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

Eur. Phys. J. E (2015) **38**: 107 DOI 10.1140/epje/i2015-15107-3

"Chameleonic" backbone hydrogen bonds in protein binding and as drug targets

C.A. Menéndez, S.R. Accordino, D.C. Gerbino and G.A. Appignanesi

edp sciences Società Italiana di Fisica Springer

Eur. Phys. J. E (2015) **38**: 107
DOI 10.1140/epje/i2015-15107-3

THE EUROPEAN PHYSICAL JOURNAL E

Regular Article

"Chameleonic" backbone hydrogen bonds in protein binding and as drug targets*

C.A. Menéndez, S.R. Accordino, D.C. Gerbino, and G.A. Appignanesi[a]

INQUISUR-UNS-CONICET and Departamento de Química, Universidad Nacional del Sur, Avenida Alem 1253, 8000-Bahía Blanca, Argentina

Received 30 May 2015 and Received in final form 27 July 2015
Published online: 19 October 2015 –

Abstract. We carry out a time-averaged contact matrix study to reveal the existence of protein backbone hydrogen bonds (BHBs) whose net persistence in time differs markedly form their corresponding PDB-reported state. We term such interactions as "chameleonic" BHBs, CBHBs, precisely to account for their tendency to change the structural prescription of the PDB for the opposite bonding propensity in solution. We also find a significant enrichment of protein binding sites in CBHBs, relate them to local water exposure and analyze their behavior as ligand/drug targets. Thus, the dynamic analysis of hydrogen bond propensity might lay the foundations for new tools of interest in protein binding-site prediction and in lead optimization for drug design.

1 Introduction

Backbone hydrogen bonds (BHBs) have been shown by Linus Pauling to constitute major determinants for protein structure, responsible for shaping their main secondary and tertiary structural motifs. In soluble proteins, such non-covalent interactions are stable provided water is significantly excluded from their local environment by aminoacid side-chains. However, this requirement is not necessarily met all along the protein chain, particularly at protein binding sites [1–10]. Precisely, the hydration properties of binding sites have been suggested to play a main role in ligand-binding or in protein-protein association [1–10] since labile hydration-water molecules are expected to be displaced from the protein binding site [1–4]. Indeed, the replacement of so-called "unfavorable" water molecules by groups of the ligand complementary to the protein surface has been established as a principal driving force for binding [1–4] and, as such, has been incorporated in computational structure-based strategies [2–4]. Also, fragment clustering approaches (binding of small molecular probes) have been combined with exclusion maps dictated by the pattern of tightly bound water molecules [5] in order to detect protein binding sites. In accord with this heterogeneous scenario for protein hydration, the existence of BHBs partially exposed to the solvent (incompletely wrapped BHBs or dehydrons [6, 7]) has been established together with their relevance for protein binding [6–10]. Such motif, which exhibits an enhanced dehydration propensity, represents a structural packing defect that is readily determined from PDB coordinates [6–10]. Thus, this concept enabled a potent novel strategy for drug design [6–10] that simply relies on structural information (that is, on the structural characterization of the hydrogen bonds *already* present in the PDB of the apo protein). However, since proteins are inherently dynamical objects, the merely binary (formed/not formed) classification of non-covalent interactions provided by PDB structures might be veiling valuable information regarding protein interactions and function, as we shall here demonstrate. Specifically, we shall find PDB BHBs that tend to be disrupted during the dynamics while other BHBs, completely absent in the PDB of the apo protein, display a significant dynamical persistence. We shall also show that such "chameleonic" BHBs (CBHBs since they change state from the PDB prescription to the opposite formation propensity in solution) are not homogeneously distributed along the protein chain but concentrate in binding regions. Additionally, we shall show that CBHBs are removed upon ligand binding, thus revealing their role as drug targets. These results imply that the dynamic analysis of BHB propensity might be easily translated into a novel drug design concept.

2 Methods

In this work we shall focus on protein-protein interactions for which disruptive small molecules or drugs have been developed [11]. Besides their main therapeutic relevance, these systems provide a more accurate means

* Supplementary material in the form of a .pdf file available from the Journal web page at
http://dx.doi.org/10.1140/epje/i2015-15107-3

