EPJ E
Soft Matter and Biological Physics
EPJ.org
your physics journal

Eur. Phys. J. E (2013) **36**: 62 DOI 10.1140/epje/i2013-13062-7

Protein packing defects "heat up" interfacial water

María Belén Sierra, Sebastián R. Accordino, J. Ariel Rodriguez-Fris, Marcela A. Morini, Gustavo A. Appignanesi and Ariel Fernández Stigliano

edp sciences Società Italiana di Fisica Springer

Eur. Phys. J. E (2013) **36**: 62
DOI 10.1140/epje/i2013-13062-7

THE EUROPEAN PHYSICAL JOURNAL E

Regular Article

Protein packing defects "heat up" interfacial water

María Belén Sierra[1,a], Sebastián R. Accordino[1], J. Ariel Rodriguez-Fris[1], Marcela A. Morini[1], Gustavo A. Appignanesi[1,b], and Ariel Fernández Stigliano[2,3]

[1] Sección Fisicoquímica, INQUISUR-UNS-CONICET-Departamento de Química, Universidad Nacional del Sur, Avda. Alem 1253, 8000 Bahía Blanca, Argentina
[2] Instituto Argentino de Matemática "Alberto P. Calderón", CONICET (National Research Council), Saavedra 15 Buenos Aires 1083, Argentina
[3] Collegium Basilea, Institute for Advanced Study, Hochstrasse 51, Basel CH4053, Switzerland

Received 24 November 2012 and Received in final form 5 April 2013
Published online: 25 June 2013 –

Abstract. Ligands must displace water molecules from their corresponding protein surface binding site during association. Thus, protein binding sites are expected to be surrounded by non-tightly-bound, easily removable water molecules. In turn, the existence of packing defects at protein binding sites has been also established. At such structural motifs, named dehydrons, the protein backbone is exposed to the solvent since the intramolecular interactions are incompletely wrapped by non-polar groups. Hence, dehydrons are sticky since they depend on additional intermolecular wrapping in order to properly protect the structure from water attack. Thus, a picture of protein binding is emerging wherein binding sites should be both dehydrons rich and surrounded by easily removable water. In this work we shall indeed confirm such a link between structure and dynamics by showing the existence of a firm correlation between the degree of underwrapping of the protein chain and the mobility of the corresponding hydration water molecules. In other words, we shall show that protein packing defects promote their local dehydration, thus producing a region of "hot" interfacial water which might be easily removed by a ligand upon association.

1 Introduction

The hydration properties of protein binding sites have been suggested to play a main role in the binding of ligands and in protein-protein association [1–14]. From one side, ligands are expected to displace hydration water molecules from their protein binding site [1–3] and the replacement of so-called "unfavorable" waters by groups of the ligand complementary to the protein surface has been established as a principal driving force for binding [2,3]. In fact, this description has been shown to hold valid for a significant fraction of receptors of pharmaceutical interest [2,3,15]. Even in some cases, a portion of the receptor active site is so unfavorable for water molecules that it tends to remain practically dry [3]. Thus, the estimation of the free energy contribution involved in the displacement of quasilocalized water molecules with unfavorable free energies in the receptor active site constitutes an issue of great interest in computational structure-based drug design [2,3]. Within this same philosophy, a recent study [16] of fragment clustering of diverse organic probes on hen egg white lysozyme showed that several regions of the protein were targeted by fragment clusters. However, by combining this strategy with water exclusion (superimposing the map of water molecules tightly bound to the protein) the experimentally known binding site, or hot spot, could be properly predicted since all other fragment clustering sites coincided with the tightly bound water molecules. This is so since such water molecules would exclude the ligands from the protein surface by blocking their binding sites. Thus, a picture of protein binding with regions of easily removable (non-tightly-bound) water molecules at small-molecule binding sites or protein-protein interaction hot spots is emerging.

On the other hand, the role of structural packing defects characterized by regions of the protein backbone exposed to the solvent has been extensively discussed in the literature [6–14,17]. Soluble proteins tend to protect their backbone hydrogen bonds, HBs, from water attack by wrapping them intramolecularly with side-chain nonpolar groups. In turn, uncompletely wrapped HBs (or dehydrons [5–14]) are sticky since they promote further *intermolecular* removal of surrounding water and have

[a] e-mail: mbsierra@uns.edu.ar
[b] e-mail: appignan@criba.edu.ar

been demonstrated to be central for the protein-protein association phenomena upon which most biological functions rely [6–14]. In fact, a decomposition of the protein-protein complex interface into a web of wrapping interactions enabled us to successfully predict the hot spots reported by alanine-scanning experimental studies for a set of protein-protein complexes [13]. Additionally, we have also shown that drugs disruptive of protein-protein interfaces tend to mimic the wrapping behavior of the protein they replace [14]. In summary, soluble proteins need to wrap their HBs but also need to retain certain level of vulnerability or backbone exposure in order to interact with other proteins or ligands. This paradox of reconciling stability with interactivity (which implies to reformulate the previous question of "how to keep dry in water?" [18] towards "how to do it without sacrificing interactivity?") implies that dehydrons must promote their dehydration so that water around them should be more labile [6, 8, 9, 12].

