Protein hydration shell formation: Dynamics of water in biological systems exhibiting nanoscopic cavities

M.C. Morón *

Instituto de Nanociencia y Materiales de Aragón (INMA), Consejo Superior de Investigaciones Científicas (CSIC) - Universidad de Zaragoza, Zaragoza 50009, Spain
Departamento de Física de la Materia Condensada, Facultad de Ciencias, Universidad de Zaragoza, Zaragoza 50009, Spain

A R T I C L E   I N F O

Article history:
Received 10 November 2020
Revised 17 May 2021
Accepted 22 May 2021
Available online 27 May 2021

Keywords:
Protein hydration shell formation
Hydrogen bond network
Nanoscopic cavities in water solutions
Bulk water dynamics
Two states model
All-atoms molecular dynamics simulations

A B S T R A C T

This work explores the singular scenarios emerging from nanoscopic cavities, located in aqueous solutions, that include biomolecules and, as a consequence, the process of biomolecule hydration shell formation. The research presents a nano-scale study, performed in various systems related with protein-water solutions, using all-atoms molecular dynamics simulations. This research shows that if a protein falls within an empty nanoscopic cavity located in an aqueous solution, it will take a time, with a magnitude significant for biological processes, to rebuild its whole network of hydrogen bonds with the solvent molecules. During that protein isolation time the dynamics of the biomolecule, and therefore the corresponding bioactivity, will be seriously compromised. In the case of the protein barstar (radius of gyration of 1.17 nm) located in the centre of cavities with radius $r$ from 2.5 to 4.5 nm, and solvent diffusion coefficients for bulk and physiologic water (2.4 and 1.5 $\text{m}^2/\text{ms}$, respectively), that time is found to be of the order of tens-hundreds of picoseconds, a significant temporal range concerning the dynamics and bioactivity of proteins. On the other hand, the dynamics of formation of the inner protein hydration shell has been followed using an atomic view. The required time has been found to decrease with $r$, as the network of water molecules approaching the biomolecule resembles that of the biological water corresponding to that biomolecule. That resemblance increases as those water molecules have previously been in close contact with the protein. The dynamics of those systems has been modelled using a two states model. Isolated bulk water is taken as reference for the computer simulations. Comparison with experimental data is also provided.

© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Water plays a ubiquitous and essential role on Earth. Its peculiar properties are crucial in relevant physical, chemical, biological and technical processes [1–3]. Many of those properties are attributable to the ability of its molecules to form hydrogen bonds with other water molecules given place to extended three-dimensional networks. Thus the dynamics of water will be dominated by the behaviour of its continuously rebuilt hydrogen-bonded network [4,5].

Although being one of the simplest molecules, water participates in nearly every process necessary for life [6–8]. Biomolecules like nucleic acids, proteins, membranes etc., are usually in close contact with water solutions. Thus the layer of water molecules in intimate interaction with a protein surface is typically called protein hydration shell or biological water [9–12]. That particular solvent supports vital biophysical and biochemical processes in living organisms, influencing key cellular functions such as protein folding and stability, enzyme catalysis, DNA packaging, molecular recognition, cellular tolerance to freezing, intracellular signalling and transport processes [5,10,13,14]. Investigating the biophysical-chemical properties of water is also essential for understanding medical phenomena [15–17]. Thus in living tissues the content of tightly bound water to biomolecules decreases with aging. Release of that particular water, characterised by slower dynamics, results in protein conformational changes and in age-related diseases [18].

Moreover a series of experimental and computational works have evidenced that nanocavities in various materials as carbon nanotubes, kaolinite, graphene or Au walls can confine water that often exhibits a dynamics significantly different from that of its bulk counterpart [19–21]. As an example concerning molecular dynamics computer simulations, a model nanopore has been built by removing some carbon atoms of one graphene sheet to analyse
the dewetting process of water encapsulated between graphene walls [22]. The particular dynamics of confined water inside nanocavities located in biological systems as membranes, micelles or lipid bilayers, has also been reported [23–25]. The archetype model of those referred studies is a nanocavity, bounded by a hydrophobic or biological system, with water molecules inside the hollow. In the present work the model we are going to consider can be regarded as the reverse one in the sense that it is the biomolecule, instead the water, which is located inside the nanocavity.

In the following that novel scenario is described. Thus let us consider an empty nanoscopic cavity situated in the network of water molecules of an aqueous solution. If this cavity was gradually filled with the surrounding water molecules, and that process would take a no negligible time, the cavity would not be immediately closed. If before that cavity closes a protein falls in it, for example due to an abrupt movement of the aqueous solution or whatever the reason was, that biomolecule would take some time to regenerate its entire protein-water hydrogen bonds network. During that time, the protein would be isolated from the solvent. In the present work such hypotheses will be analysed.

The nature of the solvent will also be taken into consideration. Furthermore two cases will be contemplated: with or without a biomolecule inside the referred cavity. In the first context the time necessary to completely fill the cavity with solvent would depend on the number of linked water molecules (hydration layer), the protein brings with it when entering the cavity. Since the pioneering works on protein hydration shell three decades ago, an open discussion remains about the water molecules that would integrate the hydration shell of proteins [10,26–30]. In addition to that absence of definition, please mind that the aim of this work is also to follow, with atomic resolution, the process of inner protein hydration shell formation, as the solvent approaches the biomolecule, from the first to the last protein-water hydrogen bond. Thus only the hypothetical situation where the protein does not carry those particular water molecules will be contemplated in the present work.

A ribonuclease inhibitor, that can be found in Bacillus amyloliq-u faciens, has been considered as biomolecule. Thus barstar is an 89 residues protein of interest in cancer cell research [31,32], molecular recognition [drug discovery [32,33]] or neurodegenerative diseases (amyloid fibrils investigations [34,35]). In order to follow the dynamics and existing interactions within the systems analysed, we have carried out all-atoms molecular dynamics simulations. Those numerical simulations are very useful for giving details of dynamical processes that are not experimentally reachable [36,37]. In all cases, the study is performed at nanoscopic level.