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

to determine the target protein binding site, since the small molecule/drug is significantly smaller than the partner protein they replace [9, 11]. This does not represent any loss in generality, since protein binding "hot spots" (the region that comprises the residues that mainly contribute to the binding free energy) are usually restricted to just a handful set of aminoacids [9, 11]. We shall study some successful cases [11]: MDM2(PDB: 1Z1M)/Nutlin-3 (PDB: 1RV1)/p53 (PDB: 1YCR) [12], IL-2 (PDB:1M4C)/ SP4206 (PDB: 1PY2)/IL-2 receptor α-chain (PDB: 1Z92) [13], BCL-XL (PDB: 1R2D)/ABT-73734 (PDB: 2YXJ)/BAD-derived peptide (amino acids 100126) (PDB: 2BZW) [14] and ZipA (PDB: 1F7W)/Compound 1 (PDB: 1Y2F)/FtsZ-derived peptide (PDB: 1F47) (amino acids 367-383) [15]. In each case the notation indicates first the target protein, then the disruptive drug and finally the partner protein, with the corresponding PDB entries for the apo protein, the drug-protein complex and the protein-protein complex. We have also chosen these cases for validation purposes (all 3D structures have been accurately determined and deposited in the PDB). However, our analysis is generally applicable to the broad realm of protein binding. To study the dynamical behavior of these systems were carried out molecular dynamics simulations by means of AMBER simulation package 10 [16], using in all cases periodic boundary conditions, TIP3P water and $T = 300\,\mathrm{K}$ and by means of the same minimization and equilibration protocol. Equilibration was tested by monitoring the behavior of thermodynamical properties like temperature, pressure and energy oscillations (in the Supplementary Material we provide the time evolution of the root mean square displacement, RMSD, for the different systems we studied; Supplementary fig. 1). For all apo proteins we performed production runs of 50 ns with 5 fs time step, recording 10000 equally spaced configurations. For all other systems production dynamics were performed for 20 ns with 5 fs time step and saving 4000 configurations. To determine the binding site for each protein we used a simple geometrical method by finding BHBs in the target protein whose distance (measured form the N amide or the carbonyl O) to any heavy atom of their partner protein is less than 6 Å [9, 10] in the PDB of the corresponding protein-protein complex (criterion 1). Another possibility (criterion 2), that yields a smaller size binding site, is to use the same geometrical criterion but with respect to the heavy atoms of the drug (small molecule) in the corresponding drug-protein complex PDB structure. We note that with these methods we determine the BHBs located within a certain interface or binding site. However, the protein might have other interacting/binding sites we are not considering here. For the four proteins studied, criterion 1 found that roughly 15% of the BHBs were located at binding sites, while criterion 2 reduced this percentage to half that value.

The stability of non-covalent interactions like protein BHBs is expected to depend strongly on their local (hydration) environment [6–10]. Thus, we first determined the existence of heterogeneities in protein hydration properties. To this end, we calculated the water vacating probability, $P(N = 0)$, in small observational (spherical) volumes of radius 4.0 Å [17, 18] at each site of interest, with N being the number of water molecules within such volume ($P(N = 0)$ is calculated by computing the number of configurations the observational volume is empty of water molecules divided by the total number of configurations considered along the long MD runs performed). This indicator (which contains an equivalent information as that of normalized water density fluctuations [18]), represents a good measure of local hydrophobicity, since hydrophobic-like surfaces present a higher water vacating probability (larger normalized density fluctuations) than the ones displayed by hydrophilic-like surfaces. In small observation volumes $\mu = k_b T \ln[p(N = 0)]$, where μ is the free energy of formation of a cavity of such small radius [18], k_b is Boltzmann constant and T is the absolute temperature. Concurrently, a high value of the normalized density fluctuations at a given place indicates a favorable work of cavity creation at such place [18] and, thus, a high hydrophobicity. We first calculated $P(N = 0)$ at observation volumes centered at every heavy atom of the protein for large simulation runs (50 ns or 20 ns according to the system) after equilibration, starting from the reported PDB structure. A value of $P(N = 0)$ close to unity implies that the atom is completely "dry" or desolvated (absence of a first hydration shell). If we focus on the protein surface, we find that most of the surface is well hydrated (hydrophilic), while a patch of enhanced vacating probability (hydrophobic behavior) is observed at the binding site. This result, already anticipated in the works of the group of S. Garde [18] is compatible with the belief that ligands are expected to displace labile hydration water molecules from their protein binding site [1–4]. However, since here we explicitly focus on the BHB as protagonist of protein binding, we decided to calculate $P(N = 0)$ at the amide (N atom) and carbonyl (O atom) moieties of each of the BHBs of the PDB.