The above-expounded facts imply that binding hot spots should be both dehydron-rich and hydrated by easily removable water. Thus, the aim of the present work is to seek for a correlation between the degree of wrapping of protein HBs and the removability of their hydration water. To this end, instead of using complicated methods to calculate the free energy of water removal (based on approximations that might involve a significant degree of inexactitude) we shall simply resort to dynamically based calculations by measuring local translational diffusivities for water molecules within the protein hydration shell. More specifically, we shall compute the mean squared displacement of the water molecules within the desolvation domains of the protein HBs, both for dehydrons and non-dehydrons. Such mobility calculations based on molecular dynamics (MD) simulations, will thus characterize the removability of the water molecules and will enable us to find "unfavorable" water molecules around the protein surface. In view of the subnanoscale ruggedness of the protein surface [12], this method is more suited than a study of water residence times in fixed regions around the superficial residues since at variance from the situation in a flat surface, there would be a whole distribution of minimum distances of water molecules to the protein. Thus, a calculation of residence times within certain sphere around α carbons or HBs would give misleading results since in some regions the water molecules would be close to the center of the calculation sphere while in others it would already be much closer to its boundary.

Our work will ascertain the existence of a clear correlation between the degree of underwrapping of the HB and the local mobility of the hydration water molecules. In turn, we shall correlate water mobility with the degree of local structural disorder of the protein chain. We shall also map the distribution of water mobility around the surface of a case study: the central DNA-binding core domain of protein p53. Thus, we shall find regions of easily removable water, which we shall find to correspond to dehydron clusters and experimentally relevant binding hot spots.

2 Methods

2.1 Model systems

In this work we studied the behavior of the hydration layers of a set of complete (without missing residues) proteins without ligands (pdb IDs: 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, 2GEQ, 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). These 62 pdbs were chosen at random, with an average residue number of 155 and standard deviation 151. The water molecules were modeled by the TIP3P model [19, 20] as explicit solvent with AMBER versions 10 and 11 [21], using periodic boundary conditions and a simulation box that extended more than 14 Å away from any protein atom. The equilibration and MD simulations were carried out according to the AMBER official tutorial (`http://ambermd.org/tutorials/basic/tutorial1/section5.htm`) in four stages and using the same parameters. We note that we used a cutoff interaction of 8 Å instead of the suggested 10 Å, and a total simulation time of 4 ns for stage 4. Basically, in the first stage the energy of the system is minimized holding the protein fixed, in the second stage the minimization is over the entire system, in the 3rd stage the system is heated to the desired temperature (300 K for this work) at constant volume (density around 1.0 kg/dm^3), with Langevin dynamics controlling the temperature and with a slight restraint in the protein, and in the 4th stage the temperature is maintained by Langevin dynamics keeping the pressure at an average value of 1 atm. SHAKE was employed in stages 3 and 4. The force field used in the simulation was ff99SB. Equilibration was tested by monitoring the behavior of thermodynamical properties like temperature, pressure and energy oscillations.

2.2 Quantifying wrapping and identifying dehydrons

To prevail in water environments, soluble proteins protect their backbone hydrogen bonds (HBs) from the disruptive effect of water attack by clustering non-polar residues around them [4–14, 17]. This exclusion of surrounding water, or wrapping effect, also enhances the electrostatic contribution by modulating the local dielectric (de-screening the partial charges) and thus stabilizes the HB. In turn, as demonstrated previously, underwrapped interactions are adhesive [6, 7], hence promoters of protein associations because their inherent stability increases upon approach of additional non-polar residues [4–14]. Thus, the integrity of the protein-protein interface in protein complexes becomes extremely reliant on intermolecular cooperativity [4–14]. To complete this description it is necessary to classify pairwise electrostatic interactions and detect underprotected

interactions (UPIs). UPIs that involve HBs are named dehydrons. This structural motif has been extensively discussed in the literature and identified in soluble proteins with pdb-reported structure [6–14]. Thus, the extent of HB protection can be determined directly from atomic coordinates. This parameter indicates the number of three-body correlations engaging the HB and is also known as the wrapping of the bond and denoted ρ. It is given by the number of side-chain carbonaceous non-polar groups (CH_n, $n = 0, 1, 2, 3$, where the carbon atom of these groups is not bonded to an electrophilic atom) contained within a desolvation domain around the HB. Each wrapping non-polar group represents the third body within a three-body correlation involving the HB. This domain is defined as the reunion of two intersecting spheres of fixed radius (∼thickness of three water layers) centered at the α carbons of the residues paired by the HB. In structures of pdb-reported soluble proteins, HBs are protected on average by $\rho = 26.6 \pm 7.5$ side-chain non-polar groups for a desolvation sphere of radius 6 Å [8, 9, 12–14]. Thus, structural deficiencies lie in the tail of the ρ distribution, *i.e.* their microenvironment contains 18 or fewer non-polar groups, so their ρ value is below the mean (=26.6) minus one standard deviation (=7.5). While the statistics on ρ values for HBs vary with the radius, the tails of the distribution remain invariant, thus enabling a robust identification of structural deficiencies [8, 9, 12–14]. Moreover, later on we present a proof of this insensitivity in the choice of the radius (within reasonable bounds). In this work, we studied 62 complete pdbs chosen at random. Therefore, we studied a total of 15681 HBs, either dehydrons or non-dehydrons. We considered a HB when the nitrogen (N) bonded to the α carbon of a residue and the oxygen (O) of the carbonyl bonded to the α carbon of another residue are less than 3.5 Å, and the (minimum) angle between H$\cdots$O and H-N (H is the hydrogen bonded to N) is greater that 140 degrees. Hydrogens, when not found in the pdbs, and missing atoms were added to the protein using AMBER software [22].