The paper is organized as follows. Section 2 provides the procedure that has been followed to carried out more than 400 computer simulations of nanocavities in protein-water solutions. Those simulations are required in order to analyse the various cases contemplated in the present work. Moreover isolated bulk water has been also simulated to be taken as a reference. In Section 3 the dynamics of those systems is studied as a function of the nature of the solvent and the ratio of the nanocavity, and modelled using empirical equations. The dynamics of the protein hydration layer formation is also followed with atomic resolution. Those results are discussed in Section 4 and compared with some available, already reported, experimental data. Finally, Section 5 presents the main conclusions of the work.

2. Methods

All-atoms molecular dynamics simulations of the protein bar-star in aqueous solutions were performed. Numerical calculations, carried out on bulk water, were also computed to guide the rest of simulations.

2.1. Simulation 1: Protein barstar in water

A molecular dynamics simulation of the protein barstar (wild type) in water was conducted at room temperature conditions (300 K and 1 atm). The initial structure of the inhibitor barstar was obtained from the RCSB Protein Data Bank (PDB ID: 1BRS). The water molecules shown in the crystallographic PDB data were not maintained. After adding the hydrogens, the whole protein was immersed in a cubic box of well equilibrated water containing more than 24000 solvent molecules. The periodic boundary condition was applied in order to prevent boundary effects. The energy of the system was minimized using steepest descent and conjugate gradient algorithms implemented in the Gromacs software package [38]. Conjugate gradient is slower than steepest descent in the early stages of the minimization, but becomes more efficient closer to the energy minimum. After that procedure the system was equilibrated during 1000 ps at room temperature conditions under constant temperature and volume (NVT ensemble) and at constant temperature and pressure (1 atm) (NPT ensemble) for additional 1000 ps. After the equilibration period, the trajectories for data production were performed within the NPT ensemble during 1000 ps at 1 atm and 300 K. The cubic box presents an averaged volume of 754 nm³ (corresponding to an edge of 9.1 nm). During all the time, the protein is centred in the box.

The pre-processing and data production were conducted using the Gromacs software package, [38] the multi-component all-atom force field OPLS-AA/L (to model the protein) [39,40] and the four site TIP4P-Ew model (for water) [41]. That solvent model was chosen since, in addition to being consistent with the used protein force field, it is able to nicely reproduce the experimental diffusion coefficient of isolated water [41]. Please mind that the properties analysed in this work relate to the translation of solvent molecules. All bond lengths were constrained using the LINCS algorithm [42]. The cutoff length for the Lenard-Jones potentials was set at 1.0 nm. The electrostatic interactions were calculated using the particle-mesh Ewald (PME) method [43].

Table 1. The temperature of the system was controlled using Langevin dynamics and the pressure with the Parrinello-Rahman method [44]. The simulation was performed using an integration time step of 2 fs (2 × 10⁻¹³ s). The production run was saved every 0.1 ps (1 × 10⁻¹³ s) for later analysis.

2.2. Simulation 2: Nanoscopic cavity in water (negative photogram of the protein)

Simulation 2 starts pulling out the protein from the solution at the final snapshot of simulation 1. Thus the nano-cavity, located at the centre of the simulation box, consists of an empty space with the size and shape of the protein barstar. The rest of the box is occupied with water molecules. That empty space (cavity) constitutes in fact a negative photogram of the protein. The simulation of that particular system was conducted for 500 ps, within the NPT ensemble under the same conditions already described for simulation 1. It can be worth to mention that in order to improve statistics and representativity, not only the final snapshot of simulation 1 was consider as initial step for simulation 2. Thirty five
additional times, regularly spaced along simulation 1, were also contemplated. In fact, simulation 2 is the generic name given to a set of 36 simulations that have been processed and the results averaged.

2.3. Simulations 3a, 3b, 3c, 3d and 3e: Larger cavity in water partially filled with the protein

In this five simulations the initial structural snapshot is an empty spherical cavity, with a radius \( r = 2.5, 3.0, 3.5, 4.0 \) and 4.5 nm located at the centre of the water box and hosting the protein inside. The gyration radius of native \( \text{barstar} \) is 1.17 nm [45]. That initial system has been created using the final snapshot of simulation 1 and removing the water molecules located at a distance smaller than \( r \) nm from the centre of the box. The biomolecule is located at that centre. The temporal evolution of that system was followed during 500 ps, within the NPT ensemble, using the same conditions of the previous simulations. As time increases, the empty space between the protein and the solvent decreases in size. At the end of the process the initial nanoscopic cavity is completely filled with water molecules. As in the case of simulation 2, in order to improve statistics and representativity, not only the final snapshot of simulation 1 was consider as initial step for simulation 3a. Thirty five additional times, regularly spaced along simulation 1, were also contemplated. Simulation 3a is in fact the generic name given to a set of 36 simulations that have been processed and the results averaged. The same consideration is valid for simulations 3b to 3e.

2.4. Simulation 4: Isolated (bulk) water

A computer simulation has been conducted on isolated (bulk) water following the conditions of simulation 1. The cubic box contains more than 9000 water molecules with an average volume of 275 nm\(^3\) (edge 6.5 nm). Simulation 4 was used as a reference for simulations 1, 2, 3a, 3b, 3c, 3d and 3e. Additional information is provided in Section 3.1.

2.5. Simulations 5, 6, 7a, 7b, 7c, 7d, 7e and 8: Physiological conditions

They are, respectively, simulations 1, 2, 3a 3b, 3c, 3d, 3e and 4 performed using a constant time taut = 0.112 ± 0.002 ps and a temperature \( T = 310 \) K, in order to simulate physiological conditions in humans and others (see Table 1 and Section 3.3). As happens with simulation 4, simulation 8 is used as a reference for simulations 5, 6, 7a, 7b, 7c, 7d and 7e. For more information see Section 3.3.

3. Results

In this section the dynamics of nanoscopic cavities, created in water solutions by pulling out either a protein or solvent molecules, is explored. The cavities may or may not contain the biomolecule inside. Such dynamics, followed using molecular dynamics simulations, is studied through temporal distributions of water molecules and hydrogen bond network analysis. Moreover the dynamics of protein hydration shell formation is analysed and modelled using a two state model. The study is performed at nanoscopic level.