3 Water vacating probability around backbone hydrogen bonds

The results of the water vacating probability around BHBs calculated during simulation are displayed in fig. 1(a), which depicts the behavior of $\langle P(N = 0)\rangle$ (the probability to find the observational volume empty of water moleucles, as already indicated) for the complex MDM2-p53 and fig. 1(b), for the apo form of protein MDM2. In fact, as for the rest of our work, we simulated the apo form of the N-terminal domain of MDM2, $\mathrm{MDM2}^N$, and a complex of $\mathrm{MDM2}^N$ with a peptide fragment taken from the transactivation domain of p53 [19]. The preponderance of red colors in fig. 1(a) ($P(N = 0)$ close to unity) speaks of the clear dehydration propensity of the backbone hydrogen bonds of the MDM2 protein in complex with p53. This result implies that the above-mentioned predominance of hydrophilic regions ($P(N = 0)$ close to zero) at the surface of the apo protein outside the binding site is given by well-hydrated side-chains that, in turn, protect the

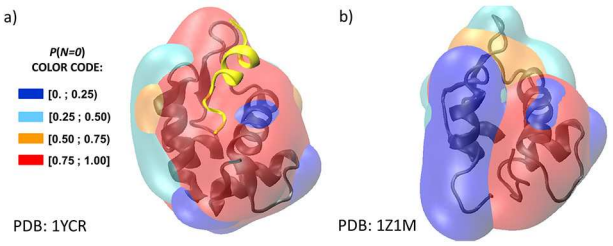

Fig. 1. (a) Water vacating probability $P(N = 0)$ for the BHBs of the MDM2 protein in complex with p53. (b) Idem for the apo MDM2 protein.

BHBs by promoting their drying. However, when we turn to fig. 1(b), where we analyze the apo MDM2 protein, we can learn on the existence of a significant number of water-accessible BHBs (light-blue spheres) at the binding site. Thus, it is evident that the local environments of BHBs along the protein chain vary significantly, with predominance of buried or dry BHBs but also with binding-site BHBS which are (at least partially) water exposed.

4 Dynamic analysis of BHB propensity

We next recorded a time-averaged, or dynamic, BHB contact matrix (DBHB-CM) for the different proteins. This was done by calculating the fraction of time each BHB is formed during long runs after equilibration. At each evaluation time, if a pair of residues i and j satisfy a hydrogen bonding criterion (N-O cutoff distance, $r < 3.5$ Å; N-H-O cutoff angle, $\theta > 140°$ [10]), the corresponding $\{i, j\}$ matrix element becomes 1, while it is 0 otherwise. Then, we averaged the results for each matrix element at all evaluation times. Thus, the DBHB-CM contains values that range from 0 (never formed) to 1 (formed all the time) for the different matrix elements (for the sake of clarity, we discarded BHBs that were formed for less than 10% of the total run; this choice is arbitrary but the results do not depend on it). With this tool we can better learn on the correlation between BHB stability and water exposure. Thus, we calculated the mean value of the water vacating probability, $P(N = 0)$, for the BHBs whose average formation time was over 0.8 (very stable) and for the ones that were formed for a fraction of time less than 0.3 (the less persistent BHBs). Restricting the analysis for the BHBs located within the binding sites of the apo proteins (criterion 1, but similar results are obtained for criterion 2) we found that the very stable BHBs are practically dry ($\langle P(N = 0)\rangle = 0.9287$; $\sigma = 0.108$, where σ is the standard deviation), while the less persistent ones are indeed water exposed ($\langle P(N = 0)\rangle = 0.3995$; with a large σ-value of 0.228). Another interesting result is that the population of less persistent BHBs was greatly depleted when we considered the complexes of the target proteins with their partner proteins or with their corresponding disruptive molecules.