2.3 Measure of water mobility around HBs

In order to get an idea of the influence of the protein surface on water dynamics, we show in fig. 1 the mean squared displacement, MSD $= \langle r^2(t)\rangle = \langle[\mathbf{r}_i(t) - \mathbf{r}_i(0)]^2\rangle$, of water molecules close to the proteins studied, where $\mathbf{r}_i(t)$ is the position of the oxygen of water molecule i at time t and $\langle\ldots\rangle$ is the average over all i water molecules. Based on previous results that indicate that only water molecules within the first peak of the water-protein radial distribution (or the surface density plot for model flat surfaces) exhibit a dynamics significantly different from that of the bulk [23–26] we have chosen to display the behavior of the molecules that are closer than 4 Å from the protein surface (minimum distances) at the initial time (time zero). For comparison, we also include the case of water molecules initially far appart from the protein, with bulk-like behavior.

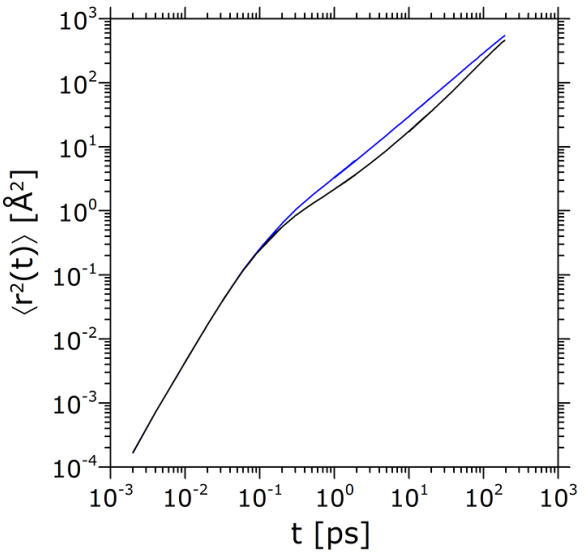

Fig. 1. Mean squared displacement plot for the water molecules close to the different proteins (water molecules whose minimum distance to the protein is lower than 4 Å, black curve) and for water molecules distant from the proteins (more than 14 Å, blue curve).

Direct inspection of fig. 1 reveals that water dynamics close to the protein is slower than bulk dynamics. While the MSD curve for the molecules far from the protein surface displays a behavior typical of bulk water, with a diffusive regime (a region with slope equal to unity following the usual short-time ballistic regime), the water molecules close to the protein evidence a slower, subdiffusive, regime and only tend to bulk behavior at large times. This fact is consistent with previous results [23, 24, 26–28] and holds also if one discriminates the parallel and perpendicular components of the MSD [26]. This behavior of the water molecules close to the protein speaks of the existence of a superficial regime at short times and distances and a bulk-like one which the molecules adopt when they get away from the protein surface. Thus, a calculation of the diffusion constant for the superficial water molecules (as obtained by the Einstein relation and which implies the extrapolation of the MSD at long times) would not be meaningful. To avoid this problem, we decided to adopt as a measure of water mobility the calculation of the MSD value at short times. We considered a timescale of $\phi = 4$ ps (a timescale significantly larger than the end time of the ballistic regime). As can be learnt from fig. 1, at such timescale the superficial water molecules have moved on average one water-water distance, thus providing a reasonable measure of the local diffusivities.

In summary, to assess the mobility of the water molecules within the desolvation domains of HBs we calculated their mean squared displacements within a fixed time interval of length ϕ, defined as $\langle r^2(t)\rangle = \langle[\mathbf{r}_i(t = \phi) - \mathbf{r}_i(t = 0)]^2\rangle$. As indicated, the time interval was chosen as $\phi = 4$ ps, a timescale short enough to be sensitive to the local diffusion of the molecules abandoning the first hydration shells, thus providing a good measure of local translational diffusivities. We applied this study to the hy-

dration shell of the large set of proteins above indicated, by discriminating between the desolvation domains of dehydrons and non-dehydrons. To get good statistics, for each of the 62 proteins we generated 400 different MD runs of length $t = \phi$.

2.4 Description of structural parameters

The description of the structural parameter used has already been provided in a previous work [25]. However, and for the sake of completness, we hereby provide an almost verbatim short description.