3.1. Nanoscopic cavity in water (with and without protein inside):

Referent water

A nanocavity has been created in a water solution by pulling out the protein from its place at the centre of the simulation box. In order to do that, a computer simulation of the inhibitor \( \text{barstar} \) in water was performed at ambient conditions (see Section 2/simulation 1). At the final step of that simulation, the inhibitor was pulled out from solution (see Section 2/simulation 2). The cavity created by following that procedure is a negative photograph of the protein, it presents its shape. The dynamics of that new system, hosting no biological material, was followed as a function of time showing that the size of the empty cavity decreases as time proceeds. The temporal evolution of the number of water molecules inside the cavity, \( < n_{\text{WM}} > \), is showed in Fig. 1 (normalized to 1). At the starting point of simulation 2, when the cavity has just been created, that number of molecules is zero. As time increases, and due to water-water interactions, the empty space is gradually occupied with water molecules until the saturation is reached indicating that the cavity has been completely filled with solvent.

Before proceeding with a numerical analysis, let us consider the following scenario: a protein falls down in one of those cavities before it closes. In order to follow the influence of its size in the dynamics, cavities with initial radius \( r = 2.5, 3.0, 3.5, 4.0 \) and 4.5 nm were considered. Please mind that the radius of gyration for native \( \text{barstar} \) has been reported as 1.17 nm [45]. With the aim of analysing that situation, simulations 3a, 3b, 3c, 3d and 3e were performed. Let us concentrate in one of them, for example the central one, that is 3c \((r = 3.5 \) nm). Following the procedure described in Section 2.3, a cavity of 7.0 nm of diameter was created. The protein centred in that cavity is shown in Fig. 2. That

![Fig. 1. Temporal evolution of the number of water molecules (normalized to 1) inside the nanoscopic cavity which has been originated by pulling out the protein \( \text{barstar} \) from its location in a bath of water molecules. The data stand for referent (orange) and physiologic (purple) water (see Table 1). The green line corresponds to the best fit of the data to Eq. (2).](image-url)
From the contemplation of the five curves in that figure, two particular times can be defined: i) $t_i$ : the maximum time for which $<\text{nHB}> = 0$, determining the moment where protein-water hydrogen bonds start to form and ii) $t_{\text{sat}}$ : the minimum time for which $<\text{nHB}> = 1$, defining the time for which the saturation of those hydrogen bonds is reached. The difference $t_{\text{sat}} - t_i$ establishes the time required since the first hydrogen bond takes place to the formation of the last one. The evolution of $t_{\text{sat}}$ and $t_{\text{sat}} - t_i$ with the radius of the nanocavity is shown, respectively, in Figs. 4 and 5.

Finally, it is essential to mention that as a very first step and since the properties studied in this work concern the translation of solvent molecules, we performed various simulations of isolated water to assure the reproduction of the experimental value of the diffusion coefficient of bulk water ($D^{\text{referent water}}_{\text{exp}} = 2.40 \text{ mm}^2/\text{ms}$ at 300 K and 1 atm [47]). For doing that we tuned the taut parameter in the Langevin equation of motion (see Section 2/simulation 1) [48,49]. Using $\text{taut} = 2.00 \pm 0.02 \text{ ps}$, a diffusion coefficient of $D^{\text{referent water}}_{\text{calc}} = 2.41 \pm 0.02 \text{ mm}^2/\text{ms}$ is obtained for simulation 4. This result is in excellent agreement with the experimental data (see Table 1). Therefore that constant time $\text{taut} = 2.00 \text{ ps}$ has been used in the six sets of simulations reported in the precedent paragraphs.

3.2. Two states model

In this section the dynamics of those systems, empty nanocavities in an aqueous solution with (or without) a biomolecule inside, will be approximated by a two states model. The water molecules that have enter the cavity (or that have established hydrogen bonds with the protein) are in a state that will be called ‘1’ (current is on). Before doing that, their state is ‘0’ (current is off). The model is binary: on (inside cavity or hydrogen bonded to the protein) and off (outside cavity or no hydrogen bonded with the protein).

Inspired in a sigmoid function, we have written the following empirical equation in the case of water molecules entering the cavities previously reported (simulations 3a to 3e) and establishing hydrogen bonds with the protein that is centred in the cavity:

$$<\text{nHB}> = \frac{e^{\frac{\text{ti}}{k_{\text{B}}T}} - e^{\frac{\text{tsat}}{k_{\text{B}}T}}}{e^{\frac{\text{ti}}{k_{\text{B}}T}} + e^{\frac{\text{tsat}}{k_{\text{B}}T}} + e^{\frac{\text{ti}}{k_{\text{B}}T}}}$$

(1)

Initial snapshot was let to evolve with time. We have calculated the number of hydrogen bonds that the protein establishes with the solvent molecules for all the snapshots of the simulation. A standard geometrical criterion for the determination of the presence of a hydrogen bond between two atoms, donor and acceptor, has been used: the hydrogen-donor–acceptor angle smaller than 30° and the distance donor–acceptor within 0.35 nm [46]. The result is shown in Fig. 3 (the temporal evolution of the variable is normalized to 1). At the beginning of the process there is no hydrogen bonds between both entities. As water molecules approaches the biomolecule, that variable, $<\text{nHB}>$, gradually starts to increase until saturation at the end of the process. In order to analyse the influence of the size of the nanoscopic cavity, an equivalent procedure has been followed for cavities with two smaller and two larger radius (see, simulations 3a, 3b, 3d and 3e in Section 2). The results are also displayed in Fig. 3.
where \(< n_{\text{HB}} >\) is the number of water-protein hydrogen bonds, \(k_B\) the Boltzmann constant, \(T\) the temperature of the system under consideration and \(t\) the simulation time along which the cavity is being filled due to water-water and protein-water interactions. The parameters \(f\) and \(E\) would be characteristic of the system under consideration. If no water-protein hydrogen bonds are stabilised (initial state) then \(< n_{\text{HB}} > = 0\), while if the protein has saturated all its hydrogen bonds possibilities with water molecules (final state), then \(< n_{\text{HB}} > = 1\) (normalization to 1). The intermediated scenarios correspond to different number of protein-water hydrogen bonds. It can be noticed that the result \(n_{\text{HB}} > = 0.5\) can be obtained giving to the variable \(t\) in Eq. (1), the particular value \(t_{1/2} = E_{\text{ff}} \left[ \text{A} \right]\) (please mind that \(\text{A}\) is the number of water-protein hydrogen bonds, \(E_{\text{ff}}\)). Thus the time \(t_{1/2}\) increases as \(E\) goes up and \(f\) decreases.

