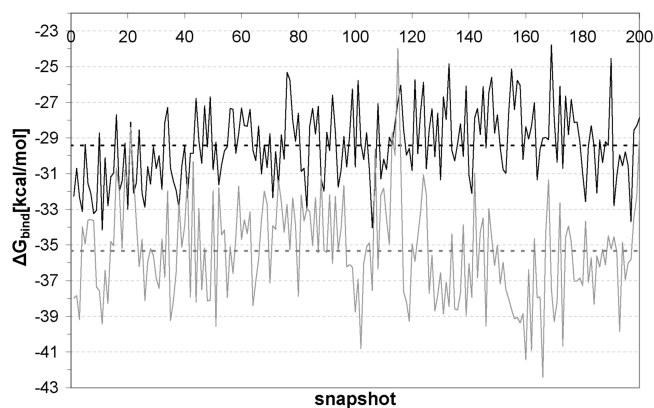


Figura 6.14: Evolución del  $\Delta G_{\text{binding}}$  durante una producción de dinámica molecular. Línea negra: T-p53C-Y220C/PhiKan083; línea gris: T-p53C-Y220C/segunda modificación de PhiKan083 (molécula B); línea punteada negra: valor medio para el complejo T-p53C-Y220C/PhiKan083; línea punteada gris: valor medio para el complejo T-p53C-Y220C/segunda modificación de PhiKan083 (molécula B).

### 6.3. Conclusiones

En este Capítulo verificamos la importancia de las interacciones de tres cuerpos o cooperativas en la interacción entre drogas y proteínas al verificar que las drogas disruptivas de interfases proteína-proteína tienden a mimetizar las interacciones cooperativas intermoleculares de la proteína que reemplazan. Sucesivamente vimos que las interacciones clásicas, como lo son los puentes de hidrógeno intermoleculares o puentes salinos intermoleculares, en general no son mimetizados por la molécula disruptiva. Esta comprensión es de gran relevancia para el diseño racional de drogas destinadas a bloquear interfases de proteínas relevantes para inhibir ciertas enfermedades. También, como vimos en este último caso particular de la mutación oncogénica de la proteína p53 y la molécula PhiKan083, estos conceptos son importantes para rediseñar racionalmente compuestos guía (*lead compounds*) y como filtro de selectividad para evitar efectos colaterales indeseables.



## Capítulo 7

# Agua en superficies hidrofóbicas modelo y superficies gráficas: la hidrofilicidad del grafeno

En los Capítulos 5 y 6 hemos estudiado procesos de unión a proteínas basados en un enfoque que codifica el comportamiento del agua nanoconfinada en un motivo estructural, los defectos de empaquetamiento dados por BHBs mal “arropados” o dehidrones. Dichos motivos estructurales muestran una propensión a la deshidratación local y, por lo tanto, representan regiones de alta hidrofobicidad local (pues como veremos luego, la labilidad del agua de hidratación constituye una medida de hidrofobicidad). Esta tendencia a la labilización del agua circundante que poseen dichos motivos (demostrada, a su vez, en el Capítulo 4) promueve la asociación de ligandos (otras proteínas o drogas) que deben reemplazar al agua de hidratación y que completan la desolvatación o secado del BHB.

Como ya hemos mencionado anteriormente, los principios que gobiernan la organización biológica a nivel molecular son los mismos que, en solución acuosa, actúan en el autoensamblado de estructuras moleculares para crear materiales sintéticos supramoleculares de dimensiones nanométricas [40, 63]. El solvente juega, por lo tanto, un papel fundamental en el diseño de nuevos materiales [40, 109, 55]. Sin embargo, si deseamos estudiar otros sistemas no proteicos,

tal como los procesos de ensamblado o autoensamble en ciencia de materiales en ambientes acuosos, nos encontramos ante contextos mucho más amplios donde la existencia de un motivo estructural prevalente no emerge claramente dada la heterogeneidad de sistemas de estudio. Por lo tanto, como primer paso, es necesario para este contexto desarrollar medidas eficaces de la hidrofobicidad local. Es de interés, así mismo, estudiar modelos simples de los cuales extraer principios generalizables. Además, en este campo los procesos de ensamble en ambientes acuosos se producen a través del establecimiento de interacciones no-covalentes dentro de entornos de nanoconfinamiento. Comprender como dichas interacciones no-covalentes dependen del entorno (principalmente del comportamiento local del solvente y, por ende, de la hidrofobicidad local) es un primer paso para intentar extraer elementos de diseño de impacto generalizado. Por lo tanto, el presente capítulo apunta a estudiar el comportamiento del agua interfacial y nanoconfinada y las condiciones de hidrofobicidad local en materiales y en procesos de ensamblado, con el objeto de inaugurar una nueva etapa en nuestras investigaciones, inclinándolas también hacia la ciencia de materiales (cuestión que, si bien excede el marco de esta Tesis, señala un camino de continuidad). Esperamos que los enfoques más generales a que apuntamos también sean de interés en contextos biológicos o de bioingeniería en que también operen otras interacciones no covalentes no contempladas en los capítulos anteriores o en que se necesite un grado de detalle o resolución mayor al que hemos utilizado anteriormente.

En este capítulo estudiamos diferentes parámetros como las fluctuaciones de densidad y tiempos de residencia del agua alrededor de una placa de grafeno, en el exterior de nanotubos de carbono de pared simple, en la superficie de monocapas autoensambladas de alcanos (SAMs) y en poros o cavidades de diferente tamaño talladas en la superficie de las SAMs. Mostramos que la curvatura convexa de las superficies gráficas no produce efectos significativos sobre la dinámica del agua de hidratación y así el grado de hidrofobicidad/hidrofilicidad de estas superficies no se ve mayormente afectado por la curvatura de las mismas [14]. Sin embargo, se observa una marcada dependencia en la hidrofobicidad con la curvatura de las superficies concavas de los poros tallados en las SAMs [14]. Demostramos que las fluctuaciones de densidad de agua alrededor de una placa de grafeno son prácticamente idénticas a las fluctuaciones sobre una superficie hidrofílica y muy diferentes de aquellas para una superficie hidrofóbica como las

SAMs. Finalmente, por medio de estudios dinámicos mostramos que dicha hidrofiliidad del grafeno se traduce en un comportamiento de interés práctico: previene la agregación de dos placas de grafeno.

## 7.1. Cuantificando la hidrofobicidad local mediante el cálculo de las fluctuaciones de densidad de agua

Entre las diferentes medidas estructurales, dinámicas y termodinámicas de hidrofobicidad que han sido propuestas [40, 70], una muy atractiva consiste en la cuantificación de las fluctuaciones de densidad de agua [70]. Se ha demostrado que la densidad superficial del agua no representa una buena medida de la hidrofobicidad de una superficie. Esto es esperable en términos del conocimiento usual de que el “agua aborrece el vacío” y así, las moléculas de agua tienden a hidratar tanto a las superficies polares como a las no polares. En consecuencia, los perfiles de densidad de agua normal al plano de la superficie presentan similares características para ambos tipos de sustratos. Sin embargo, a diferencia de las superficies hidrofílicas, donde las moléculas de agua están sujetas a interacciones atractivas significativas, en las superficies hidrofóbicas estas interacciones son muy débiles, en consecuencia las moléculas de agua de la capa de hidratación presentan menores tiempos de residencia y son más fáciles de remover. Por lo tanto, dichas capas de hidratación tendrán una dinámica superior a las de las superficies hidrofílicas [70], una mayor compresibilidad [40, 70] y mayores fluctuaciones de densidad [70]. En particular, se han caracterizado las fluctuaciones de densidad de agua en diferentes monocapas autoensambladas (SAMs) funcionalizadas, demostrando que aquellas superficies de naturaleza hidrofóbica presentan mayores fluctuaciones que las correspondientes a superficies hidrofílicas, proporcionando así una excelente medida cuantitativa de la hidrofobicidad [70]. La fluctuación en la densidad numérica de moléculas de agua ( $\sigma^2$ ) en un volumen de observación pequeño y fijo, dividida por el valor medio de moléculas (en dicho volumen) elevado al cuadrado ( $\langle N \rangle^2$ ) es aproximadamente igual a  $kT/2\mu^{ex}$ , donde  $\mu^{ex}$  es el exceso de energía libre involucrado en la formación de una cavidad de un determinado radio  $r$  (el radio del volumen de observación esférico) [70]. Así, un valor elevado de  $\sigma^2/\langle N \rangle^2$ , en un determinado volumen, indica un trabajo favorable

en la creación de una cavidad en dicho volumen, y por lo tanto una elevada hidrofobicidad. En consecuencia esta medida es de gran relevancia en el ámbito de procesos de autoensamblado de superficies no polares en agua, de modo de cuantificar la propensión a la deshidratación de dichas interfases. También en el caso de sitios de *binding* se puede explotar esta medida, ya que se espera que en dichos sitios los ligandos desplacen fácilmente el agua de hidratación.

## 7.2. Hidratación y geometría

Las propiedades de hidratación de un soluto no solo están definidas por la naturaleza química del soluto sino también por su geometría. Por ejemplo, un soluto no polar muy pequeño puede ser clatrado por el agua mientras que un soluto no polar de mayores dimensiones se agrega por medio del efecto hidrofóbico. Las moléculas de agua en el clatrato básicamente retienen la coordinación de cuatro puentes de hidrógeno (HBs) típica del agua *bulk*, mientras que en superficies no polares extendidas, la red de HBs se ve interrumpida y así las moléculas de agua en la superficie se encuentran vinculadas por menos de cuatro HBs. El proceso de agregación de solutos hidrofóbicos disminuye el tamaño total de superficie expuesta al agua de modo que cierto número de moléculas de agua superficiales (condicionadas a un arreglo geométrico determinado a los efectos de excluir al hidrófobo minimizando la pérdida de puentes de hidrógeno) se liberan hacia el seno del líquido (*bulk*) donde gozan de mayor libertad conformacional. Por lo tanto, el efecto hidrofóbico es de naturaleza entrópica. A su vez, se ha demostrado que cavidades no polares subnanométricas se mantienen secas, mientras que el agua sí penetra aquellas cavidades de dimensiones nanométricas [70, 23]. Este comportamiento se debe a la resistencia de las moléculas de agua a resignar la coordinación por puente de hidrógeno con otras moléculas de agua (solo las cavidades de tamaño nanométrico permiten la penetración de moléculas de agua reteniendo la coordinación típica en las superficies). Sin embargo, diferentes simulaciones han mostrado que el agua penetra en nanotubos de carbono de pared simple con diámetros de hasta 7 Å [130, 23]. A diferencia del comportamiento típico en superficies planas [22, 24, 17], donde las moléculas de agua resignan la mínima coordinación de HBs posible, dentro de estos nanotubos las moléculas de agua penetran con un arreglo uni-dimensional, en el cual cada molécula posee solo dos HBs con sus moléculas

de agua vecinas [130, 105, 70, 23]. Dado que la interacción entre moléculas de agua dentro del nanotubo con aquellas moléculas fuera del mismo es insignificante, la disminución de la resistencia típica del agua a la pérdida de coordinación solo puede atribuirse a las interacciones entre las moléculas de agua y las paredes del nanotubo, por lo tanto los nanotubos de carbono son a lo sumo menos hidrofóbicos de lo que anteriormente se esperaba. Esto es compatible con los resultados presentados en las Refs. [130, 105, 70], donde mediante simulaciones muestran que las moléculas de agua se desorben de nanotubos de carbono de diámetro subnanométrico cuando se reducen las interacciones con la pared del mismo. Por otro lado, trabajos de nuestro grupo de investigación han demostrado que tuneles de diametro similar tallados en las SAMs hidrofóbicas persisten secos dado a que estos sistemas carecen de la interacción significativa del agua con las paredes, la cual es necesaria para contrarestar el costo de 10 kcal/mol debido a la pérdida de un HB. Por lo tanto, la penetración del agua demanda un diametro de tunel lo suficientemente grande como para permitirle a las moléculas conservar un número significativo de HBs.