Liquid water is known to present several anomalies which become more prominent as it is supercooled [29–32]. Such anomalies have been tentatively associated to structural facts: the presence of two competing preferential local structures, identified with molecules characterized by high or low local density [30,33,34]. Support for this idea comes from the existence of at least two different forms of amorphous glass states, namely low-density amorphous ice and (very) high-density amorphous ice [30,35,36]. Different parameters have been proposed to study the local structural order of the water molecules on a quantitative basis. One of them, proposed by Shiratani and Sasai [33, 34], associates a local structure index I to each molecule to quantify the degree of local order. The key observation is the existence of certain molecules which show an unoccupied gap between 3.2 Å and 3.8 Å in their radial-neighbor distribution for certain periods of time. Such low-density molecules are well structured and coordinated in a highly tetrahedral manner with four other water molecules. Occupancy of such gap increases the local density and distorts the tetrahedral order of the central molecule. Shiratani and Sasai [33, 34] defined $I(i,t)$ for molecule i at time t. For each molecule i one orders the rest of the molecules depending on the radial distance R_j between the oxygen of the molecule i and the oxygen of molecule j: $R_1 < R_2 < R_j < R_{j+1} < \cdots < R_{n(i,t)} < 3.7\,\text{Å} < R_{n(i,t)+1}$. Then, $I(i,t)$ is defined as [33, 34]

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

where $\Delta(j;i,t) = R_{j+1} - R_j$ and $\overline{\Delta}(i,t)$ is the average over all molecules of $\Delta(j;i,t)$. Thus, $I(i,t)$ expresses the inhomogeneity in the radial distribution within the sphere of radius around 3.7 Å.

A high value of $I(i,t)$ implies that molecule i at time t has a good tetrahedral local order and low local density (and thus, a low local potential energy since it is able bind to its first four neighbors by geometrically well-shaped hydrogen bonds), while on the contrary, values of $I(i,t) \approx 0$ indicate a molecule with defective tetrahedral order and high local density (and thus, high local potential energy), even allowing for a fifth or more neighbors within the coordination shell. This abnormal coordination could also promote the formation of bifurcated hydrogen bonds (when a water molecule binds to two others via the same hydrogen), a feature that has been shown to promote local mobility [37, 38]. Liquid water both in the normal liquid state and in the supercooled liquid state has been shown to present I-distributions with a peak at low values and a tail to the right. Over certain temperature range the curves for the different temperatures intersect at a value of $I(i,t) \approx 0.04\,\text{Å}^2$, which can be regarded as a limit between structured (the ones with $I(i,t) > 0.04\,\text{Å}^2$) and unstructured water molecules (with values of I lower than such threshold). The fraction of structured molecules increases as temperature is decreased [33, 34, 39, 40]. In a previous work for pure water [39, 40] we have employed the inherent dynamics formalism (which implies minimizing each instantaneous configuration of the MD simulation, that is, the one for the real dynamics, to reach the basin of attraction or local minimum of the potential energy to which it belongs; that is, by subtracting the kinetic energy). This procedure renders a I-distribution clearly bimodal, with a kind of isosbestic point or local minimum in the distribution whose position is temperature independent, separating two peaks: the one for the unstructured molecules (to the left) and the one for the structured ones (to the right) [39]. However, in this work we shall employ directly the real dynamics to characterize the local structure of the water molecules since it suffices to this end and thus there is no need to resort to the more computationally involved inherent dynamics technique. Additionally, we note that the local structure index I yields similar information on bulk water as other quantities that have been calculated, like the orientational order parameter q, which depends on the values of the angles between the lines connecting the oxygen of a given molecule with those of its four nearest neighbors [39, 41]. However, while I can be directly employed for water at interfaces (since it measures the quality of the local interactions of the water molecules regardless of their number), the index q cannot be used in its original form (which demands the evaluation of the positions of the first four neighbors) and should be reformulated in order to be suited for an interface [23–25]. In fact, we have already made use of I for protein hydration water [25]. The use of these kind of indices (like I) to characterize the local structure of the water molecules at an interface (instead of the study of density profiles and other parameters already used and which provide less detailed information) is also precisely relevant since they can point to connections with dynamical quantities (for example, they have proven useful in determining the existence of a causal link between structure and dynamic propensity in glassy water, an issue of great interest in glass physics [42–44]).

3 Results and discussion

We calculated the mobility of the water molecules within the desolvation domain of each HB of the different proteins studied. We recall that, as defined before, this desolvation domain represents the reunion of two intersecting spheres with radius 6 Å centered at the α carbons of the residues