The increase of \(< n_{\text{HB}} >\) with time for nanocavities with initial radius \(r = 2.5, 3.0, 3.5, 4.0\) and 4.5 nm, has been fitted using equation (1) and the free software Gnuplot. The fits are shown in Fig. 3. The corresponding fitted parameters \(E_{\text{ff}}\) and \(f_{\text{sat}}\) are presented in Table 2 while the linear dependence with the radius of the cavity is depicted in Figs. 6 and 7. As an example, \(f_{\text{sat}} = 12.98 \pm 0.35 \times 10^{-22}\) J ps\(^{-1}\) and \(E_{\text{ff}} = 2.41 \pm 0.17 \times 10^{-20}\) J for \(r = 3.0\) nm. The values of \(t_{1/2}\) corresponding to the various cavities are also shown in Table 2.

Although \(f_{\text{sat}}\) keeps constant, \(E_{\text{ff}}\) clearly increases with \(r\) and therefore \(t_{1/2}\) also goes up with the radius of the cavity.

In addition, let us consider the case in which the initial state of the model corresponds to the instant previous to the formation of the first protein-water hydrogen bond and the final state to the establishment of the last one. The interest of that particular case will be discussed in Section 4.3. Thus the temporal evolution of \(< n_{\text{HB}} >\) restricted to the particular time interval \(t_{\text{sat}} - t_i\), has been also fitted to equation (1). In fact, that temporal evolution corresponds to a zoom of Fig. 3 in the temporal range \(t_{\text{sat}} - t_i\). As a function of the radius of the nanocavity \(r\), the fitted parameters \(f_{\text{sat}-t_i}\) and \(E_{\text{sat}-t_i}\) are shown in Table 3.

### Table 2

<table>
<thead>
<tr>
<th>Type of water</th>
<th>Nanocavity Radius (nm)</th>
<th>(f_{\text{sat}} \times 10^{22}) (J ps(^{-1}))</th>
<th>(E_{\text{sat}} \times 10^{20}) (J)</th>
<th>(t_{1/2} = E_{\text{sat}} / f_{\text{sat}}) (ps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>referent</td>
<td>2.5</td>
<td>12.98 ± 0.35</td>
<td>1.32 ± 0.16</td>
<td>10.2 ± 0.9</td>
</tr>
<tr>
<td>referent</td>
<td>3.0</td>
<td>12.98 ± 0.35</td>
<td>2.41 ± 0.17</td>
<td>18.6 ± 1.3</td>
</tr>
<tr>
<td>referent</td>
<td>3.5</td>
<td>12.50 ± 0.35</td>
<td>3.22 ± 0.18</td>
<td>25.8 ± 1.6</td>
</tr>
<tr>
<td>referent</td>
<td>4.0</td>
<td>12.78 ± 0.35</td>
<td>3.81 ± 0.19</td>
<td>29.8 ± 1.8</td>
</tr>
<tr>
<td>referent</td>
<td>4.5</td>
<td>13.03 ± 0.35</td>
<td>4.78 ± 0.20</td>
<td>36.7 ± 2.7</td>
</tr>
<tr>
<td>physiologic</td>
<td>2.5</td>
<td>2.90 ± 0.20</td>
<td>1.01 ± 0.16</td>
<td>34.8 ± 2.7</td>
</tr>
<tr>
<td>physiologic</td>
<td>3.0</td>
<td>2.62 ± 0.20</td>
<td>2.10 ± 0.17</td>
<td>80.2 ± 6.2</td>
</tr>
<tr>
<td>physiologic</td>
<td>3.5</td>
<td>2.69 ± 0.20</td>
<td>3.21 ± 0.18</td>
<td>119.3 ± 6.9</td>
</tr>
<tr>
<td>physiologic</td>
<td>4.0</td>
<td>2.77 ± 0.20</td>
<td>4.00 ± 0.19</td>
<td>144.4 ± 7.4</td>
</tr>
<tr>
<td>physiologic</td>
<td>4.5</td>
<td>2.50 ± 0.20</td>
<td>5.17 ± 0.20</td>
<td>206.8 ± 10.5</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Type of water</th>
<th>Nanocavity Radius (nm)</th>
<th>(f_{\text{sat}-t_i} \times 10^{22}) (J ps(^{-1}))</th>
<th>(E_{\text{sat}-t_i}) (J)</th>
<th>(t_{1/2} = E_{\text{sat}-t_i} / f_{\text{sat}-t_i}) (ps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>referent</td>
<td>2.5</td>
<td>12.74 ± 0.39</td>
<td>10.2 ± 0.9</td>
<td></td>
</tr>
<tr>
<td>referent</td>
<td>3.0</td>
<td>12.74 ± 0.39</td>
<td>18.6 ± 1.3</td>
<td></td>
</tr>
<tr>
<td>referent</td>
<td>3.5</td>
<td>12.74 ± 0.39</td>
<td>25.8 ± 1.6</td>
<td></td>
</tr>
<tr>
<td>referent</td>
<td>4.0</td>
<td>12.74 ± 0.39</td>
<td>29.8 ± 1.8</td>
<td></td>
</tr>
<tr>
<td>referent</td>
<td>4.5</td>
<td>12.74 ± 0.39</td>
<td>36.7 ± 2.7</td>
<td></td>
</tr>
<tr>
<td>physiologic</td>
<td>2.5</td>
<td>12.74 ± 0.39</td>
<td>34.8 ± 2.7</td>
<td></td>
</tr>
<tr>
<td>physiologic</td>
<td>3.0</td>
<td>12.74 ± 0.39</td>
<td>80.2 ± 6.2</td>
<td></td>
</tr>
<tr>
<td>physiologic</td>
<td>3.5</td>
<td>12.74 ± 0.39</td>
<td>119.3 ± 6.9</td>
<td></td>
</tr>
<tr>
<td>physiologic</td>
<td>4.0</td>
<td>12.74 ± 0.39</td>
<td>144.4 ± 7.4</td>
<td></td>
</tr>
<tr>
<td>physiologic</td>
<td>4.5</td>
<td>12.74 ± 0.39</td>
<td>206.8 ± 10.5</td>
<td></td>
</tr>
</tbody>
</table>