To contrast the dynamical behavior of the protein with the information provided by the PDB structure we calculated, for each matrix element of the DBHB-CM, its (absolute value) distance, D, with respect to the corresponding matrix element of the BHB-CM of the PDB (PDB-BHB-CM, with values that are either strictly 1 or 0 if the BHB under consideration is formed or not in the PDB structure, respectively). In other words, for each BHB we calculated the (absolute value) difference between its formation propensity during the dynamics (its time-averaged state in the dynamics) and its corresponding state-value in the PDB structure. The resulting Distance Matrix (DM) is shown in fig. 2(a) for the case of the protein MDM2 in its apo form, while in fig. 2(b) we show the DM for the complex MDM2-p53 (calculated by comparing the DBHB-CM and the PDB-BHB-CM of the MDM2-p53 complex; we recall that we are studying $MDM2^N$ and the complex of $MDM2^N$ with a peptide fragment taken from the transactivation domain of p53 [19]). In turn, in fig. 2(c) we present the case for the complex MDM2-Nutlin3. The color intensity indicates the size of the distance value, D: intense means large distance, while pale indicates low distance. If we focus on the case of the apo protein, we can learn that while most BHBs of the PDB are stable during the dynamics and thus yield a low distance value, D, there exist a significant number of BHBs that indeed display large distance values. Such BHBs represent either interactions that are present in the PDB but that disappear during the dynamics or interactions that are absent in the PDB but that are persistently formed during the dynamics. Such BHBs that "change color" (roughly from black to white or vice versa) between their PDB-BHB-CM (the matrix corresponding to the PDB) and the DBHB-CM (the matrix for the dynamics), will be termed as "chameleonic" BHBs, CBHBs. In this sense, the actual dynamical state in solution of such interactions is "hidden" in the PDB structure of the apo protein under an opposite state-value (which might thus confer a misleading dynamical expectation). We note that the determination of CBHBs resorts simply to the PDB structure of the apo protein and a dynamical analysis based on a molecular dynamics simulation form such structure (if possible, it might be worthy to employ, instead, experimental dynamical information from techniques like NMR). In turn, when we focus on the DM for the complex of MDM2 with its natural ligand, the MDM2-p53 complex, we can see that the dynamical and the PDB BHB-CMs are more similar, since most of the CBHBs of the binding site have disappeared as a result of the stabilization gained by the binding site upon ligand association. A similar result is found when we study the case of the complex MDM2-Nutlin3, a disruptive drug that acts by mimicking the behavior of p53 and represents a potent inhibitor. It is interesting to note that certain regions of secondary structure are quite disordered in the apo PDB (for example, the region between residues 101 and 107), while they are nicely structured in the complex, as can be learnt from the helical structure in the MDM2-p53 complex. These residues are located in regions of the apo protein with water accessible BHBs, as can be learnt from fig. 1. However, when in complex with p53 (or with the disruptive molecule Nutlin-3) these BHBs become sta-

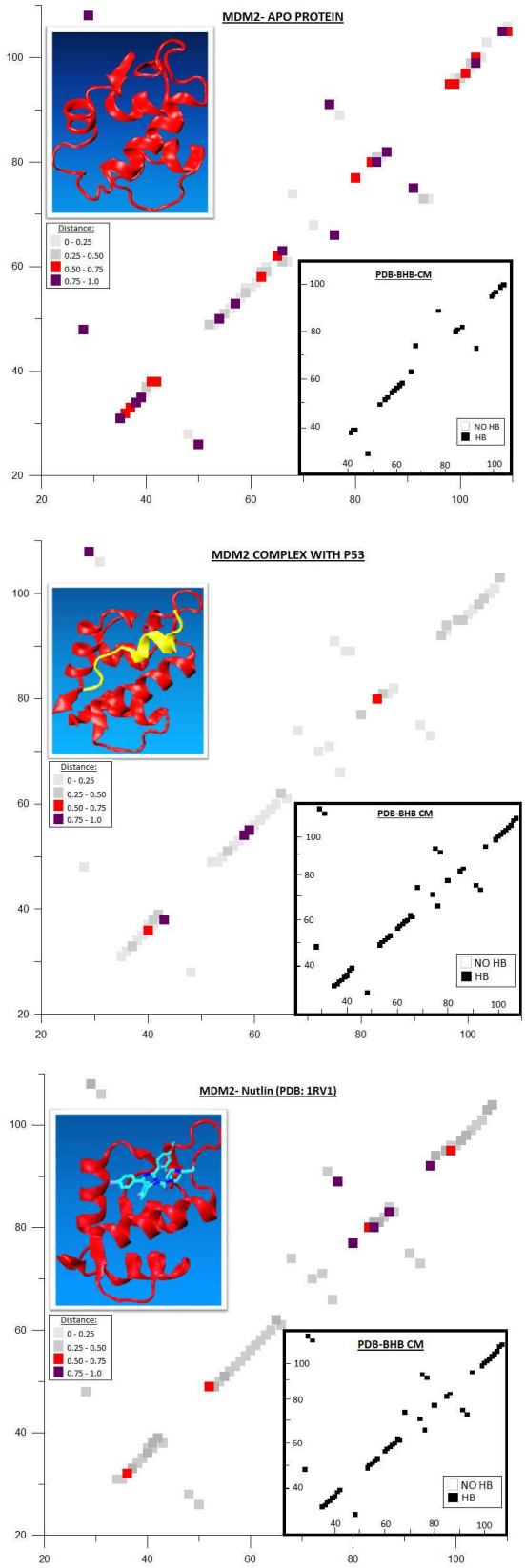

Fig. 2. (a) Distance matrix for a region containing the binding site of the apo form of protein MDM2. (b) Idem for the MDM2-p53 complex. (c) Idem for MDM2-Nutlin3. The insets show the 3D structure and the BHB-CM of the corresponding PDBs.