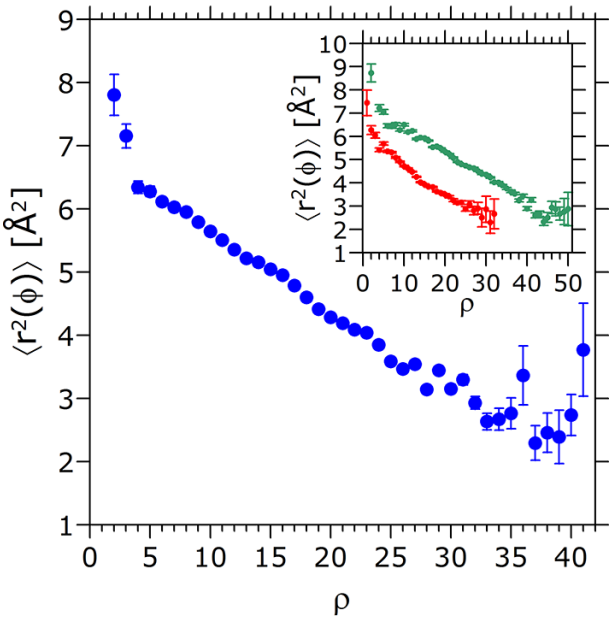

Fig. 2. $\langle r^2(\phi)\rangle$ *vs.* wrapping, ρ, for water molecules i within desolvation domains of HBs. Error bars: standard deviation. $\phi = 4$ ps. The inset shows the curves when, instead of using a value of 6.0 Å for the radius of the desolvation spheres, we employ 5.4 Å (red dots) or 6.6 Å (green dots).

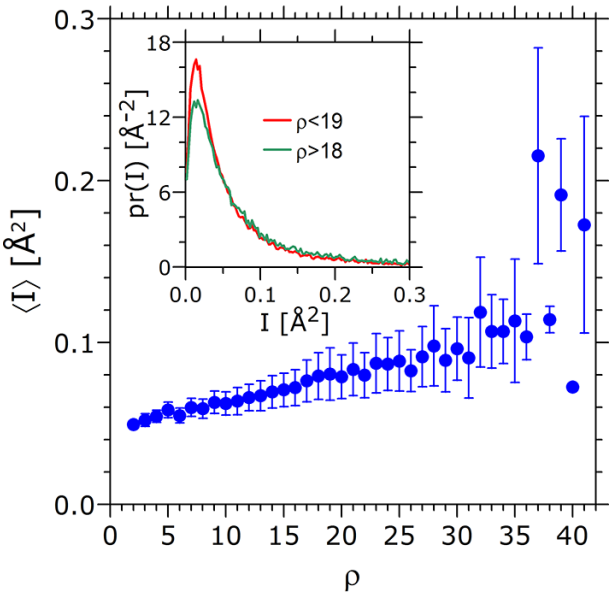

Fig. 3. Averaged value of the local structure index I as a function of the wrapping of the HB (blue dots). Inset: distribution of the local structure index I (where $pr(I)$ indicates the probability of finding a value I for the local structure index) for the water molecules within the desolvation domains of dehydrons (red curve) and well-wrapped HBs (green curve). Data for water molecules with only one neighbor within 3.7 Å, which yield a value of $I = 0$, were not considered.

paired by the mainchain HB [8, 9, 12–14]. Thus, most of the water molecules within this region are very close to the protein backbone and hence, are expected to present very low mobility values.

The main result of our study is presented in fig. 2. In such graph we display the average mobility of the water molecules within the desolvation domain of HBs as a function of the wrapping ρ-value of the HB, averaged over all the proteins and configurations studied. From such a picture we can learn on the existence of a clear correlation between wrapping and water local diffusivity, since the mobility value monotonically decreases as the level of wrapping of the HB increases, following a fairly linear dependence. This means that water molecules close to the highly wrapped HBs are practically immobile (very low mobility values) or, in other words, such water molecules are tightly bound to the surface and thus are hard to remove. On the other hand, the water molecules at the local environment of an underwrapped HB are clearly more labile. When we just discriminate between dehydrons ($\rho < 19$) and well-wrapped HBs ($\rho \geq 19$), we find that the average mobility of the water molecules around the latter ones is around $\langle r^2(\phi)\rangle = 3.5\,\text{Å}^2$, while it amounts to $\langle r^2(\phi)\rangle = 5\,\text{Å}^2$ for the water molecules surrounding dehydrons. These results are consistent with the notion that dehydron motifs promote their local desolvation. In turn, since dehydrons have been determined as main components of protein binding sites [4–14], this behavior might be instrumental in the binding of proteins or ligands that must remove hydration water molecules for the association.

The inset in fig. 2 proves that the tendency also holds valid when we change the value of 6.0 Å for the radius of the desolvation spheres to values of 5.4 and 6.6 Å.

It is interesting to test whether, in addition to their enhanced mobility, water molecules surrounding dehydrons also exhibit different features in their strucutral properties as compared to water molecules around well-wrapped HBs. Specifically, it is expected that such molecules would be less structured since, as described in the Methods section, a molecule with a poorer local tetrahedral arrangement should present a faster dynamics. The desolvation domains of dehydrons usually present more water molecules than that of well-wrapped HBs but it is not obvious whether these water molecules do in fact present a less structured local environment. To evidence the increase in water structural order with the wrapping of the HB, in fig. 3 we present the average value of the local structural index I for water molecules within the desolvation domain of the HB as a function of the wrapping of the HB. A clear monotonic increase can be observed, as expected. The inset of such figure shows the distribution of the local structure index of the water molecules within the desolvation domain of the HBs, I, by discriminating between dehydrons and non-dehydrons. As expected, the curve for the dehydrons (red) is a bit displaced to the left (to lower I values), thus speaking of a less structured local arrangement. In this case, the difference between the distributions is not very marked given that we have coarsely discriminated between dehydrons and non-dehydrons, but the results are still consistent.