Debido a las diferencias en el comportamiento del agua de hidratación de las superficies gráficas con el de las superficies hidrofóbicas modelo como las SAMs, en las secciones y subsecciones siguientes estudiaremos las fluctuaciones de densidad y tiempos de residencia del agua en ambos tipos de sistemas, presentando en una primer subsección los resultados para las superficies gráficas y en una segunda para las SAMs. De esta manera podremos analizar como la geometría afecta a estas propiedades en ambos tipos de superficies por separado. Luego en la sección 7.3 compararemos estos resultados entre ambos tipos de superficies y a su vez mostraremos que dos placas de grafeno paralelas no tienden a excluir la capa intermedia de agua entre ellas, sino que se mantienen totalmente hidratadas y no se autoensamblan, mientras que dos superficies hidrofóbicas enfrentadas, como dos SAMs, se deshidratan y se autoensamblan.

Tanto para las simulaciones de las SAMs como para las superficies gráficas se utilizaron los campos de fuerza GAFF y ff99SB simultaneamente [84, 98] mediante la plataforma de simulación AMBER 10 [83]. En todas las simulaciones se utilizó el modelo de agua TIP3P [131]. Solo en algunos casos se utilizó el modelo de agua SPC/E [161] para corroborar si una parametrización diferente del mismo sistema influye o no en los resultados obtenidos. Esta cuestión

se discutirá más en detalle en la sección 7.3. La equilibración de los diferentes sistemas y el modelado de la placa de grafeno y de los nanotubos de carbono se llevó a cabo tal cual ya fue descrito para una placa de grafeno en el Capítulo 3, sección 3.2, subsección 3.2.1.

### 7.2.1. Agua en superficies gráficas: grafeno y nanotubos

Resultados previos en el cálculo del índice de estructura local  $I$  [145, 125] (ver definición de  $I$  en subsección 2.2.1) obtenidos por nuestro grupo de investigación, muestran que las moléculas de agua cercanas a la superficie del grafeno poseen una estructura similar a las moléculas de agua cercanas a la pared exterior de nanotubos de gran radio. Sin embargo, cuando se consideran radios dentro del rango subnanométrico se observa una significativa pérdida en la estructura del agua de hidratación del nanotubo [24]. Por lo tanto estamos interesados en determinar si la desestructuración geométrica del agua de hidratación, inducida por la curvatura de la superficie, está acoplada a un cambio en el grado de hidrofobicidad/hidrofilicidad de las superficies gráficas. La Fig. 7.1 muestra la distribución de probabilidad ( $p(N)$ ) de observar  $N$  moléculas de agua dentro de un pequeño volumen de observación de radio  $r = 3,3 \text{ \AA}$  (similar a una molécula de metano) tangente a las diferentes superficies gráficas estudiadas. Cuanto más ancha sea la distribución de  $p(N)$  mayores serán las fluctuaciones de densidad de agua en dicho volumen de observación. Podemos ver que las fluctuaciones en la densidad del agua no dependen apreciablemente de la curvatura de las superficies gráficas, y por lo tanto el grado de hidrofobicidad/hidrofilicidad es prácticamente constante con el cambio de curvatura de estas superficies [14].

Otro parámetro que refleja el grado de hidrofobicidad/hidrofilicidad son los tiempos de residencia de las moléculas agua. A medida que la hidrofobicidad de la superficie aumenta se espera que los tiempos de residencia de las moléculas de agua de hidratación decrezcan (mientras que serán mayores para el caso de superficies hidrofílicas). En consecuencia, calculamos los tiempos de residencia de las moléculas agua dentro de volúmenes esféricos tangentes a las superficies por medio de la función de autocorrelación. Dicha función se define como:

$$C(t) = \langle R(0)R(t) \rangle / \langle R(0) \rangle^2$$

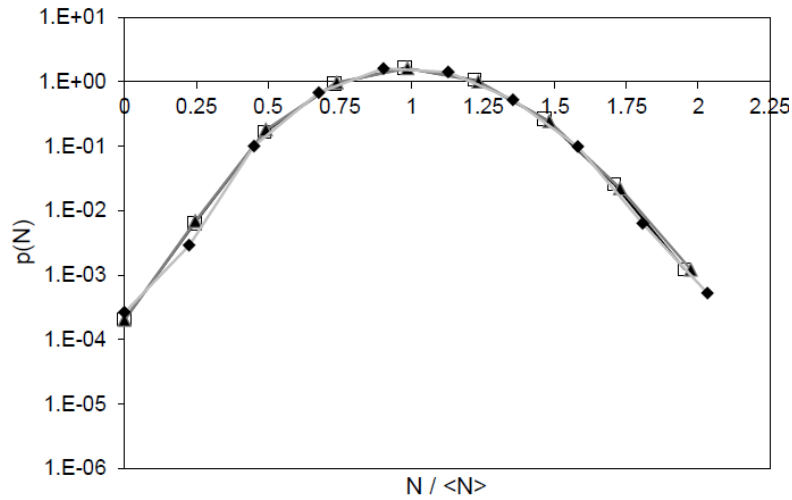


Figura 7.1: Distribución de probabilidad  $p(N)$  para una placa de grafeno (cuadrados) y para los nanotubos de radio  $r = 2 \text{ \AA}$  (triángulos) y  $r = 28 \text{ \AA}$  (diamantes negros).

donde  $R(t)$  es una función de residencia definida por:

$$R(t) = \sum_{i=1}^n R_i(t)$$

El índice  $i$  corre sobre las  $n$  moléculas de agua que al tiempo  $t = 0$  se encuentran dentro del volumen ( $V$ ) de observación esférico, tangente a la superficie y de radio  $r = 3 \text{ \AA}$ . Por lo tanto  $R_i(t) = 1$  cuando la molécula  $i$  al tiempo  $t$  se encuentra todavía dentro del volumen  $V$ , de lo contrario  $R_i(t) = 0$ .

La Fig. 7.2 muestra la función de autocorrelación para esferas de radio  $r = 3 \text{ \AA}$  tangentes a la superficie del grafeno y de los diferentes nanotubos. Nuevamente se puede apreciar que la curvatura de las superficies gráficas no afecta significativamente al grado de hidrofobicidad/hidrofilicidad de las mismas [14], y así se refuerza el concepto de que las fluctuaciones de densidad de agua representan una buena medida de la hidrofobicidad/hidrofilicidad de las superficies.

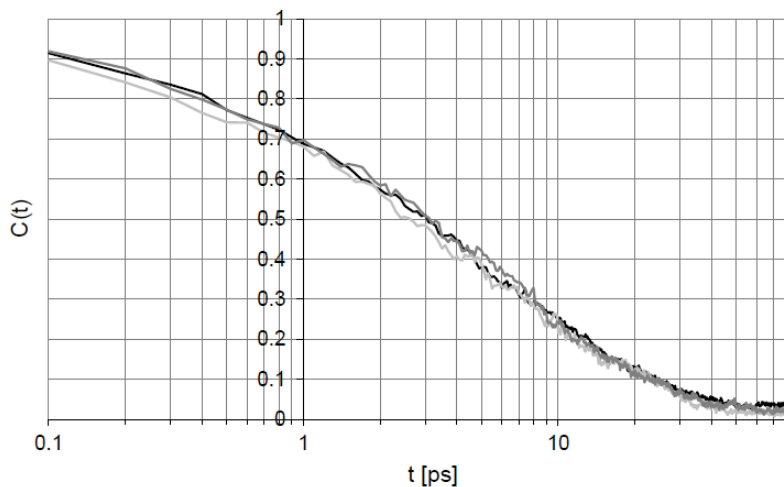


Figura 7.2: Función de autocorrelación,  $C(t)$ , para el grafeno (curva negra) y para los nanotubos de radio  $r = 2 \text{ \AA}$  (curva gris claro) y  $r = 28 \text{ \AA}$  (curva gris oscuro).

### 7.2.2. Agua en las monocapas autoensambladas de alcanos (SAMs) y el llenado de cavidades

Para el modelado de las cavidades hidrofóbicas construimos monocapas de n-decano de 81 cadenas ( $9 \times 9$ ) y de 144 cadenas ( $12 \times 12$ ), alineadas de forma paralela en un arreglo cúbico. Este arreglo imita la estructura de las monocapas de ácido esteárico en una interfase con agua, pero reemplazando los grupos ácido ( $-\text{COOH}$ ) por un átomo de hidrógeno, de forma que las cadenas terminen con un grupo metilo ( $-\text{CH}_3$ ), es decir transformando el ácido esteárico en n-decano. En estos casos, las monocapas o SAMs constituyen solo un modelo de un arreglo hidrofóbico, el cual utilizaremos para enfocarnos en el comportamiento del agua que las rodea. La distancia de separación original entre las cadenas es de  $4,53 \text{ \AA}$ , la típica distancia en las monocapas de ácidos grasos. Sin embargo, luego de la equilibración del sistema esta distancia se reduce a  $4,20 \text{ \AA}$ .

De aquí en adelante a la monocapa sin ninguna cavidad la llamaremos monocapa “perfecta”. Para generar las cavidades se tallaron agujeros cuadrados de diferente tamaño en el centro de la monocapa cortando el número correspondiente de cadenas y luego saturando con hidrógeno a los carbonos en el fondo de la cavidad, de manera que todas las cadenas terminen con un grupo metilo. Para las cavidades utilizamos la siguiente nomenclatura:  $A \times B$ , donde  $A$  corresponde al



número de grupos metilo ( $-\text{CH}_3$ ) de la superficie que son eliminados conformando un cuadrado en el centro de la SAM; mientras que  $B$  equivale al número de átomos de carbono eliminados en profundidad (la profundidad de la cavidad estará determinada entonces por  $B \times 1,2 \text{ \AA}$ , donde  $1,2 \text{ \AA}$  equivale a la longitud de un monómero ( $-\text{CH}_2-$ ). Estudiamos una monocapa perfecta junto con las siguientes cavidades:  $1 \times 5$ ,  $4 \times 5$ ,  $9 \times 5$ ,  $16 \times 5$ ,  $25 \times 5$ ,  $36 \times 5$ ,  $49 \times 5$  y  $64 \times 5$ . Cada una de ellas fueron talladas en el centro de una de las caras superficiales de la monocapa. De esta manera nos aseguramos que se encuentran lo suficientemente alejadas de los bordes. Las cavidades de  $1 \times 5$ ,  $4 \times 5$ ,  $16 \times 5$  y  $25 \times 5$  se tallaron sobre la monocapa de 81 cadenas ( $9 \times 9$  cadenas), mientras que las cavidades de mayor tamaño fueron talladas sobre la monocapa de 144 cadenas ( $12 \times 12$  cadenas). La Tabla 7.1 muestra las dimensiones de cada una de las cavidades talladas.

$A \times B$	Ancho ( $\text{\AA}$ )	Profundidad ( $\text{\AA}$ )
$1 \times 5$	8,5	6,4
$4 \times 5$	12,7	6,4
$9 \times 5$	16,5	6,4
$16 \times 5$	21,3	6,4
$25 \times 5$	25,1	6,4
$36 \times 5$	29,5	6,4
$49 \times 5$	33,8	6,4
$64 \times 5$	38,0	6,4

Cuadro 7.1: Dimensiones de las diferentes cavidades talladas en las SAMs. El ancho de las mismas se mide entre átomos de carbono que se encuentran a la misma altura en paredes opuestas de la cavidad. Todas las cavidades poseen la misma profundidad, ya que en todos los casos se cortaron las cadenas correspondientes en cinco unidades o átomos de carbono.