The corresponding fitted parameters \(f_{\text{sat}}\) with \(r\), \(t_{\text{sat}} - t_i\) represents the time associated to the inner protein hydration shell formation and \(E_{\text{sat}-t_i}\) the corresponding energy. More information in Sections 3.1, 3.2, 4.2 and 4.3.
< $n_{wat} > = \frac{e^{\frac{\Delta E}{k_B T}} - e^{\frac{E_{wm}}{k_B T}}}{e^{\frac{\Delta E}{k_B T}} + e^{\frac{E_{wm}}{k_B T}}} \ \ \ \ \ \ \ \ (2)$

The best fit of the corresponding data to Eq. (2) is shown in Fig. 1 and the corresponding fitted parameters, $f_{wm}$ and $E_{wm}$, are presented in Table 4.

3.3. Nanoscopic cavity in water (with and without protein inside): Physiologic water

In the previous sections the dynamics of a nanocavity created in the network of water molecules has been analysed. That cavity may, or not, include a biomolecule. The diffusion coefficient of bulk water at ambient temperature and pressure was used ($D_{ref}^{footnote}$ = 2.40 μm²/μs [47] see Table 1). However when considering animal or human bodies the aqueous scenario is generally more crowded due to the presence of intrinsic biomolecules, metabolites or extrinsic entities like drugs etc. [6,10]. Crowded conditions are expected to slow down the dynamics of water. In this section the analysis of the previous paragraphs is performed considering that the water solution is one corresponding to a physiologic milieu. Thus in-vivo Magnetic Resonance Imaging (MRI) experiments performed in humans, monkeys and rats show that the diffusion coefficient of water significantly decreases when compared with the value of isolated water (2.40 μm²/μs). Thus the diffusion coefficient of blood in humans amounts 1.4 ± 0.4 μm²/μs [50]. The concrete value depends of various factors as blood composition, hematocrit (volume percentage of red blood cells in blood), hemoglobin concentration etc. [50]. That value agrees with the diffusion coefficient of water reported in rat skeletal muscle, m²/ms [51,52]. However in the case of the human, monkey or rat brains the diffusion coefficient of water shows a smaller value, 0.8 ± 0.2 μm²/μs [53–56]. Gordon et al. reported that these results for brains would be underestimated (nearly 40% lower) due to particular experimental conditions [57]. Therefore a corrected value of around 1.3 ± 0.2 μm²/μs could be estimated in the case of water in brain.

With all that information in mind a value for the water diffusion coefficient $D_{phys}^{footnote}$ = 1.5 ± 0.5 μm²/μs seems reasonable to simulate the translation of water in human and related animal scenarios. Regarding the molecular dynamics simulations, that diffusion coefficient can be obtained by tuning the constant time $\tau_{aut}$ in the Langevin equation of motion (see Methods/simulation 1) [48,49]. Thus various simulations of bulk water were performed to tune the $\tau_{aut}$ parameter. A $D_{phys}^{footnote}$ = 1.52 ± 0.02 μm²/μs was obtained when using $\tau_{aut}$ = 0.112 ± 0.002 ps (Section 2/simulation 8). Therefore the simulations reported below in the present section (simulations 5 to 8) were performed using that value $\tau_{aut}$ = 0.112 ps (see Table 1). In addition the experiments reported in the previous paragraph, were carried out at 37 ± 1C and therefore the simulations considered in this section (simulations 5 to 8) have been performed at that temperature, that is 310 K.

Simulation 5 is the equivalent of simulation 1 (temporal evolution of protein in water). Simulation 6 is the equivalent of simulation 2 (temporal evolution of the snapshot of depart: an empty nanocavity, with the shape of the protein, located in the middle of the water molecules simulation box). Simulations 7a, 7b, 7c, 7d and 7e are the equivalent of simulations 3a, 3b, 3c, 3d and 3e (temporal evolution of the snapshot of depart: the protein centred in a cavity with an initial radius $r$ = 2.5, 3.0, 3.5, 4.0 or 4.5 nm, respectively; the solvent surrounds such cavity). Finally, simulation 8 is the equivalent of simulation 4 (temporal evolution of bulk water). Computer simulations 8 and 4 are taken as references for the rest of simulations.

The temporal evolution of the number of water molecules entering the cavity with the shape of the protein (simulation 6) is depicted in Fig. 1 while that corresponding to the number of hydrogen bonds between the protein and the solvent molecules entering the cavities with five different diameters, is shown in Fig. 8 (simulations 7a to 7e). These 1 + 5 sets of data have been fitted to Eqs. (2) and (1), respectively (see Figs. 1 and 8 for the fits). The parameters obtained from those fits are reflected in Tables 2–4. The evolution of $E_{sat} E_{sat}$, $E_{sat}$, $E_{sat}$ – $t$ and $E_{sat}$–$a$ are presented in Tables 3.1 and 3.2 for their definition) with the radius of the nanocavity is depicted in Figs. 4, 6, 7 and 9. Although $E_{sat}$ and $E_{sat}$–$a$ vary with $r$, $E_{sat}$ and $E_{sat}$–$a$ are independent of the radius of the cavity. It can be also concluded that $E_{sat}$–$a$ = $E_{sat}$ since $E_{sat}$ = 2.70 ± 0.20 x 10$^{-22}$ J ps$^{-1}$ (mean value) and $E_{sat}$–$a$ = 2.71 ± 0.27 x 10$^{-22}$ J ps$^{-1}$ (mean value). As in the case of the referent water (Section 3.1) and for a given $r$ value, $E_{sat}$ < $E_{sat}$ and such difference increases with $r$. Moreover, $E_{sat}$ ($r$ = 2.5 nm) = $E_{sat}$ ($r$ = 2.5 nm) since $t_i$ ($r$ = 2.5 nm) ≥ 0.