It is worth mentioning that, on average, in the outermost limits of the ρ spectrum, HBs with $\rho = 4$ have 21 water molecules in their desolvation domains, while HBs with $\rho = 33$ have only 3. Thus, obviously the local environments of non-dehydrons and dehydrons differ from each other, somehow implying sparse water in hydrophobic confinement and interconnected higher-density water close to a more hydrophilic-like surface, respectively. However, the distribution of distances of water molecules (within desolvation domains) to the geometrical center of the HBs is almost identical for dehydrons and for well-wrapped HBs ($\rho \geq 19$). That is, even though there are on average more water molecules within desolvation domains in dehydrons than in well-wrapped HBs, the probability to find a water molecule to a certain distance of the geometrical center of a HB is well-nigh the same for both cases. Additionally, we wish to note that unlike models usually employed to study water under hydrophobic or hydrophilic confinement, the HB desolvation domains we are dealing with do not imply a fixed geometry. Some portions of the protein chain (as unstructured regions or loops which are usually dehydron-rich) are expected to exhibit enhanced structural fluctuations. In fact, certain HBs can even be subject to the disruptive effect of water hydration, thus promoting main chain motions. Hence, the situation we face, while might share certain features with simple model settings of hydrophobic/hydrophilic confinement, is more complicated given the particular characteristics of the protein structure and dynamics and the distinctive water-protein interactions.

Noteworthy, an inverse correlation similar to that of fig. 2 has been determined between protein structural disorder and wrapping [9]. In such work, disorder propensity was quantified by a sequence-based score generated by the program PONDR-VLXT [45, 46]. In this sense, it is expected that dehydrons, being packing defects or motifs with structural disorder, would be related to regions of enhanced protein chain dynamics and a greater conformational freedom as denoted by a high B-factor (B-factors, or temperature factors, quantify the displacement of the atomic positions from an average or mean value). In turn, the local "equilibration" of the molecular degrees of freedom of protein and water would also be consistent with the existence of labile water molecules ("hot" water) around dehydrons. Thus, to test whether the dehydronic HBs indeed present an enhanced local dynamics and are less stable than well-wrapped HBs, we have monitored the time evolution of the HB distance for dehydrons in large (200 ps) molecular dynamics simulations for the different proteins under study. In such runs we have detected that many dehydronic HBs show significant fluctuations and eventual breaking and reforming events (at some times the HB distance clearly exceeds the HB distance threshold of 3.5 Å and adopts values larger than 5 Å or more, even with intercalation of a water molecule between the two N and O partners or with different water molecules hydrogen-bonding to each partner). In fig. 4A we plot the distribution of the HB distance, d, for the different trajectories of the 62 proteins studied, by discriminating between HBs that at the corresponding initial configuration are dehydrons or non-dehydrons. From such graph we can note that the curve for the dehydrons (red) is a bit displaced to the right as compared to that for the well-wrapped HBs (green) which decays earlier. This fact denotes an enhanced protein chain dynamics at dehydronic regions, a behavior consistent with the existence of more labile water molecules around them, as we have demonstrated above. This locally enhanced conformational freedom or structural plasticity at dehydrons represents a property that might be useful in shaping protein binding sites. Addi-