Como en el caso de las superficies gráficas, estudiamos las fluctuaciones de densidad de agua en un volumen de observación esférico y tangente a la superficie. Por lo tanto, para el caso de las cavidades, dicha esfera de observación se localiza centrada dentro de la misma, tangente a la línea definida por los átomos de hidrógeno de los grupos metilo superficiales de las cadenas alquílicas que enmarcan las cavidades (ver Fig. 7.3 para el caso de la cavidad de  $1 \times 5$  y la Fig. 7.4 para el caso de la cavidad de  $25 \times 5$ ).

La Fig. 7.5 a) muestra el valor medio,  $\langle N \rangle$ , de moléculas de agua dentro del volumen de

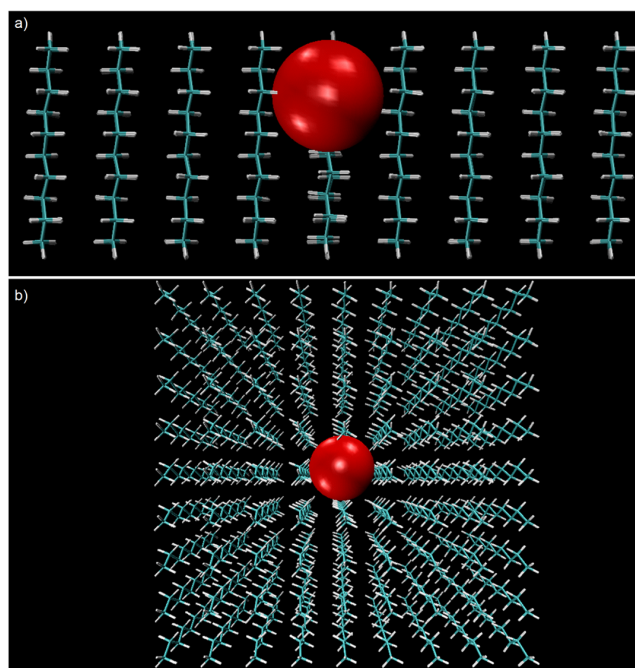


Figura 7.3: Ilustración de la SAM con una cavidad de  $1 \times 5$ . La nomenclatura  $A \times B$  significa que un número de cadenas  $A$ , conformando un cuadrado en el centro de la monocapa, fueron acortadas en  $B$  unidades. En este caso una sola cadena alquílica (la cadena central) fue acortada en cinco unidades para crear la cavidad. a) Vista lateral de la monocapa cortada transversalmente en la mitad de manera de poder observar mejor la cavidad. La esfera roja indica el volumen de observación utilizado de radio  $r = 3,3 \text{ \AA}$ . b) Vista superior de la monocapa.

observación en las cavidades, en función del ancho de la cavidad. La línea segmentada negra representa dicho valor medio para una monocapa perfecta. Vemos que a medida que el ancho de la cavidad aumenta, el valor de  $\langle N \rangle$  tiende al valor medio para una superficie sin cavidad. La Fig. 7.5 b) muestra las fluctuaciones de densidad de agua normalizadas,  $\sigma^2 / \langle N \rangle^2$ , en las esferas de observación (ver Fig. 7.3 y Fig. 7.4) para todas las cavidades talladas. Podemos ver que el grado de hidrofobicidad depende fuertemente de la curvatura en las SAMs. Las cavidades son más hidrofóbicas que la monocapa perfecta, aunque el rol de la geometría se vuelve más conspicuo a medida que el tamaño de la cavidad (el ancho de la cavidad,  $L$ ) se aproxima el régimen subnanométrico. Las cavidades subnanométricas no se llenan de agua debido a la reticencia de las moléculas de agua a romper su coordinación de puente de hidrógeno para penetrar en la cavidad subnanométrica. Cavidades de mayor tamaño (hasta aproximadamente  $L = 25 \text{ \AA}$ ),

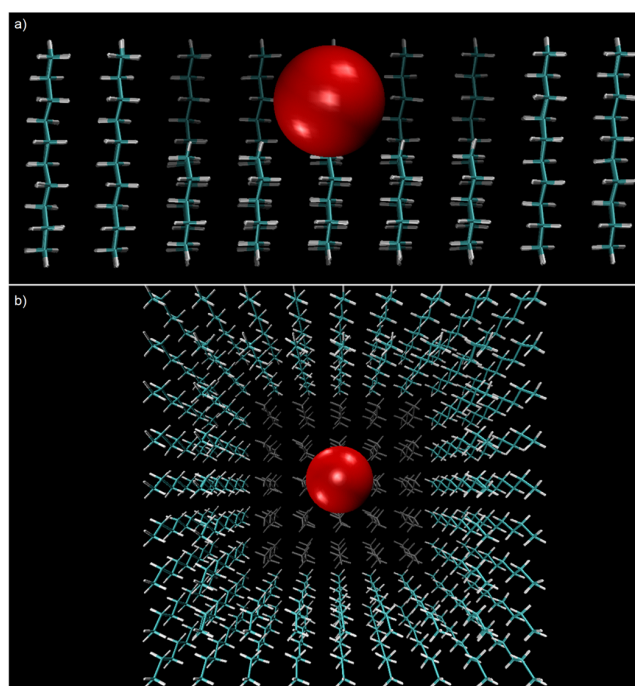


Figura 7.4: Ídem Fig. 7.3 pero para la cavidad de 25 x 5. En este caso las 25 cadenas alquílicas centrales de la monocapa fueron acortadas en cinco unidades, como lo indica la nomenclatura  $AxB$ . a) Vista lateral. b) Vista superior.

donde las moléculas de agua pueden penetrar sin perder su coordinación, son sin embargo más hidrofóbicas que la monocapa perfecta [14]. Por lo tanto la hidrofobicidad es claramente dependiente de la curvatura para superficies concavas en las SAMs. Este resultado es interesante, por ejemplo, en el contexto del *binding* de proteínas, dado que se espera que los sitios de *binding* (muchas veces constituídos por “bolsillos”) contengan moléculas de agua fácilmente removibles, las cuales deberían ser desplazadas por el ligando durante la asociación [119, 72, 16].

### 7.3. El comportamiento hidrofílico del grafeno y de los materiales gráficos

Diversos estudios del ángulo de contacto de agua sobre superficies gráficas han mostrado que estas superficies presentan un comportamiento hidrofóbico, con valores que rondan los  $90^\circ$  [37, 52, 60, 168, 170, 73, 120, 107, 38]. Sin embargo, un estudio reciente señala la posibilidad de

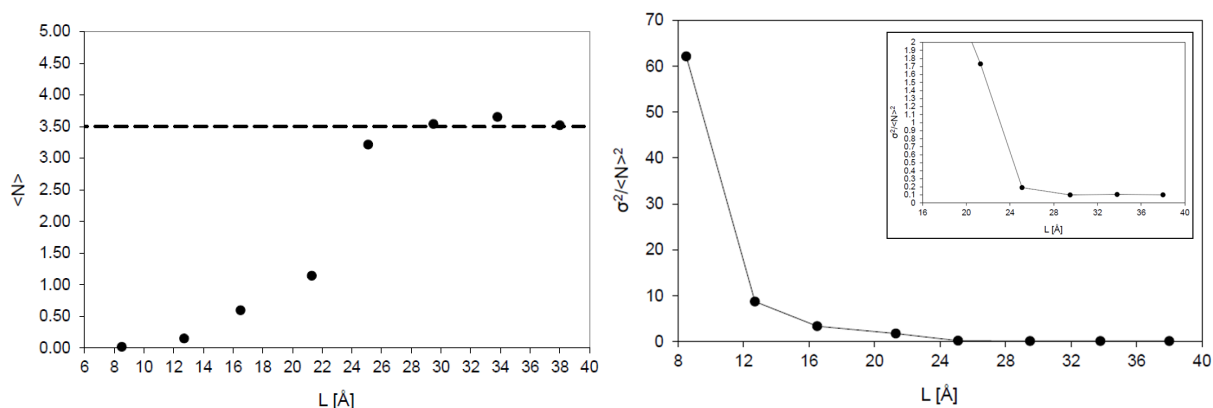


Figura 7.5: a) Valor medio,  $\langle N \rangle$ , del número de moléculas agua dentro del volumen de observación para las cavidades en función del ancho,  $L$ , de las mismas. La línea negra segmentada representa el valor medio para una monocapa perfecta. b) Fluctuaciones de densidad de agua normalizadas,  $\sigma^2 / \langle N \rangle^2$ , en función del ancho,  $L$ , de la cavidad. En el inset se puede observar que los valores de  $\sigma^2 / \langle N \rangle^2$  tienden a un valor constante (el valor para la monocapa “perfecta”) a medida que  $L$  aumenta.

que las medidas experimentales del ángulo de gota realizadas sobre superficies gráficas estén siendo afectadas por la susceptibilidad de estas superficies a la contaminación directa desde el aire [37]. Por lo tanto tomando los recaudos necesarios, en la Ref. [37] han obtenido un ángulo de contacto cercano a los  $64^\circ$  para grafito pirolítico altamente orientado (*highly ordered pyrolytic graphite*, HOPG), y valores un poco menores para grafeno soportado tanto en níquel como en cobre (sin embargo en estos casos la medida podría estar reflejando el efecto de transparencia parcial en el mojado del grafeno, y así estar afectando al valor resultante del ángulo medido [42, 43]). Estos valores obtenidos aumentan con el envejecimiento de la superficie, reflejando así los efectos de la contaminación sobre esta medida [37].

En esta sección comparamos los resultados obtenidos para las fluctuaciones de densidad de agua alrededor de la placa de grafeno, de la SAM (de  $9 \times 9$  cadenas alquílicas de n-decano) hidrofóbica perfecta (sin cavidad) y además incluimos los resultados para una SAM de las mismas dimensiones pero funcionalizada con grupos hidrofílicos -OH en la superficie. Es decir que en este último caso, reemplazamos un átomo de hidrógeno de los grupos  $-\text{CH}_3$  localizados en la superficie, por un -OH, de manera de obtener el grupo  $-\text{CH}_2\text{OH}$ .

En el caso particular de la placa de grafeno el sistema se modeló de acuerdo a las tres variantes siguientes:

- Utilizamos los parámetros de Lennard-Jones,  $\sigma_{CO} = 3,275 \text{ \AA}$  y  $\epsilon_{CO} = 0,114 \text{ kcal/mol}$ , para modelar las interacciones agua-carbono, resultantes de combinar el campo de fuerza GAFF con el modelo de agua TIP3P.
- Debilitamos las interacciones de atracción de van der Waals agua-carbono en un factor de aproximadamente igual a dos [130] ( $\sigma'_{CO} = 3,41 \text{ \AA}$  y  $\epsilon'_{CO} = 0,065 \text{ kcal/mol}$ ) utilizando el modelo de agua TIP3P y el campo de fuerza GAFF.
- Utilizamos el modelo de agua SPC/E junto con el campo de fuerza GAFF y los mismos parámetros de Lennard-Jones que en el primer ítem ( $\sigma_{CO} = 3,275 \text{ \AA}$  y  $\epsilon_{CO} = 0,114 \text{ kcal/mol}$ ) [120].

Diferentes modelos de agua poseen diferente energía potencial de interacción agua-agua en el *bulk*,  $E_{\text{pot}}$  ( $\text{kJ mol}^{-1}$ ), y para una interacción agua-carbono constante, cambiar el modelo de agua equivale a cambiar el balance de energía entre las interacciones agua-carbono y agua-agua [120]. Entre los modelos TIP3P y SPC/E, el último posee la menor energía potencial de interacción en el *bulk* ( $-47,2 \text{ kJ/mol}$  para el modelo SPC/E contra  $-39,8 \text{ kJ/mol}$  para el modelo TIP3P) [120].