### 4. Discussion

In the following paragraphs the results presented in the previous sections are discussed. The analysis of Section 3 indicates that two different processes can be described when considering the scenario with the protein located inside the cavity. The first one corresponds to the movement of the water molecules through the cavity until the first protein-water hydrogen is formed. That process takes a $t_i$ time. As expected $t_i$ increases with the radius of the cavity (see Figs. 3 and 8). The second one, $t_{sat}$ – $t_i$, corresponds to the time interval from the first protein–water hydrogen bond is formed to the last one takes place (see Section 3.1). That time can be associated to the inner protein hydration shell formation. The sum of both $t_i$ and $t_{sat}$ – $t_i$ provides $t_{sat}$, the total time necessary

### Table 4

Results from applying Eq. (2) to the dynamics of an empty nanoscopic cavity created after pulling out the protein burstor from its place in a water solution. As the cavity is being filled with the solvent, the magnitude < $n_{wat}$ > provides the number of water molecules inside the cavity (normalized to 1). The parameters $f_{wm}$ and $E_{wm}$ are obtained from fitting the data to Eq. (2). The fits are shown in Fig. 1. The relevant characteristics of referent and physiologic water were already detailed in Table 1.

<table>
<thead>
<tr>
<th>Type of water</th>
<th>$f_{wm}$ x10²² (J ps⁻¹)</th>
<th>$E_{wm}$ x10²⁰ (J)</th>
<th>$t_{\tau/2}$ = $E_{wm}$/fwm (ps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>referent</td>
<td>15.60 ± 0.40</td>
<td>2.01 ± 0.16</td>
<td>12.9 ± 1.0</td>
</tr>
<tr>
<td>physiologic</td>
<td>5.57 ± 0.20</td>
<td>2.21 ± 0.16</td>
<td>39.7 ± 2.9</td>
</tr>
</tbody>
</table>

### Table 3

Results from applying Eq. (1) to the dynamics of a nanoscopic cavity, located in a water solution, with protein barstar in the centre. The examined time interval corresponds to that associated to the inner protein hydration shell formation, and flows from the formation of the first protein – water hydrogen bond to the last one. That time interval corresponds to a zoom of Figs. 3 and 8 in the temporal range $t_{sat}$ – $t_i$. The parameters $f_{sat}$–$a$ and $f_{sat}$ are obtained by fitting the data to that equation (Section 3.2). The relevant characteristics of referent and physiologic water are detailed in Table 1.

<table>
<thead>
<tr>
<th>Type of water</th>
<th>Nanocavity Radius (nm)</th>
<th>$f_{sat}$ x10²² (J ps⁻¹)</th>
<th>$E_{sat}$ x10²⁰ (J)</th>
</tr>
</thead>
<tbody>
<tr>
<td>referent</td>
<td>2.5</td>
<td>13.05 ± 0.39</td>
<td>1.21 ± 0.16</td>
</tr>
<tr>
<td>referent</td>
<td>3.0</td>
<td>12.78 ± 0.39</td>
<td>1.61 ± 0.20</td>
</tr>
<tr>
<td>referent</td>
<td>3.5</td>
<td>12.31 ± 0.39</td>
<td>1.65 ± 0.25</td>
</tr>
<tr>
<td>referent</td>
<td>4.0</td>
<td>12.56 ± 0.39</td>
<td>1.65 ± 0.27</td>
</tr>
<tr>
<td>referent</td>
<td>4.5</td>
<td>13.00 ± 0.39</td>
<td>1.73 ± 0.30</td>
</tr>
<tr>
<td>physiologic</td>
<td>2.5</td>
<td>3.01 ± 0.27</td>
<td>1.00 ± 0.16</td>
</tr>
<tr>
<td>physiologic</td>
<td>3.0</td>
<td>2.60 ± 0.27</td>
<td>1.70 ± 0.20</td>
</tr>
<tr>
<td>physiologic</td>
<td>3.5</td>
<td>2.69 ± 0.27</td>
<td>2.08 ± 0.25</td>
</tr>
<tr>
<td>physiologic</td>
<td>4.0</td>
<td>2.76 ± 0.27</td>
<td>2.05 ± 0.27</td>
</tr>
<tr>
<td>physiologic</td>
<td>4.5</td>
<td>2.49 ± 0.27</td>
<td>2.28 ± 0.30</td>
</tr>
</tbody>
</table>
to complete both processes. In the following, the two scenarios associated to times \( t_{\text{sat}} \) and \( t_{\text{sat}} - t_i \) are separately discussed.

### 4.1. Global scenario: Saturation time \( t_{\text{sat}} \)

Let us consider the magnitude \( < n_{\text{HB}} > \) which represents the percentage of hydrogen bonds (normalized to 1) between the solvent and the protein during the process of filling the nanoscopic cavity with water. When changing from referent to physiologic water, the time necessary to reach the saturation \( t_{\text{sat}} \) notably increases (see Fig. 4). An equivalent scenario can be contemplated when considering the magnitude \( < n_{\text{WM}} > \) which represents the percentage of water molecules (normalized to 1) entering the empty nanoscopic cavity which exhibits the shape of the protein (see Fig. 1).

In the cases under study, times of the order of picoseconds are necessary to saturate the number of protein-water hydrogen bonds, or to completely fill the empty cavity with water molecules (Fig. 4). The relevant question is if that time can be associated to significant times concerning the dynamics and bioactivity of proteins. Main considerations are given in the two following paragraphs.

Protein function relies on structural dynamics, like motions of side chains, with time scales ranging from femtoseconds to beyond seconds [58–61]. In particular, how an organism benefits from motions of a protein in the range of picoseconds can not be promptly evident. Nevertheless, it is accepted that release of oxygen to tissues is essential for survival. As an example, myoglobin is a protein that muscle cells of vertebrates use to accelerate oxygen diffusion for times of intense respiration. The oxygen-binding site of that protein is occluded, thereby pointing to necessary movement during the course of the protein’s activity [62]. Brunori et al have concluded that the overall \( \text{O}_2 \) dissociation rate is governed not only by the iron-ligand bonding but also by fluctuations of internal side chains of myoglobin in the range of picoseconds [62]. Those fluctuations can open or close the access to the referred oxygen-binding site of the protein. Therefore picosecond dynamics has a survival character for organisms containing hemoproteins as myoglobin.