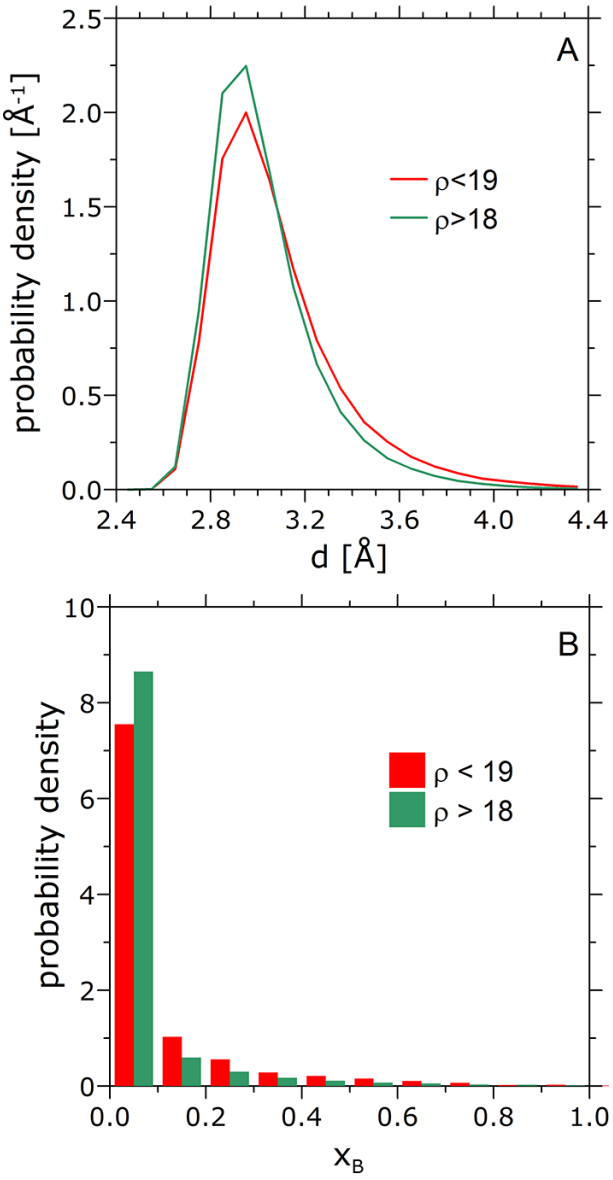

Fig. 4. A) Backbone hydrogen bonds (HB) distance d distribution for dehydrons (red curve) and well-wrapped HBs (green curve). B) The distribution of the quantity X_B, the fraction of time the HB is broken, calculated as the fraction of time the N-O distance is larger than 3.5 Å. Red bars are for dehydrons while the green ones correspond to non-dehydrons. In both cases the results were obtained over 400 trajectories of each of the 62 proteins. The duration of each trajectory was $\phi = 4$ ps.

tionally, in fig. 4B we show the quantity X_B, the fraction of initial HBs that are broken ($d > 3.5\,\text{Å}$) at the end of the simulation runs of $\phi = 4\,\text{ps}$ (although similar results were obtained for MD runs of 200 ps). This picture also indicates that dehydrons are on average more labile than non-dehydrons.

In order to better illustrate the role of dehydrons as dewetting sites and binding promoters, we comment explicitly the situation for one of the proteins studied: the p53 protein (PDB 2GEQ). The case was chosen given that, besides the practical relevance of the p53 protein (p53 is mutated in most of the cases of human cancer), the DNA binding site of this molecule represents one of the largest dehydron clusters in the Protein Data Bank [9]. From a 200 ps molecular dynamics simulation of p53 we found that while the well-wrapped HBs of the protein are on average 97% of the time formed, its dehydrons are broken on average 15% of the time. No well-wrapped HB shows an X_B value larger than 0.25 while one quarter of the dehydrons exhibit X_B larger than such value. In fact, 3 dehydrons (10% of all dehydrons) are broken more than 75% of the time. These results clearly denote the existence of a transference of constraints between regions of the protein chain with an enhanced dynamics and mobile hydration water molecules. We next study in detail the DNA recognition site of this molecule. The dehydronic nature of this recognition site is evident: the three Arginines (Arg 245, 270 and 277) involved in the p53-DNA contacts take part in a cluster of nine dehydrons in the protein chain (a dehydron cluster is defined as the maximal set of dehydrons with intersecting desolvation domains, with the definition of desolvation domain given in our Methods section). Such cluster is formed by the dehydrons (132, 270), (237, 271), (244, 239), (245, 237), (278, 274), (280, 276), (281, 277), (282, 279) and (284, 280), where we indicate both residues (by the residue number) involved in the dehydronic HB. The averaged mobility value for the water molecules hydrating such dehydron cluster (that is, within the desolvation domains of all these HBs) is roughly $\langle r^2(\phi)\rangle = 5\,\text{Å}^2$. This value, consistent with the average mobility value around dehydrons as indicated previously, is significantly larger than the mean mobility value averaged over the water molecules within the desolvation domains of all the HBs of the protein which gives $\langle r^2(\phi)\rangle = 3.5\,\text{Å}^2$, close to the average value for non-dehydrons. These results show that the recognition site of p53 is indeed surrounded by labile water molecules that would be more easily displaced upon association with DNA. Given the large size of this dehydron cluster (9 dehydrons), this molecule presents a large patch of easily removable hydration water in its contact region, thus providing an expedient for DNA approaching and binding. Moreover, in this case the dehydration tendency of the binding site is even more relevant if we consider the particularly electrostatic nature of this binding process. This is so since a recognition process based on electrostatic interactions might not be operative in a bulk water environment, where the high dielectric medium should effectively screen the electrostatic charges. However, by promoting the local desolvation of the p53 DNA-recognition site, this large dehydron cluster would play a main role in quenching the local dielectric and triggering the electrostatic interaction between the positively charged Arginines and the negatively charged backbone phosphates of the DNA.

4 Conclusions

In this work we have presented firm evidence on the existence of a link between the wrapping of the HBs of the protein chain (a structural parameter) and the mobility of the hydrating water molecules (cf. fig. 2). We have shown that protein-backbone exposure is correlated with a looser hydration of the protein surface and thus, that underwrapped intramolecular interactions (dehydrons) are surrounded by tightless-bound, easily removable water molecules. Thus, by promoting their local dehydration, such motifs provide an expedient by which partner proteins or ligands might displace water molecules during the binding process.

The authors acknowledge financial support from ANPCyT and CONICET. AF, GAA, MAM, JAR-F and MBS are research fellows of CONICET. SRA acknowledges a fellowship by CONICET.