El valor obtenido para el ángulo de contacto de agua sobre grafito, mediante simulaciones por computadora, utilizando el modelo de agua TIP3P y los parámetros de Lennard-Jones  $\sigma_{CO} = 3,275 \text{ \AA}$  y  $\epsilon_{CO} = 0,114 \text{ kcal/mol}$ , fue  $48,0^\circ$  [120]; mientras que en el caso del modelo de agua SPC/E, con los mismos parámetros de interacción agua-carbono, fue  $65,4^\circ$  [120]. Este último valor está de acuerdo con los recientes valores experimentales obtenidos en la Ref. [37]. Por esta razón en esta sección utilizamos también el modelo de agua SPC/E para estudiar las fluctuaciones de densidad de agua sobre grafeno y así verificar la robustez de nuestros resultados. En el caso de las SAMs siempre utilizamos el modelo de agua TIP3P.

La Fig. 7.6 muestra las distribuciones de  $p(N)$  para las diferentes superficies estudiadas. Podemos ver que las fluctuaciones de densidad de agua en el grafeno, con los parámetros de

Lennard-Jones  $\sigma_{CO} = 3,275 \text{ \AA}$  y  $\epsilon_{CO} = 0,114 \text{ kcal/mol}$  resultantes de la combinación del campo de fuerza GAFF con el modelo de agua TIP3P, son casi idénticas a aquellas para la SAM hidrofílica (la SAM funcionalizada con grupos -OH en la superficie); y muy diferentes de aquellas para la SAM hidrofóbica (la SAM con grupos metilo,  $-\text{CH}_3$ , en la superficie). Adicionalmente, podemos notar que si disminuimos las interacciones agua-carbono en el sistema grafeno-agua (TIP3P), utilizando los parámetros  $\sigma'_{CO} = 3,41 \text{ \AA}$  y  $\epsilon'_{CO} = 0,065 \text{ kcal/mol}$ , las fluctuaciones de densidad de agua aumentan (la distribución se hace más ancha), lo cual habla de un comportamiento más hidrofóbico para el grafeno modificado. Estos resultados son compatibles con los encontrados en las simulaciones de las Refs. [130, 105, 70], donde el agua se desorbe de nanotubos de carbono de radio pequeño si las interacciones entre el agua y la pared del nanotubo se reducen. Finalmente, podemos ver que si utilizamos el modelo de agua SPC/E para solvatar a la placa de grafeno, y utilizamos los parámetros de Lennard-Jones  $\sigma_{CO} = 3,275 \text{ \AA}$  y  $\epsilon_{CO} = 0,114 \text{ kcal/mol}$ , la distribución sigue siendo similar a la de la superficie hidrofílica (SAM con grupos -OH) y así podemos asegurar la robustez de nuestros resultados frente a una parametrización diferente del mismo sistema. En este último caso utilizamos los mismos parámetros y modelo de agua que los autores usaron en la Ref. [120] para reproducir un valor teórico de ángulo de contacto de gota de  $65,4^\circ$ . Este ángulo es muy similar al valor medido experimentalmente en la Ref. [37].

Por último estudiamos un sistema constituido por dos placas de grafeno paralelas en un medio acuoso separadas inicialmente una distancia  $d$  (ver Fig. 7.7 a) y b)) y otro sistema similar de dos SAMs enfrentadas, cada una de 25 cadenas alquílicas (5 x 5) de n-nonano alineadas de forma paralela en un arreglo cúbico (ver Fig. 7.8 a), b), c) y d)). La temperatura de simulación en ambos casos es  $T = 300 \text{ K}$ .

El sistema de dos placas paralelas de grafeno se modeló de tres maneras diferentes, de acuerdo a los tres ítems descritos anteriormente al comienzo de la sección. En este caso en particular las superficies gráficas consisten en un par de placas de grafeno paralelas, cada una de ellas de dimensiones  $A = 20 \text{ \AA} \times 20 \text{ \AA}$  con terminación en átomos de hidrógeno. Las superficies fueron centradas en el medio una caja cúbica ortogonal de moléculas de agua, paralelas al plano  $XY$  y separadas una de otra inicialmente por  $d_1 = 5 \text{ \AA}$  y  $d_2 = 7 \text{ \AA}$  (se utilizaron dos distancias iniciales

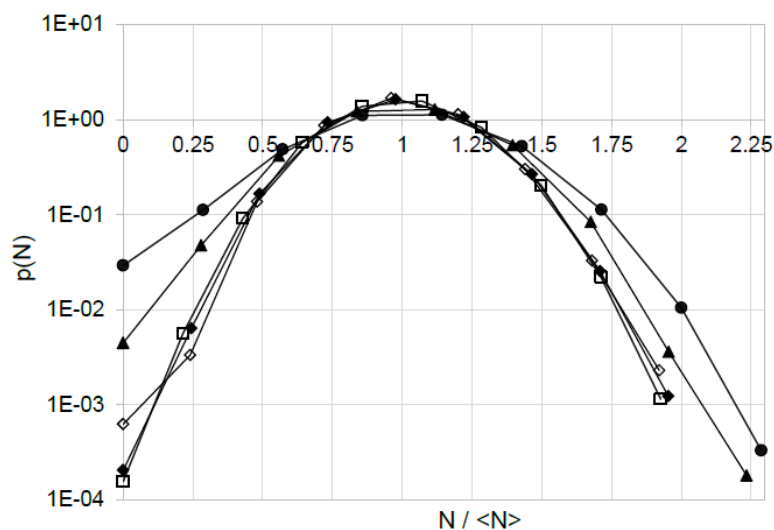


Figura 7.6: Distribuciones de probabilidad de observar  $N$  moléculas de agua,  $p(N)$ , dentro de un pequeño volumen de observación esférico de radio  $r = 3,3 \text{ \AA}$  tangente a las diferentes superficies estudiadas. Círculos negros: SAM con terminación en grupos  $-\text{CH}_3$  (campo de fuerza GAFF y modelo de agua TIP3P). Cuadrados: SAM con terminación en grupos  $-\text{CH}_2\text{OH}$  (campo de fuerza GAFF y modelo de agua TIP3P). Rombos negros: Grafeno con los parámetros de Lennard-Jones  $\sigma_{CO} = 3,275 \text{ \AA}$  y  $\epsilon_{CO} = 0,114 \text{ kcal/mol}$  resultantes de combinar el campo de fuerza GAFF con el modelo de agua TIP3P. Rombos vacíos: Grafeno con los parámetros de Lennard-Jones  $\sigma_{CO} = 3,275 \text{ \AA}$  y  $\epsilon_{CO} = 0,114 \text{ kcal/mol}$  (campo de fuerza GAFF y modelo de agua SPC/E). Triángulos: Grafeno con los parámetros de Lennard-Jones  $\sigma'_{CO} = 3,41 \text{ \AA}$  y  $\epsilon'_{CO} = 0,065 \text{ kcal/mol}$  (campo de fuerza GAFF y modelo de agua TIP3P).

diferentes  $d_1$  y  $d_2$ , ver Fig. 7.7 a) y b)). Tanto en el proceso de minimización como de equilibración del sistema fijamos las coordenadas  $XYZ$  de los átomos de carbono de ambas placas de grafeno, con el fin de conservar la separación inicial entre ambas superficies. Luego, durante la simulación de producción en el ensamble  $NpT$ , liberamos ambas placas para comprobar si excluyen la capa de agua intermedia o si se mantienen completamente hidratadas.

En el caso de las SAMs enfrentadas los protocolos de minimización y equilibración son los mismos que para las placas de grafeno, es decir que se fijaron las coordenadas  $XYZ$  de los átomos de ambas SAMs para conservar la distancia inicial entre ambas superficies. Luego, durante la simulación de producción solo fijamos las distancias entre átomos de carbono adyacentes (C9, C13, C17, ver Fig. 7.8 c)) dentro de una misma SAM mediante un potencial de pozo cuadrado

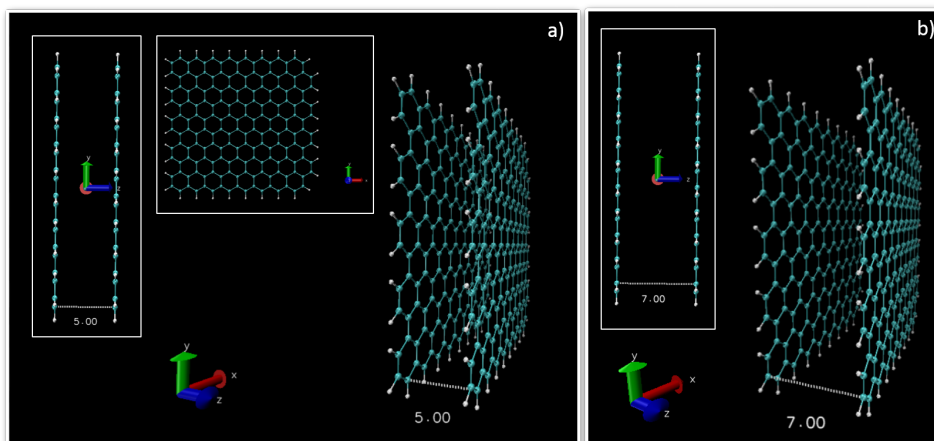


Figura 7.7: Dos placas de grafeno paralelas. Las distancias iniciales son: a)  $d_1 = 5 \text{ \AA}$  y b)  $d_2 = 7 \text{ \AA}$ .

con lados parabólicos. Los límites, inferior y superior, de este potencial son  $r_2 = 4,33 \text{ \AA}$  y  $r_3 = 4,53 \text{ \AA}$  respectivamente cuyas constantes de fuerza en cada extremo son  $rk_2 = 20 \text{ kcal/mol}$  y  $rk_3 = 20 \text{ kcal/mol}$  [13]. De esta manera impedimos que las cadenas de n-nonano colapsen dentro de una misma SAM. Aquí también partimos de dos distancias iniciales diferentes entre ambas superficies,  $d_{1-hh} = 7 \text{ \AA}$  y  $d_{2-cc} = 12 \text{ \AA}$  (los sufijos *hh* y *cc* significan que en el primer caso la distancia entre ambas superficies se mide entre los átomos de hidrógeno, mientras que en el segundo caso entre los átomos de carbono, ver Fig. 7.8 a) y b)). Las superficies fueron centradas en el medio de una caja cúbica ortogonal de agua y en este caso solo se utilizó el modelo de agua TIP3P.