On the other hand the structure, dynamics and function of biomolecules are fully dependent of the changes taking place in their hydration layer. Hydration water, including protein-water hydrogen bonds, is fundamental for preserving the bioactivity of proteins [63]. A full structural relaxation of the biomolecule demands relaxation of the protein-water hydrogen bonds by way of solvent movement [64]. The large magnitude motions of biomolecules that are needed for developing their biological activity, are stimulated by diffusion of the solvent molecules on the protein surface [65]. In fact restraining that water translational movement is equivalent to dehydrate the biomolecule [64]. Disabling their biochemistry, bacterial spores can remain alive for a long time under hard circumstances. The nucleus of spores accommodates very few water molecules (biomolecules immobilization) [66].

As a result of the two precedent paragraphs, if a protein falls in a nanocavity taking a time in the range of picoseconds to rebuit its hydrogen bonds networks with the solvent, and if the dynamics of that protein develops in that temporal range, such a dynamics, and therefore the corresponding bioactivity, will be highly compromised during that period of time.

In a physiologic milieu protein isolation times on the range of nanoseconds would be reached by enlarging the radius of the cavity to one ten of nanometers (see Fig. 4). To give an idea what that size means in biological systems, please mind that the dimensions of globular proteins range from 2 to 10 nm, antibodies: 10 nm, viruses: 100 nm, bacteria: \( 10^3 \) nm, cancer cells: \( 10^4-10^5 \) nm etc [17,67]. In addition typical inorganic nanoparticles range between 1 and 100 nm; if they are functionalised with ligands and biomolecules for diagnostic and therapeutic applications, their size can notably increase [67,68]. As an example, dipolar relaxtion on the nanosecond time scale has been described to occur for membrane proteins from fluorescence spectroscopy experiments [58]. A significant number of computational and experimental analysis have shown that the structural dynamics of proteins often play a significant role in their functionality [60]. Thus the relaxation dynamics, and therefore the corresponding biofunction, of one of those membrane proteins located inside a cavity with associated protein isolation times in the range of nanoseconds will be seriously damaged during that period of time.

### 4.2. Two states model

The dynamics of those systems has been modelled using a two states model. Such analysis shows that dynamics is governed by two parameters, \( f \) and \( E \) (Eqs. (1) and (2); Section 3.2). Thus Fig. 7
indicates that the parameter $E_{\text{tsat}}$ does not basically change with the type of water (referent or physiologic), and therefore with the diffusion coefficient of the solvent, however it clearly depends on the initial radius of the cavity. An equivalent behaviour can be observed in the case of $E_{\text{tsat}-\text{ti}}$ (Table 3) and $E_{\text{wm}}$ (Table 4). An inspection of Eqs. (1) and (2) indicates that for large $E$ values, in any case $E \gg k_B T$, $e^E$ will be also very large and therefore $<n_{\text{sam}}>$ and $<n_{\text{tot}}>$ will be nearly zero. No protein-water hydrogen bonds or no water molecules inside the cavity as time evolves. The system keeps essentially located in the initial state. The final state is not reached since $E$ is too large. Moreover the analysis of those equations shows that the physical dimensions of $E$ correspond to those of an energy (J). As a consequence of those considerations, the parameter $E$ might be related to the difference in energy in the two states of the model referred in Section 3.2. Thus $E_{\text{sat}}$ would correspond to the difference in energy between the initial state associated to the falling of the biomolecule into the cavity and the final state corresponding to the saturation of the protein-water hydrogen bonds. $E_{\text{sat}}$ would be the energy associated to the protein isolation time $t_{\text{tsat}}$. In parallel $E_{\text{sat}-\text{ti}}$ would correspond to the difference in energy between the initial state associated to the instant previous to the formation of the first protein-water hydrogen bond and the final state corresponding to the establishment of the last one (time interval $t_{\text{tsat}} - t_i$). $E_{\text{sat}-\text{ti}}$ would be the energy associated to the formation of the inner protein hydration shell (for additional information see Section 4.3, third paragraph). In good agreement with that thesis (that is the parameter $E$ relates to the difference in energy between the two states considered), $E_{\text{tsat}-\text{ti}} < E_{\text{tsat}}$ and that difference increases with $r$ (see Tables 2 and 3).

On the contrary, the parameter $f$ keeps constant as the radius of the cavity changes (see Fig. 6). However $f$ is nearly smaller for physiologic than for referent water, it decreases with the diffusion coefficient of the solvent. Those two statements are true even if different cases of the model are considered. Thus Tables 2 and 3 indicate that $f_{\text{sat}} = f_{\text{sat}-\text{ti}} = 12.80 \pm 0.39 \times 10^{-22}$ J ps$^{-1}$ for referent water but $f_{\text{tsat}} = f_{\text{tsat}-\text{ti}} = 2.70 \pm 0.27 \times 10^{-22}$ J ps$^{-1}$ for physiologic water. When analysing Eqs. (1) and (2), it can be observed that for very small $f$ values, in any case $f \ll k_B T$, $<n_{\text{sam}}>$ and $<n_{\text{tot}}>$ will be nearly zero since $e^{-E/f} = e^E$ will be also very small. As time evolves, the system is basically located in the initial state. Those equations also evidence that the physical dimensions of $f$ correspond to those of a flux of energy (J ps$^{-1}$). Thus $f$ might be related to the flux of energy during the process of stablishing protein-water hydrogen bonds (filling the cavity with solvent in the case of $<n_{\text{sam}}>$). If $f$ is too small, the water molecules can not move in an effective way and the time necessary to reach the final state of the model will be infinite.

4.3. Dynamics of the protein hydration shell formation: Time interval $t_{\text{sat}} - t_i$

The time interval from the first protein-water hydrogen bond is formed to the last one takes place, that is the time interval for inner protein hydration layer formation, corresponds to $t_{\text{sat}} - t_i$ (see Section 3.1). At first glance, and for a given type of water, the time $t_{\text{sat}} - t_i$ should not significantly change with the radius of the cavity. However Figs. 5 and 9 indicate that this is not the case. Thus as moving from $r = 2.5$ to 4.5 nm, $t_{\text{sat}} - t_i$ nearly doubles its value (physiologic water). In the following paragraphs a possible explanation is discussed.