References

1. J. Qvist, M. Davidovic, D. Hamelberg, B. Halle, Proc. Natl. Acad. Sci. U.S.A. **105**, 6296 (2008).
2. T. Young, R. Abel, B. Kim, B.J. Berne, R.A. Friesner, Proc. Natl. Acad. Sci. U.S.A. **104**, 808 (2007).
3. C. Wang, B.J. Berne, R.A. Friesner, Proc. Natl. Acad. Sci. U.S.A. **108**, 1326 (2011).
4. A. Fernández, H.A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **100**, 113 (2003).
5. A. Fernández, R. Scott, Biophysical. J. **85**, 1914 (2003).
6. A. Fernández, in *Transformative Concepts for Drug Design: Target Wrapping* Vol. **1** (Springer: Heidelberg, 2010) pp. 1-224.
7. A. Fernández, R. Scott, Phys. Rev. Lett. **91**, 018102 (2003).
8. A. Fernández, J. Chen, A. Crespo, J. Chem. Phys. **126**, 245103 (2007).
9. N. Pietrosemoli, A. Crespo, A. Fernández, J. Prot. Res. **6**, 3519 (2007).
10. A. Fernández, Nat. Biotech. **22**, 1081 (2004).
11. A. Fernández, M. Lynch, Nature **474**, 502 (2011).
12. E. Schulz, M. Frechero, G. Appignanesi, Ariel Fernández, PLoS ONE **5**, e12844 (2010).
13. S.R. Accordino, J.A. Rodríguez-Fris, G.A. Appignanesi, A. Fernández, Eur. Phys. J. E **35**, 59 (2012) and arXiv:1108.2618v1 [cond-mat.soft].
14. S.R. Accordino, M.A. Morini, M.B. Sierra, J.A. Rodríguez Fris, G.A. Appignanesi, A. Fernández, Proteins: Struct., Funct., Bioinf. **80**, 1755 (2012).
15. R.A. Friesner *et al.*, J. Med. Chem. **49**, 6177 (2006).
16. J.L. Kulp *et al.*, JACS **133**, 10740 (2011).
17. S.R. Accordino, J.A. Rodríguez Fris, G.A. Appignanesi, PLoS ONE **8**, e55123 (2013).

18. P. Ball, Nature **423**, 25 (2003).
19. W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. **79**, 926 (1983);
20. M.W. Mahoney, W.L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).
21. D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B.P. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvry, K.F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P.A. Kollman, AMBER 11, University of California, San Francisco (2010).
22. H. Zhong, H.A. Carlson, Proteins: Struct. Funct. Bioinf. **58**, 222 (2005).
23. D.C. Malaspina, E.P. Schulz, L.M. Alarcón, M.A. Frechero, G.A. Appignanesi, Eur. Phys. J. E **32**, 35 (2010).
24. S.R. Accordino, D.C. Malaspina, J.A. Rodríguez Fris, G.A. Appignanesi, Phys. Rev. Lett. **106**, 029801 (2011).
25. S.R. Accordino, D.C. Malaspina, J.A. Rodriguez Fris, L.M. Alarcón, G.A. Appignanesi, Phys. Rev. E **85**, 031503 (2012).
26. A.R. Bizzarri, S. Cannistraro, J. Phys. Chem. B **106**, 6617 (2002).
27. A.R. Bizzarri, A. Paciaroni, S. Cannistraro, Phys. Rev. E **62**, 3991 (2000).
28. 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).
29. P.G. Debenedetti, *Metastable Liquids* (Princeton University Press, Princeton, NJ, 1996).
30. O. Mishima, H.E. Stanley, Nature **396**, 329 (1998).
31. C.A. Angell, Chem. Rev. **102**, 2627 (2002).
32. C.A. Angell, Annu. Rev. Phys. Chem. **55**, 559 (2004).
33. E. Shiratani, M. Sasai, J. Chem. Phys. **104**, 7671 (1996).
34. E. Shiratani, M. Sasai, J. Chem. Phys. **108**, 3264 (1998).
35. H.-G. Heide, Ultramicroscopy **14**, 271 (1984).
36. T. Loerting, N. Giovambattista, J. Phys.: Condens. Matter **18**, 919 (2006).
37. F. Sciortino, H. Geiger, H.E. Stanley, Phys. Rev. Lett. **65**, 3452 (1990).
38. J.A. Rodriguez Fris, G.A. Appignanesi, E. La Nave, F. Sciortino, Phys. Rev. E **75**, 041501 (2007).
39. G.A. Appignanesi, J.A. Rodriguez Fris, F. Sciortino, Eur. Phys. J. E **29**, 305 (2009).
40. S.R. Accordino, J.A. Rodriguez Fris, F. Sciortino, G.A. Appignanesi, Eur. Phys. J. E **34**, 48 (2011).
41. J.R. Errington, P.G. Debenedetti, Nature **409**, 318 (2001).
42. D.C. Malaspina, J.A. Rodríguez Fris, G.A. Appignanesi, F. Sciortino, Europhys. Lett. **88**, 16003 (2009).
43. G.A. Appignanesi, J.A. Rodríguez-Fris, R.A. Montani, W. Kob, Phys. Rev. Lett. **96**, 057801 (2006).
44. G.A. Appignanesi, J.A. Rodríguez-Fris, M.A. Frechero, Phys. Rev. Lett. **96**, 237803 (2006).
45. L.M. Iakoucheva, A.K. Dunker, Structure (London) **11**, 1316 (2003).
46. A.K. Dunker, Z. Obradovic, Z. Nat. Biotechnol. **19**, 805 (2001).