En la Fig. 7.9 a) y b) podemos ver las dos configuraciones de partida utilizadas para las dos placas de grafeno paralelas ( $d_1 = 5 \text{ \AA}$ ,  $d_2 = 7 \text{ \AA}$ ). Utilizamos en primer lugar el modelo de agua TIP3P y los parámetros de Lennard-Jones  $\sigma_{CO} = 3,275 \text{ \AA}$  y  $\epsilon_{CO} = 0,114 \text{ kcal/mol}$  para las interacciones agua-carbono. En un trabajo anterior de nuestro grupo de investigación [23], mostramos que todos los nanotubos de carbono con un diámetro mayor que  $7 \text{ \AA}$  (distancia carbono-carbono) se llenan de agua, mientras que nanotubos con un diámetro menor a los  $6,2 \text{ \AA}$  persisten vacíos. Por lo tanto, en el primer caso donde la distancia inicial entre ambas placas es  $d_1 = 5 \text{ \AA}$ , como esperábamos, el volumen entre ambas placas de grafeno se encuentra seco luego del proceso de minimización y equilibración (ver Fig. 7.9 a)). Sucesivamente, durante la



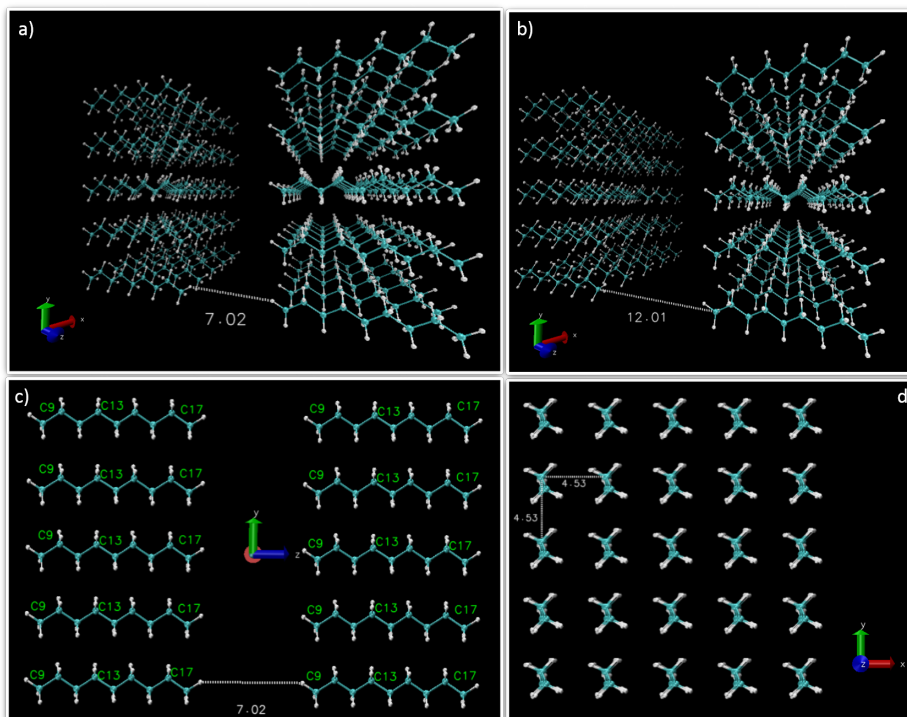


Figura 7.8: a) y b) Dos SAMs hidrofóbicas enfrentadas. Las distancias iniciales son:  $d_{1-hh} = 7 \text{ \AA}$  y  $d_{2-cc} = 12 \text{ \AA}$ , respectivamente; c) vista lateral del sistema a), se fijaron las distancias relativas entre los carbonos adyacentes C9, C13 y C17 dentro de una misma SAM mediante un potencial de pozo cuadrado con lados parabólicos. d) Vista superior del sistema a) o b), podemos ver el arreglo cúbico (5 x 5) de las cadenas de n-nonano. La separación inicial entre las cadenas de n-nonano dentro de una misma SAM es 4,53 Å. Luego de los procesos de minimización y equilibración la distancia entre la cadenas alquílicas dentro de una misma SAM se reduce a 4,33 Å.

dinámica de producción, liberamos ambas placas de grafeno y las mismas se autoensamblan durante los primeros 5 ps. Sin embargo, la situación es muy diferente para la condición inicial  $d_2$ . Luego de la minimización y equilibración del sistema (la distancia inicial entre ambas placas es  $d_2 = 7 \text{ \AA}$ ) podemos ver una capa de moléculas de agua entre ambas placas (ver Fig. 7.9 b)). Este resultado es esperable, dado que sabemos por los perfiles de densidad de agua alrededor de una placa de grafeno, que la primera capa de hidratación comienza alrededor de los 2,5 Å - 2,7 Å de la superficie (distancia medida desde el centro de los átomos de carbono al centro de los átomos de oxígeno de las moléculas de agua) [22, 24, 17]. Luego cuando liberamos las placas en

la dinámica de producción, las mismas no excluyen la capa intermedia de agua entre ellas, sino que persisten completamente hidratadas durante 3 ns de simulación. La Fig. 7.10 a) muestra la distancia entre los centros de ambas placas durante los 3 ns de simulación.

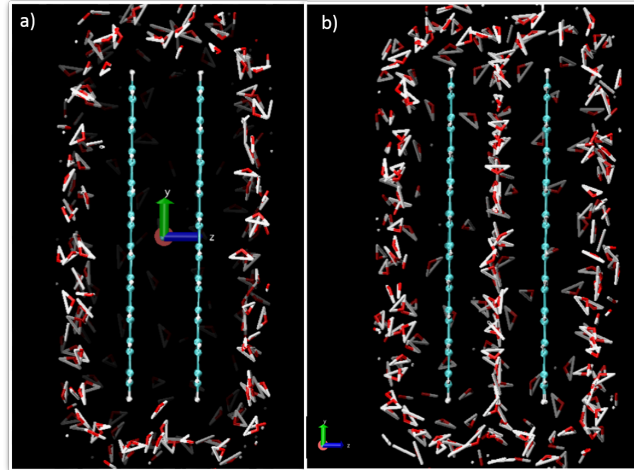


Figura 7.9: Configuración final luego de la minimización y equilibración de los sistemas grafeno-agua a)  $d_1 = 5 \text{ \AA}$  y b)  $d_2 = 7 \text{ \AA}$ .

También realizamos el mismo experimento solvatando las placas de grafeno inicialmente a la distancia  $d_2 = 7 \text{ \AA}$ , pero con el modelo de agua SPC/E. Utilizamos los mismos parámetros de Lennard-Jones  $\sigma_{CO} = 3,275 \text{ \AA}$  y  $\epsilon_{CO} = 0,114 \text{ kcal/mol}$ . El resultado general en este caso es el mismo que con el modelo de agua TIP3P: las dos placas de grafeno nunca colapsan durante los 3 ns de simulación. La Fig. 7.10 b) muestra la distancia entre los centros de ambas placas en función del tiempo de simulación para esta última parametrización del sistema.

Por último modelamos el sistema de placas de grafeno paralelas separadas inicialmente una distancia  $d_2$  inmersas en una caja de agua TIP3P, pero con los parámetros de Lennard-Jones  $\sigma'_{CO} = 3,41 \text{ \AA}$  y  $\epsilon'_{CO} = 0,065 \text{ kcal/mol}$ . Luego del proceso de minimización y equilibración el volumen entre ambas placas permanece seco (durante la minimización y la equilibración las coordenadas  $XYZ$  de los átomos de carbono de ambas placas son fijas, por lo tanto la distancia entre ambas es  $d_2$ ). A continuación, cuando liberamos ambas placas en el ensamble  $NpT$ , las mismas colapsan y se autoensamblan durante los primeros 4 ps de la simulación. Recordemos que en la Ref. [130], los autores mostraban que una pequeña reducción en los parámetros de

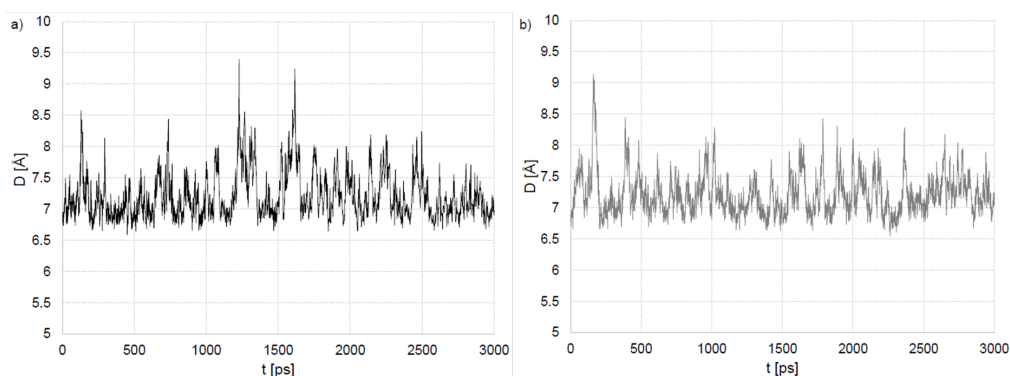


Figura 7.10: Distancia entre los centros de las dos placas de grafeno durante el tiempo completo de simulación. a) y b) Los parámetros de Lennard-Jones son  $\sigma_{CO} = 3,275 \text{ \AA}$  y  $\epsilon_{CO} = 0,114 \text{ kcal/mol}$ . a) El modelo de agua es TIP3P. b) El modelo de agua es SPC/E.

atracción entre las paredes del nanotubo de carbono y el oxígeno de las moléculas de agua (TIP3P) afectaba dramáticamente a la hidratación del poro. Por lo tanto queda claro que la interacción agua-carbono es un factor determinante en el comportamiento del agua dentro de nanotubos de carbono e incluso entre dos placas de grafeno. Así nuestros resultados son robustos y a la vez consistentes con las observaciones hechas por otros autores.

Finalmente comparamos la situación del grafeno con la de las SAMs hidrofóbicas. En trabajos previos de nuestro grupo de investigación mostramos que el mínimo tamaño de tunel, tallado en una monocapa hidrofóbica de alcanos, que puede ser llenado de agua tiene aproximadamente el doble de ancho que el mínimo diámetro de un nanotubo de carbono que permite la penetración de agua [23]. Por lo tanto, como esperábamos, el primer sistema de SAMs enfrentadas, donde  $d_{1-hh} = 7 \text{ \AA}$  (ver Fig. 7.8 a)), se comporta de manera muy diferente a lo mostrado anteriormente para dos placas de grafeno paralelas inicialmente a una distancia  $d_2 = 7 \text{ \AA}$ . Luego de los procesos de minimización y equilibración no se encontraron moléculas de agua entre ambas superficies, y finalmente cuando liberamos ambas monocapas (solo restringiendo la distancia entre los átomos C9, C13 y C17 adyacentes de una misma SAM para evitar que la estructura colapse) ambas se autoensamblan durante los primeros 6 ps de simulación. Cuando consideramos el segundo sistema, con una separación inicial  $d_{2-cc} = 12 \text{ \AA}$ , el volumen entre ambas SAMs se llena de moléculas de agua durante la minimización y equilibración del sistema. Luego, durante la dinámica de producción la distancia entre las dos superficies decrece con el tiempo Fig. 7.11

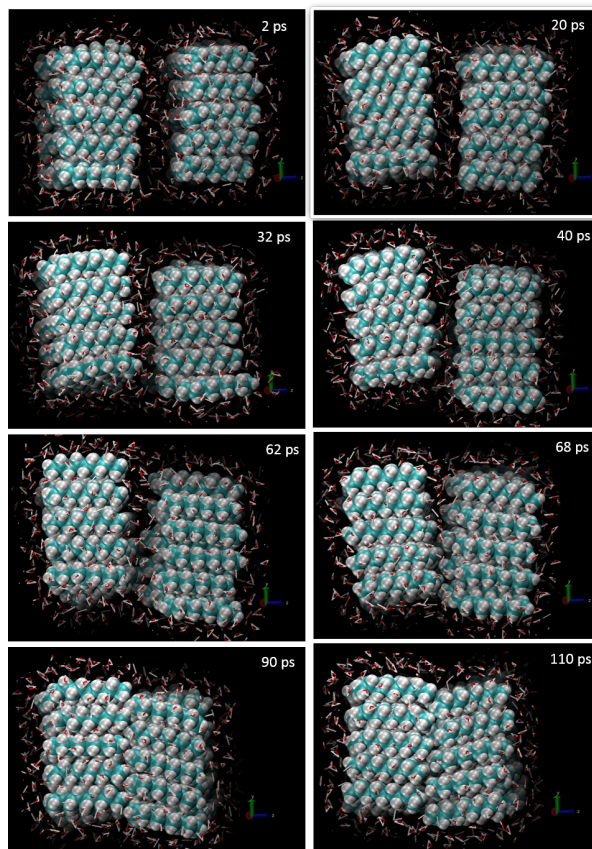


Figura 7.11: Proceso de autoensamblado entre las dos superficies hidrofóbicas inicialmente a una distancia  $d_{2-cc} = 12 \text{ \AA}$ .

y las dos SAMs colapsan durante aproximadamente los 100 ps de simulación.