The shell of water molecules in close interaction with a protein surface is named protein hydration layer or biological water. The structure and dynamics of that particular water differs from that of bulk water which is situated at higher distances from the protein (see Section 1 and references therein). As an example, the number of water-water hydrogen bonds that the protein hydration shell exhibits diverges from that of isolated water [49]. The reason of that difference: the dynamics of the biological water is forced to be the result of a compromise between the dynamics of the biomolecule and that of bulk water. The protein hydration layer is in touch with both ‘scenarios’ and tries to connect them [49]. Biological water acts as an intermediary entity.

This special water is then located closer to the protein than bulk water is. Therefore the network of solvent molecules that closely surrounds the cavity with radius $r = 2.5$ nm (simulations 3a and 7a), should resemble that of the biological water of the protein in a higher degree that the network of water molecules that are situated close around the cavity with radius $r = 4.5$ nm (simulations 3e and 7e). In the way that network of water molecules resembles that of the biological water corresponding to the protein, the dynamics of those water molecules would connect more rapidly with the dynamics of the biomolecule and the protein-water hydrogen bonds would quickly form. That implies a smaller $t_{\text{sat}} - t_i$ time, as Figs. 5 and 9 confirm to occur. The decrease of the slope of the $E_{\text{sat}} - t_i$ curve as $r$ increases (Figs. 5 and 9), would suggest a decrease of the biological character of the solvent as $r$ goes up, in agreement with our thesis. That thesis is also supported by the reduced decrease of $E_{\text{sat}-\text{ti}}$, the energy corresponding to the inner protein hydration shell formation, with $r$ (see Figs. 5 and 9, and Section 4.2). At higher $r$ values, the resemblance of the water molecules approaches that of bulk water and $E_{\text{sat}-\text{ti}}$ tends to be constant.

Although not the same, the scenario would remind that of antibodies originated by a close contact of the immune system with the biomolecules integrating vaccines against, for example, viruses. After applying the vaccine, if an infection occurs, the created antibodies will detect the target biomolecules integrating the virus, recognise them and proceed in consequence. In our case the antibodies would be the network of solvent molecules resembling that of the protein hydration water. That special network can be created by a previous close contact with the protein. Thus in our simulations a network of solvent molecules resembling that of the biological water has been built by removing a smaller amount of solvent around the biomolecule to form the initial cavity; that corresponds to a smaller radius of the hollow. If a large quantity of solvent was removed, the remaining network of water molecules will be mostly bulky in character since they were situated far away from the protein.

5. Main conclusions

This research explores the particular scenarios arising from nanoscopic cavities, placed in an aqueous solution, that includes proteins and, as a consequence, the formation of the corresponding protein hydration shell.

The dynamics of those systems, with (or without) a biomolecule inside the nanocavity, has been modelled using a two states model. The filling of that cavity with water molecules of the adjacent solvent and the formation of protein–water hydrogen bonds is followed as a function of time. Such analysis indicates that the dynamics is governed by two parameters $f$ and $E$ that can be related, respectively, to the nature of the solvent and to the initial radius of the cavity $r$. The parameter $E$ (J) might be also related to the difference in energy between the initial and the final state of the model (no protein–water hydrogen bonds and saturation of those hydrogen bonds; cavity empty or cavity completely filled with water molecules) while the parameter $f$ (J ps$^{-1}$) could be associated with the corresponding flux of energy.

In addition the dynamics of formation of the inner protein hydration shell has been followed using an atomic view. The time
interval associated to that formation, \( t_{\text{sat}} = t \), should not significantly change with the radius of the nanocavity \( r \). However, and contrary to what might be expected, \( t_{\text{sat}} = t \) clearly changes with \( r \). In the way the network of water molecules approaching the bio-molecule resembles that of the biological water corresponding to the protein, the dynamics of those water molecules would connect more rapidly with the dynamics of the biomolecule and the protein-water hydrogen bonds would quickly form. The result is a smaller \( t_{\text{sat}} = t \) time. That resemblance increases as those water molecules have previously formed part of the biological water of the protein by, for example, a previous close contact with it. That thesis is also supported by the related decrease of the energy corresponding to the inner protein hydration shell formation, as \( r \) goes down.

On the other hand and in the case of the protein barstar (radius of gyration of 1.17 nm) located in the centre of cavities with radius from 2.5 to 4.5 nm, and solvent diffusion coefficients for bulk and physiologic water (2.4 and 1.5 \( \mu \text{m/s} \), respectively), the total time \( t_{\text{sat}} \), required to reach the final state corresponding to the saturation scenario, counted since the instant the protein falls down into the empty cavity, is found to be of the order of tens–hundreds of picoseconds.

Thus if a protein falls down into one of those empty cavities, presenting a volume larger than the volume of the protein, this work shows that such protein will take a significant time, \( t_{\text{sat}} \), to recover its entire hydration layer and to rebuild its whole network of hydrogen bonds with the water molecules. Since i) protein hydration shell is essential for maintaining the functional structure, dynamics and biological activity of biomolecules, ii) the time scales regarding the bioactivity of proteins range from femtoseconds to seconds and iii) the present work reports the existence of protein isolation times (\( t_{\text{sat}} \)) in the range of picoseconds, it can be concluded that during that time \( t_{\text{sat}} \), the dynamics of the biomolecule and the associated bioactivity would be seriously compromised.

The present contribution attempts to focus the attention to novel scenarios concerning the field of the dynamics of nanoscopic cavities in aqueous solutions that include biomolecules, and to protein hydration shell formation, aiming to act as a guide for future experimental works on those research areas.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

This work was funded by the European Regional Development Fund (EU) together with the Government of Spain, under common Project FIS2017-87519-P. Financial support from the Government of Aragón (Aragón Research Group 'Fenol', E36_20R, Spain) and the European Union (Feder 2014-2020, Building Europe from the European Union Fund (EU) together with the Government of Spain, under common Project FIS2017-87519-P. Financial support from the Government of Aragón (Aragón Research Group 'Fenol', E36_20R, Spain) and the European Union (Feder 2014-2020, Building Europe from the
W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and testing of the 13C-M.
M.C. Morón Journal of Molecular Liquids 337 (2021) 116584
[34] A.C. Bhasikuttan, J. Mohanty, Detection, inhibition and disintegration of [36] P. Dutta, M. Botlani, S. Varma, Water Dynamics at Protein/C0
[38] M. J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the GROMACS