## 7.4. Conclusiones

En este Capítulo estudiamos las fluctuaciones de densidad y tiempos de residencia del agua alrededor de una placa de grafeno, en el exterior de nanotubos de carbono de pared simple, en la superficie de monocapas de alcanos (SAMs) y en poros o cavidades de diferente tamaño talladas en la superficie de las SAMs. Mostramos que el grado de hidrofobicidad/hidrofilicidad de las superficies gráficas es prácticamente independiente de la curvatura convexa de las mismas. Sin embargo en el caso de las SAMs, a medida que la curvatura se incrementa, y así el tamaño de la cavidad se aproxima al régimen sub-nanométrico, la hidrofobicidad de la superficie aumenta significativamente. Adicionalmente, las fluctuaciones de densidad de agua alrededor del grafeno

coinciden con las de una superficie hidrofílica (SAM con grupos -OH en la superficie), un comportamiento muy diferente a aquel encontrado para la SAM hidrofóbica. Además mostramos que el término de atracción agua-carbono es relevante en los sistemas gráfiticos, ya que al atenuar artificialmente dichas interacciones, las fluctuaciones de densidad de agua aumentan en la superficie del grafeno, invirtiendo su comportamiento hacia uno típico hidrofóbico. Finalmente mostramos que mediante una apropiada parametrización del sistema grafeno-agua (basada en resultados teóricos y experimentales previos), dos placas de grafeno paralelas presentan una fuerte tendencia a permanecer completamente hidratadas y a no autoensamblarse. En cambio dos SAMs hidrofóbicas enfrentadas presentan el comportamiento opuesto, excluyen la capa de agua intermedia y se autoensamblan durante un típico colapso hidrofóbico. Se espera que estos conocimientos sean de interés tanto en procesos de autoensamblado de materiales específicos como en procesos de organización biológica como el *protein binding*. Esperamos aplicar los parámetros y procedimientos desarrollados en este capítulo en dicho tipo de sistemas de mayor complejidad.

En este tema se colaboró con el Prof. Pablo Debenedetti de Princeton University. Sebastián Accordino realizó una visita a Princeton en el marco de un proyecto de cooperación bilateral CONICET - NSF.

## Capítulo 8

# Conclusiones generales

En esta Tesis estudiamos el comportamiento estructural y dinámico del agua en diferentes contextos de nanoconfinamiento como así también en el *bulk*. Específicamente en el Capítulo 2 exploramos la robustez y validez (para diferentes densidades y temperaturas) de un método para determinar moléculas de agua estructuradas (baja densidad local, LDL) y desestructuradas (alta densidad local, HDL) en agua *bulk*. Determinamos la calidad (geométrica) de la estructura local de ambos tipos de moléculas. Observamos que la dependencia con la temperatura de la fracción de moléculas estructuradas puede ajustarse con una exponencial negativa, lo cual brinda un sustento a modelos *two state* que postulan que la misma puede ser gobernada por un peso de Boltzmann debido a la ganancia energética al perfeccionar las interacciones locales al estructurarse, y al costo dado en el incremento local del volumen. Estos trabajos se realizaron en colaboración con el Prof. Francesco Sciortino, de la Universidad de Roma, La Sapienza.

En el Capítulo 3, estudiamos la estructura y orientación (en función de la temperatura) del agua de hidratación de la proteína lisozima y comparamos dichos resultados con el de una superficie modelo como lo es una placa de grafeno. Mostramos que en los dos casos, estas primeras capas de hidratación presentan una mejor estructura local que el *bulk* a temperatura ambiente, sin embargo la interfase introduce una restricción geométrica que reduce tanto la coordinación tetraédrica típica del agua como la densidad local de las moléculas de agua adyacentes a la superficie. Observamos que el ordenamiento local tetraédrico mejora con el descenso de la temperatura y, en el caso de la placa de grafeno, muestra una tendencia a asemejarse al hielo

hexagonal Ih (similar al comportamiento en la interfase agua-aire pero en el sentido contrario a lo largo del eje-c del plano basal). En cambio, el caso de las primeras capas de hidratación de la lisozima parece ser más complicado, pero aun así, encontramos indicios de ambos tipos de comportamientos antes mencionados, junto con una tendencia de algunas de las moléculas de agua más próximas a la superficie a formar puentes de hidrógeno con la proteína. Basandonos en estudios previos de otros autores, en el Capítulo 3 demostramos (mediante el cálculo de índices de estructura local de agua) que la supuesta transición vítrea de proteínas se debe en realidad a una transición dinámica del solvente, es decir del agua, entre dos fases de agua líquida con diferente densidad local, alta y baja densidad local (HDL y LDL).

El conocimiento ganado sobre la estructura y dinámica del agua, en particular del agua de hidratación, nos condujo a estudiar el papel que juega en diferentes sistemas biológicos. En colaboración con el Prof. Ariel Fernández hemos estudiado procesos de unión a proteínas y entre proteínas basados en un enfoque que codifica el comportamiento del agua nanoconfinada en un motivo estructural, los defectos de empaquetamiento dados por puentes de hidrógeno (*backbone hydrogen bonds*, BHBs) mal “arropados” o dehidrones. Los dehidrones son BHBs no completamente sustraídos intramolecularmente por grupos hidrofóbicos del efecto disruptivo del agua, de modo que resultan “pegajosos” al requerir “arropamiento” intermolecular y, por lo tanto, promover la asociación con ligandos. En ese contexto en el Capítulo 4 mostramos que dichos motivos estructurales muestran una propensión a la deshidratación local y, por lo tanto, representan regiones de alta hidrofobicidad local en proteínas. Esta tendencia a la labilización del agua circundante que poseen dichos motivos (demostrada en el Capítulo 4) promueve la asociación de ligandos (otras proteínas o drogas) que deben reemplazar al agua de hidratación y que completan la desolvatación o secado del BHB. Presentamos firme evidencia de la existencia de un *link* o conexión entre el grado de exclusión o “arropamiento” de los BHBs (el parámetro estructural  $\rho$ ) y la movilidad de las moléculas de agua de hidratación. Específicamente mostramos que la exposición del *backbone* está correlacionada con la pérdida de hidratación de la superficie de la proteína y así, estas interacciones intramoleculares desprotegidas o expuestas al agua de hidratación (dehidrones) se encuentran rodeadas de moléculas de agua debilmente unidas y facilmente removibles. De esta manera, promoviendo su deshidratación local, estos mo-

tivos estructurales llamados dehidrones proporcionan un recurso por el cual un ligando podría desplazar a las moléculas de agua durante el proceso de *binding*.

Estos resultados nos llevaron a estudiar diferentes interfases proteína-proteína como así también entre proteínas y pequeñas moléculas o drogas disruptivas de interfases proteína-proteína. En el Capítulo 5 mostramos que las interfases proteína-proteína eluden el análisis fisicoquímico estandar. Demostramos que un simple motivo molecular, los defectos de empaquetamiento en proteínas o dehidrones, presentado en el Capítulo 4, y edificado sobre los conceptos desarrollados por el Prof. Ariel Fernández, demuestra ser ubicuo en la determinación de asociaciones entre proteínas. Verificamos que dicho motivo es un indicador de la naturaleza cooperativa de las interacciones de tres cuerpos. En esencia, el concepto de cooperativismo surge de la estabilización proporcionada por un grupo hidrofóbico, de una cadena proteica, sobre un BHB de la proteína *partner* o un BHB intermolecular. De esta manera mediante el desarrollo de un *alanine scanning in silico* o computacional, basado en estos conceptos, logramos predecir satisfactoriamente los *hot spots* de un conjunto de complejos proteína-proteína, y así verificar la importancia de las interacciones de tres cuerpos o cooperativas en la asociación de proteínas. Específicamente logramos racionalizar la naturaleza de los *hot spots* y así de las interacciones en las interfases proteína-proteína, un problema que desde hace mucho tiempo no ha podido resolverse dado que solo se han considerado enfoques basados en la aditividad de las interacciones de a pares. Estos conceptos son de vital ayuda para el diseño racional de pequeñas moléculas disruptivas de interfases proteína-proteína, las cuales, en base a un diseño racional, pueden imitar la capacidad “protectora” de los *hot spots*. Por lo tanto en el Capítulo 6 verificamos la importancia de estas interacciones en la interacción entre drogas y proteínas al verificar que un conjunto de drogas disruptivas de interfases proteína-proteína tienden a mimetizar las interacciones cooperativas intermoleculares de la proteína que reemplazan. Sucesivamente vimos que las interacciones clásicas, como lo son los puentes de hidrógeno intermoleculares o puentes salinos intermoleculares, en general no son mimetizados por la molécula disruptiva. También, en base a estos conceptos, estudiamos el caso particular de la mutación oncogénica Y220C de la proteína p53, conocido como el “guardián del genoma”, el cual se encuentra mutado en aproximadamente la mitad de los casos de cancer en seres humanos. Mostramos que la mutación Y220C, que provoca



la formación de una cavidad en la superficie de la proteína, también conduce a la creación de un dehidrón y a la pérdida de la estructura local en dicho sitio de la proteína. De esta manera racionalizamos la naturaleza desestabilizante de esta mutación en particular. Por otro lado, la molécula PhiKan083 reestablece parcialmente la función del p53 estabilizando la estructura de la proteína. Mostramos que la molécula se une a la proteína precisamente en el sitio donde se encuentran los dehidrones Cys 220 - Thr 155 y Leu 145 - Thr 230 interactuando cooperativamente con estos BHBs. Mostramos que la protección suministrada por la molécula PhiKan083, en términos de interacciones cooperativas, es incompleta y por lo tanto realizamos una modificación de la molécula PhiKan083 con el fin de optimizarla. Mostramos que nuestra modificación de la molécula PhiKan083 de hecho fortalece la estructura de la proteína al estabilizar el BHB que se ve afectado (generando un dehidrón) tras la mutación oncogénica Y220C, donde la protección suministrada por la molécula PhiKan083 es incompleta.

Dado que los principios que gobiernan la organización biológica son similares a los que determinan el autoensamblado de materiales sintéticos en solución acuosa, en el Capítulo 7 extendemos preliminarmente las nociones anteriores a sistemas que habitan la vasta provincia de la ciencia de materiales en medio acuoso. Sin embargo, en estos contextos mucho más amplios la existencia de un motivo estructural prevalente no emerge claramente dada la heterogeneidad de sistemas de estudio. En particular, estudiamos el comportamiento del agua interfacial y nanoconfinada y las condiciones de hidrofobicidad local en ciertos materiales y en procesos de ensamblado en estudios sistemáticos con química y geometría controladas. Específicamente estudiamos diferentes parámetros como las fluctuaciones de densidad y tiempos de residencia del agua alrededor de una placa de grafeno, en el exterior de nanotubos de carbono de pared simple, en la superficie de monocapas autoensambladas de alcanos (SAMs) y en poros o cavidades de diferente tamaño talladas en la superficie de las SAMs. Mostramos que la curvatura convexa de las superficies gráficas no produce efectos significativos sobre la dinámica del agua de hidratación y así el grado de hidrofobicidad/hidrofilicidad de estas superficies no se ve mayormente afectado por la curvatura de las mismas. Sin embargo, se observa una marcada dependencia en la hidrofobicidad con la curvatura de las superficies concavas de los poros tallados en las SAMs. Demostramos que las fluctuaciones de densidad de agua alrededor de una placa de grafeno son

prácticamente idénticas a las fluctuaciones sobre una superficie hidrofílica y muy diferentes de aquellas para una superficie hidrofóbica como las SAMs. Por último, por medio de estudios dinámicos mostramos que dicha hidrofiliidad del grafeno se traduce en un comportamiento de interés práctico: previene la agregación de dos placas de grafeno. En este tema colaboramos con el Prof. Pablo Debenedetti de Princeton University. Sebastián Accordino realizó una visita a Princeton en el marco de un proyecto de cooperación bilateral CONICET - NSF.

En definitiva, en esta Tesis hemos estudiado el comportamiento del agua *bulk* (en el régimen de líquido normal y sobreenfriado) y en ambientes de nanoconfinamiento. Esta última cuestión resulta de gran interés en campos de enorme relevancia pues, como ya se ha indicado, el agua oficia de mediador entre las complejas superficies que se autoensamblan en los procesos de organización biológica (*protein binding*), como así también en el autoensamblado de materiales sintéticos en medio acuoso. El consiguiente nanoconfinamiento afecta las propiedades termodinámicas del agua de modo de producir el secado que requieren los procesos de autoensamble y unión. Sin embargo, hoy no existe una teoría de la hidrofobicidad en la nanoescala. Como primer paso, es necesario entonces la realización de estudios como los que aporta este trabajo de Tesis de modo de identificar principios generalizables. Asimismo los procesos de autoensamble en ambientes acuosos se producen a través del establecimiento de interacciones no-covalentes dentro de entornos de nanoconfinamiento. Por lo tanto comprender como dichas interacciones no-covalentes dependen del entorno (principalmente del comportamiento local del solvente y, por ende, de la hidrofobicidad local así como también su contribución posiblemente no aditiva) es un primer paso para intentar extraer elementos de diseño de impacto generalizado. Sin un profundo conocimiento del comportamiento del agua en la nanoescala (es decir del agua nanoconfinada) no será posible comprender las bases moleculares tanto de la biología como de la ciencia de materiales en medio acuoso, ni trascender los esfuerzos de diseño en dichos contextos hacia una etapa de pleno diseño racional. Este trabajo de Tesis intenta humildemente, aportar de manera preliminar a dicho conocimiento.

# Bibliografía

## Libros, Tesis de doctorado y páginas web

- [1] D. C. Malaspina, Tesis Doctoral, Universidad Nacional del Sur (Argentina), (2011).
- [2] A. Fernández, *Transformative Concepts for Drug Design: Target Wrapping* Springer, Heidelberg Dordrecht Londres Nueva York, 2010.
- [3] J. A. Rodríguez Fris, Tesis Doctoral, Universidad Nacional del Sur (Argentina), (2008).
- [4] Nicole Zache, Tesis Doctoral, Karolinska Institute (Sweden), (2007).
- [5] P. Atkins y J. de Paula, *Physical Chemistry for the Life Sciences* W. H. Freeman and Company, Nueva York, 2006.
- [6] P. G. Debenedetti, *Metaestable Liquids: Concepts and Principles*, Princeton University Press, Princeton, 1996.
- [7] C. R. Cantor y P. R. Schimmel, *Biophysical Chemistry Part I: The conformation of biological macromolecules* W. H. Freeman and Company, Nueva York, 1980.
- [8] C. R. Cantor y P. R. Schimmel, *Biophysical Chemistry Part II: Techniques for the study of biological structure and function* W. H. Freeman and Company, Nueva York, Spring 1980.
- [9] <http://www.lsbu.ac.uk/water/protein.html>
- [10] <http://www.lsbu.ac.uk/water/models.html>
- [11] <http://ambermd.org/tutorials/basic/tutorial1/section5/>

[12] <http://ambermd.org/tutorials/advanced/tutorial3/>

[13] <http://ambermd.org/tutorials/advanced/tutorial4/>

### Artículos de nuestro grupo de investigación

[14] L. M. Alarcón, J. M. Montes de Oca, S. R. Accordino, J. A. Rodríguez Fris y G. A. Appignanesi, *Fluid Phase Equil.* **362**, 81 (2014).

[15] S. R. Accordino, J. A. Rodríguez Fris, G. A. Appignanesi, *PLoS ONE* **8**, e55123 (2013).

[16] M. B. Sierra, S. R. Accordino, J. A. Rodríguez Fris, M. A. Morini, G. A. Appignanesi y A. Fernández, *Eur. Phys. J. E.* **36**, 62 (2013).

[17] S. R. Accordino, D. C. Malaspina, J. A. Rodríguez Fris, L. M. Alarcón y G. A. Appignanesi, *Phys. Rev. E* **85**, 031503 (2012).

[18] S. R. Accordino, J. A. Rodríguez Fris, G. A. Appignanesi y A. Fernández, *Eur. Phys. J. E.* **35**, 59 (2012).

[19] S. R. Accordino, M. A. Morini, M. B. Sierra, J. A. Rodríguez Fris, G. A. Appignanesi y A. Fernández, *Proteins: Struct. Funct. and Bioinf.* **80**, 1755 (2012).

[20] S. R. Accordino, D. C. Malaspina, J. A. Rodríguez Fris y G. A. Appignanesi, *Phys. Rev. Lett.* **106**, 029801 (2011).

[21] S. R. Accordino, J. A. Rodríguez Fris, F. Sciortino y G. A. Appignanesi, *Eur. Phys. J. E.* **34**, 48 (2011).

[22] L. M. Alarcón, D. C. Malaspina, E. P. Schulz, M. A. Frechero y G. A. Appignanesi, *Chem. Phys.* **388**, 47 (2011).

[23] E. P. Schulz, L. M. Alarcón y G. A. Appignanesi, *Eur. Phys. J. E.* **34**, 114 (2011).

[24] D. Malaspina, E. P. Schulz, L. M. Alarcón, M. A. Frechero y G. A. Appignanesi, *Eur. Phys. J. E.* **32**, 35 (2010).

- [25] D. C. Malaspina, E. P. Schulz, L. M. Alarcón, M. A. Frechero y G. A. Appignanesi, *Eur. Phys. J. E* **32**, 35 (2010).
- [26] E. Schulz, M. Frechero, G. A. Appignanesi y A. Fernández, *PLoS ONE* **5**, 12844 (2010).
- [27] G. A. Appignanesi, J. A. Rodríguez Fris y F. Sciortino, *Eur. Phys. J. E* **29**, 305 (2009).
- [28] D. C. Malaspina, J. A. Rodríguez Fris, G. A. Appignanesi y F. Sciortino, *Europhys. Lett.* **88**, 16003 (2009).
- [29] J. A. Rodríguez Fris, G. A. Appignanesi, E. La Nave y F. Sciortino, *Phys. Rev. E* **75**, 041501 (2007).
- [30] G. A. Appignanesi, J. A. Rodríguez Fris, R. A. Montani y W. Kob, *Phys. Rev. Lett.* **96**, 057801 (2006).
- [31] G. A. Appignanesi, J. A. Rodríguez Fris y M. A. Frechero, *Phys. Rev. Lett.* **96**, 237803 (2006).

### Artículos de otros autores

- [32] J. C. Palmer, F. Martelli, Y. Liu, R. Car, A. Z. Panagiotopoulos y P. G. Debenedetti, *Nature* **510**, 385 (2014).
- [33] D. Chandler arXiv: 1407.6854v2 (2014).
- [34] J. Russo y H. Tanaka, *Nature Comm.* **5**, 4556 (2014).
- [35] E. A. Disalvo, M. F. Martini, A. M. Bouchet, A. Hollmann, M. A. Frías, *Adv. Colloid Interface Sci.* **211**, 17 (2014).
- [36] D. T. Limmer y D. Chandler, *J. Chem. Phys.* **138**, 214504 (2013).
- [37] Z. Li et al. *Nature Materials* **12**, 925 (2013).
- [38] F. Taherian, V. Marcon y N. F. A. van der Vegt, *Langmuir* **29**, 1457 (2013).

- [39] E. A. Disalvo, M. A. Frías, *Langmuir* **29**, 6969 (2013).
- [40] N. Giovambattista, P. J. Rossky, P. G. Debenedetti, *Annu. Rev. Phys. Chem.* **63**, 179 (2012).
- [41] R. Wilcken, X. Liu, M. O. Zimmermann, T. J. Rutherford, A. R. Fersht, et al. *J. Am. Chem. Soc.* **134**, 6810 (2012).
- [42] J. Rafiee et al. *Nature Matter.* **11**, 217 (2012).
- [43] C. J. Shih et al. *Phys. Rev. Lett.* **109**, 176101 (2012).
- [44] A. Nilsson, C. Huang y L. G. M. Pettersson, *J. Mol. Liq.* **176**, 2 (2012).
- [45] T. Loerting et al. *Phys. Chem. Chem. Phys.* **13**, 8783 (2011).
- [46] J. L. Kulp III, et al., *J. Am. Chem. Soc.* **133**, 10740 (2011).
- [47] T. Geppert, B. Hoy, S. Wessler y G. Schneider, *Chem. Biol.* **18**, 344 (2011).
- [48] A. Fernández y M. Lynch, *Nature* **474**, 502 (2011).
- [49] C. J. Brown, C. F. Cheok, C. S. Verma y D. P. Lane, *Trends in Pharm. Science* **32**, 53 (2011).
- [50] C. J. Brown, C. F. Cheok, C. S. Verma, D. P. Lane, *Trends in Pharm. Sci.* **32**, 53 (2011).
- [51] C. Wang, B. J. Berne, R. A. Friesner, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1326 (2011).
- [52] K. S. Kim et al. *ACS Nano* **5**, 5107 (2011).
- [53] S. Dong, Z. Chen y Y. Wang, *Chem. Phys.* **369**, 8 (2010).
- [54] H. Acharya et al., *Faraday Discuss.* **146**, 353 (2010).
- [55] T.H. Rehm, C. Schmuck, *Chem. Soc. Rev.* **39**, 3597 (2010).
- [56] W. Doster, *Biochimica et Biophysica Acta* **1804**, 3 (2010).
- [57] G-Y Chuang et al., *Protein Science* **19**, 1662 (2010).

- [58] Z. Li y J. Li, *Bioinformatics* **78**, 3304 (2010).
- [59] N. Basse, J. L. Kaar, G. Settanni, A. C. Joerger, T. J. Rutherford, et al. *Chem. and Biol.* **17**, 46 (2010).
- [60] Y. J. Shin et al. *Langmuir* **26**, 3798 (2010).
- [61] C. U. Kim, B. Barstow, M. W. Tate y S. M. Gruner, *Proc. Nat. Acad. Sci.* **106**, 4596 (2009).
- [62] N. Giovambattista, P. G. Debenedetti, P. J. Rossky, *Proc. Natl. Acad. SCI USA* **106**, 15181 (2009).
- [63] B. J. Berne et al., *Annu. Rev. Phys. Chem.* **60**, 85 (2009).
- [64] M. Vogel, *J. Phys. Chem. B* **113**, 28 (2009).
- [65] J. Li y Q. Liu, *Bioinformatics* **25**, 743 (2009).
- [66] G. Rastelli, A. Del Rio, G. Degliesposti, M. Sgobba, *J. Comp. Chem.* **31**, 797 (2009).
- [67] M. P. Longinotti y H. R. Corti, *J. Phys. Chem. B* **113**, 5500 (2009).
- [68] K. L. Ngai, S. Capaccioli y N. Shinyashiki, *J. Phys. Chem. B* **112**, 3826 (2008).
- [69] F. M. Boeckler, A. C. Joerger, G. Jaggi, T. J. Rutherford, D. B. Veprintsev, et al. *Proc. Natl. Acad. Sci. U.S.A* **105**, 10360 (2008).
- [70] J. C. Rasaiah, S. Garde, G. Hummer, *Annu. Rev. Phys. Chem.* **59**, 713 (2008).
- [71] G. Cicero et al. *J. Am. Chem. Soc.* **130**, 1871 (2008).
- [72] J. Qvist, M. Davidovic, D. Hamelberg, B. Halle, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 6296 (2008).
- [73] R. Garcia, K. Osborne y E. Subashi *J. Phys. Chem. B* **112**, 8114 (2008).
- [74] M. P. Longinotti y H. R. Corti, *J. Phys. Chem. Ref. Data* **37**, 1503 (2008).

- [75] H. E. Stanley, P. Kumar, L. Xu, Z. Yan, M. G. Mazza, S. V. Buldyrev, S. H. Chen y F. Mallamace, *Physica A* **386**, 729 (2007).
- [76] J. A. Wells, C. L. McClendon, *Nature* **450**, 1001 (2007).
- [77] Y. Ofran y B. Rost, *PLoS Comp. Biol.* **3**, 1169 (2007).
- [78] P. Privalov et al., *J. Mol. Biol.* **365**, 1 (2007).
- [79] N. Pietrosevoli, A. Crespo y A. Fernández, *J. Prot. Res.* **6**, 3519 (2007).
- [80] M. Bruncko, et al. *J. Med. Chem.* **50**, 641 (2007).
- [81] T. Young, R. Abel, B. Kim, B. J. Berne, R. A. Friesner, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 808 (2007).
- [82] A. Fernández, J. Chen, A. Crespo, *J. Chem. Phys.* **126**, 245103 (2007).
- [83] D. A. Case *et al.*, AMBER 9, University of California, San Francisco, CA (2006).
- [84] V. Hornak *et al.* *Proteins* **65**, 712-725 (2006).
- [85] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. and Mod.* **25** 247260 (2006).
- [86] Y. R. Shen y V. Ostroverkhov, *Chem. Rev.* **106**, 1140 (2006).
- [87] P. Kumar, Z. Yan, L. Xu, M. G. Mazza, S. V. Buldyrev, S. H. Chen, S. Sastry y H. E. Stanley, *Phys. Rev. Lett.* **97**, 177802 (2006).
- [88] T. Loerting, N. Giovambattista, *J. Phys.: Condens. Matter* **18**, 919 (2006).
- [89] L. T. Vassilev, *Trends Mol. Med.* **13**, 23 (2006).
- [90] A. C. Joerger, H. C. Ang, A. R. Fersht, *Proc. Natl. Acad. Sci. U.S.A* **103**, 15056 (2006).
- [91] Y. R. Shen y V. Ostroverkhov, *Chem. Rev.* **106**, 1140 (2006).
- [92] R. A. Friesner et al., *J. Med. Chem.* **49**, 6177 (2006).



- [93] P. Kumar, Z. Yan, L. Xu, M. G. Mazza, S. V. Buldyrev, S. H. Chen, S. Sastry, H. E. Stanley, *Phys. Rev. Lett.* **97**, 177802 (2006).
- [94] P. W. Fenimore, H. Frauenfelder, B. H. McMahon y R. D. Young, *Phys. A* **351**, 113 (2005).
- [95] O. Keskin, B. Ma y R. Nussinov, *J. Mol. Biol.* **345**, 1281 (2005).
- [96] T. S. Rush, J. A. Grant, L. Mosyak, A. Nicholls, *J. Med. Chem.* **48**, 1489 (2005).
- [97] N. Choudhury y B. Montgomery Pettitt, *J. Phys. Chem. B* **109**, 6422 (2005).
- [98] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **25**, 1157-1174 (2004).
- [99] P. W. Fenimore y H. Frauenfelder, *PNAS* **101**, 40 (2004).
- [100] M. R. Arkin y J. A. Wells, *Nature Rev. Drug Disc.* **3**, 301 (2004).
- [101] L. T. Vassilev, et al. *Science* **303**, 844 (2004).
- [102] B. C. Raimundo, et al. *J. Med. Chem.* **47**, 3111 (2004).
- [103] M. Feig, A. Onufriev, M. S. Lee, W. Im, D. A. Case, et al. *J. Comp. Chem* **25**, 265 (2004).
- [104] M. Feig, W. Im, C. L. Brooks, *J. Chem. Phys.* **120**, 903 (2004).
- [105] S. Vaitheeswaran, H. Yin, J. C. Rasaiah, G. Hummer, *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17002 (2004).
- [106] A. Fernández, *Nat. Biotech.* **22**, 1081 (2004).
- [107] A. H. Barber, S. R. Cohen y H. D. Wagner, *Phys. Rev. Lett.* **92**, 186103 (2004).
- [108] P. G. Debenedetti, H. E. Stanley, *Phys. Today* **56**, 40 (2003).
- [109] C. F. J. Faul, M. Antonietti, *Adv. Mater.* **15**, 673 (2003).
- [110] P. Ball, *Nature (London)* **423**, 25 (2003).
- [111] M. Sasai, *J. Chem. Phys.* **118**, 10651 (2003).

- [112] E. G. Ponyatovsky, *J. Phys.: Condens. Matter* **15**, 6123 (2003).
- [113] A. Fernández y R. Scott, *Phys. Rev. Lett.* **91**, 018102 (2003).
- [114] A. Fernández y R. Scott, *Biophysical J.* **85**, 1914 (2003).
- [115] A. Friedler, D. B. Veprintsev, L. O. Hansson, A. R. Fersht, *J. Biol. Chem.* **278**, 24108 (2003).
- [116] R. Steitz, T. Gutberlet, T. Hauss, B. Klösgen, R. Hratev, S. Schemmel, A. C. Simonsen y G. H. Findenegg, *Langmuir* **19**, 2409 (2003).
- [117] T. R. Jensen, M. Ø. Jensen, N. Reitzel, K. Balashev, G. H. Peters, K. Kjaer y T. Bjørnholm, *Phys. Rev. Lett.* **90**, 086101 (2003).
- [118] A. Fernández, H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 113 (2003).
- [119] X. Huang, C. J. Margulis, B. J. Berne, *Proc. Natl. Acad. Sci. U. S. A.* **100**, 11953 (2003).
- [120] T. Werder et al. *J. Phys. Chem. B* **197**, 1345 (2003).
- [121] A. R. Bizarri y S. Cannistraro, *J. Phys. Chem. B* **106** (2002).
- [122] T. Kortemme y D. Baker, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14116 (2002).
- [123] P. Chakrabarti y J. Janin, *Proteins* **47**, 334 (2002).
- [124] K. R. Wilson, R. D. Schaller, D. T. Co, R. J. Saykally, B. S. Rude, T. Catalano y J. D. Bozek, *J. Chem. Phys.* **117**, 7738 (2002).
- [125] J. R. Errington y P. G. Debenedetti, *Nature (London)* **409**, 318 (2001).
- [126] P. G. Debenedetti y F. H. Stillinger, *Nature* **410**, 259 (2001).
- [127] T. Loerting, C. Salzmann, I. Kohl, E. Mayer, A. Hallbrucker, *Phys. Chem. Chem. Phys. (Inc. Faraday Trans.)* **3**, 5355 (2001).
- [128] K. S. Thorn y A. A. Bogan, *Bioinformatics* **17**, 284 (2001).

- [129] T. V. Chalikian, *J. Phys. Chem B* **105**, 12566 (2001).
- [130] G. Hummer, J. C. Rasaiah, J. P. Noworyta, *Nature* **414**, 188 (2001).
- [131] M. W. Mahoney and W. L. Jorgensen, *J. Chem. Phys.* **112**, 8910 (2000).
- [132] A. R. Bizarri, A. Paciaroni, S. Cannistraro, *Phys. Rev. E* **62**, 3991 (2000).
- [133] M. Sasai, *Physica A* **285**, 315 (2000).
- [134] H. Tanaka, *Europhys. Lett.* **50**, 340 (2000).
- [135] H. Tanaka, *J. Chem. Phys.* **112**, 799 (2000).
- [136] U. C. Khlmann et al., *J. Mol. Biol.* **301**, 1163 (2000).
- [137] A. N. Bullock, J. Henckel, A. R. Fersht, *Oncogene* **19**, 1245 (2000).
- [138] M. R. Lee, Y. Duan, P. A. Kollman, *Proteins* **39**, 309 (2000).
- [139] J. Urquidi, S. Singh, C. H. Cho, G. W. Robinson, *Phys. Rev. Lett.* **83**, 2348 (1999).
- [140] O. Mishima, H. E. Stanley, *Nature* **396**, 329 (1998).
- [141] E. G. Ponyatovsky, V. V. Sinitsyn, T. A. Pozdnyakova, *J. Chem. Phys.* **109**, 2413 (1998).
- [142] H. Tanaka, *Phys. Rev. Lett.* **80**, 5750 (1998).
- [143] A. A. Bogan y K. S. Thorn, *J. Mol. Biol.* **280**, 1 (1998).
- [144] C. V. Sindelar, Z. S. Hendsch y B. Tidor, *Protein Science* **7**, 1898 (1998).
- [145] E. Shiratani y M. Sasai, *J. Chem. Phys.* **108**, 3264 (1998); E. Shiratani y M. Sasai, *J. Chem. Phys.* **104**, 7671 (1996).
- [146] A. Böttger et al., *J. Mol. Biol.* **269**, 744 (1997).
- [147] C-Z Chen y R. Shapiro, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1761 (1997).
- [148] A. Böttger, V. Böttger, C. Gracia Echeverría, P. Chène, H. K. Hochkeppel, W. Sampson, K. Ang, S. F. Howard, S. M. Pickles, D. P. Lane, *J. Mol. Biol.* **269**, 744 (1997).

- [149] A. N. Bullock, J. Henckel, B. S. DeDecker, C. M. Johnson, P. V. Nikolova, et al. Proc. Natl. Acad. Sci. U.S.A **94**, 14338 (1997).
- [150] C. H. Cho, S. Singh, G. W. Robinson, Phys. Rev. Lett. **76**, 1651 (1996).
- [151] M. J. M. Castro y S. Anderson, Biochemistry **35**, 11435 (1996).
- [152] W. Kob y H. C. Andersen, Phys. Rev. E **51**, 4626 (1995).
- [153] T. Clackson y J. A. Wells, Science **267**, 383 (1995).
- [154] G. Schreiber y A. R. Fersht, J. Mol. Biol. **248**, 478 (1995).
- [155] E. G. Ponyatovsky, V. V. Sinitsyn, T. A. Pozdnyakova, JETP Lett. **60**, 360 (1994).
- [156] R. J. Loncharich, B. R. Brooks y R. W. Pastor, Biopolymers **32**, 523-535 (1992).
- [157] S. Miyamoto, P. A. Kollman, J. Comput. Chem. **13**, 952-962 (1992).
- [158] P. H. Poole, F. Sciortino, U. Essmann y H. E. Stanley, Nature (London) **360**, 324 (1992).
- [159] F. H. Stillinger, Phys. Rev. B **41**, 2409 (1990).
- [160] A. Ashkenazi et al., Proc. Nat. Acad. Sci. U.S.A. **87**, 7150 (1990).
- [161] H. J. C. Berendsen, J. R. Grigera, T. P. Stroatsma, J. Phys. Chem. **91**, 6269 (1987).
- [162] O. Mishima, L. D. Calvert, E. Whalley, Nature **310**, 393 (1984).
- [163] H. G. Heide, Ultramicroscopy **14**, 271 (1984).
- [164] M. L. Connolly, J. Appl. Cryst. **16**, 548 (1983).
- [165] P. J. Artymiuk, C. C. F. Blake, D. W. Rice, K. S. Wilson, Acta Crystallogr. Sect. B **38**, 778-783 (1982).
- [166] O. Steinhauser, Mol. Phys. **45**, 335 (1982).
- [167] A. Rahman y F. H. Stillinger, J. Chem. Phys. **55**, 3336 (1971).

[168] I. Morcos, *J. Colloid Interf. Sci.* **34**, 469 (1970).

[169] J. A. Barker y R. O. Watts, *Chem. Phys. Lett.* **3**, 144 (1969).

[170] F. M. Fowkes y W. D. Harkins, *J. Am. Chem. Soc.* **62**, 3377 (1940).