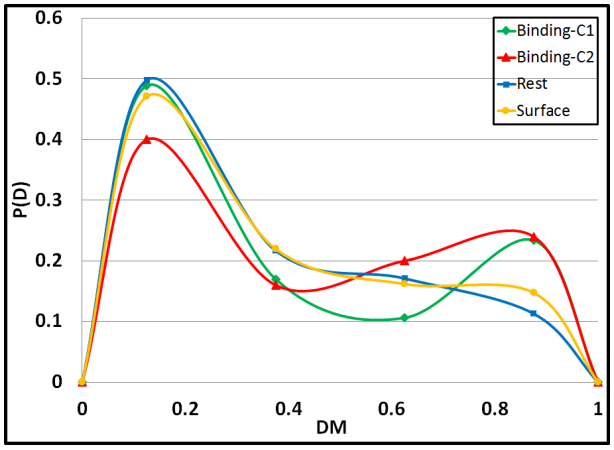

Fig. 3. Probability density distribution of distance values, D, for all the BHBs of the four apo proteins studied. We discriminate between the BHBs located at protein binding sites (criteria 1 and 2) from the rest of the protein BHBs (BHBs outside the binding site). The latter are classified as "Rest" (including all the BHBs outside the binding site, regardless if they are located at the protein surface or buried at the protein interior) or "Surface" (only the ones that are located at the protein surface).

bilized in the helical structure since they are now additionally inter-molecularly protected from hydration. Additionally, when we performed a simulation starting from the PDB structure of the apo MDM2 protein but with Nutlin-3 located at the binding site, we found that the secondary structure of this helix is significantly improved during the simulation. The plasticity of the MDM2 binding site, as made evident by the conformational changes that accompany the binding of p53 to MDM2^N, had been already determined experimentally [19]. In such work [19] it was shown that upon binding of p53, the binding cleft of MDM2^N undergoes an expansion, achieved through a rearrangement of its two pseudosymetrically related subdomains, resulting in the outward displacements of the secondary structural elements that comprise the walls and floor of the p53-binding cleft. Additionally, it has been stated that MDM2^N becomes more rigid and stable upon binding p53 [19], also in accord with our results. We then repeated the analysis of fig. 2 for the other three proteins and their complexes both with their natural protein ligand and with the corresponding disruptive drug or small molecule, finding in all cases similar results (shown in the Supplementary Material; Supplementary figs. 2, 3 and 4).

Additionally, in fig. 3 we provide the probability distribution of distance values, D, of all the matrix elements of the DM of the four apo proteins studied. We distinguish between BHBs at the binding sites of all proteins studied and for the rest ones (BHBs located outside the binding sites studied, indicated as "Rest"). We used criterion 1, but we also include the distribution for the binding site BHBs as determined by criterion 2, which provides similar results. A direct inspection reveals the dominance of two different populations, large or low distance values, together with a conspicuous depletion in the region of

Mol	CBHBs		Apo Prot	Complex (Prot.-Prot.)		Complex (Prot.-Mol.)	
MDM2	54	50	0.89	0.12	✓	0.06	✓
	57	53	0.90	0.16	✓	0.11	✓
	62	58	0.69	0.09	✓	0.12	✓
	99	95	0.56	0.29	✓	**0.59**	✗
	100	96	0.45	0.10	✓	0.05	✓
	103	99	0.85	0.29	✓	0.17	✓
IL-2	39	35	0.84	0.40	✓	Not formed	—
	41	38	0.87	**0.66**	✗	0.36	✓
	61	57	0.56	**0.61**	✗	0.37	✓
	68	64	0.77	**0.77**	✗	**0.90**	✗
	69	65	0.77	0.41	✓	**0.57**	✗
BCL	100	96	0.94	0.09	✓	0.10	✓
	125	121	0.54	0.29	✓	0.41	✓
	126	122	0.52	0.39	✓	**0.81**	✗
ZIP	40	44	0.58	0.38	✓		
	44	41	0.45	0.39	✓		
	64	81	0.85	0.17	✓		

Fig. 4. Distance values, D, for the CBHBs for the four apo proteins studied, indicated by the pairs of residues comprised in the hydrogen bond (a value larger than 0.5 corresponds to a CBHB). We also provide the corresponding D-value for the protein-protein and protein-small molecule complexes. Whenever this value falls below 0.5 it indicates a CBHB removal upon complex formation. For the ZipA protein we do not provide the D values for the complex with the small molecule since its binding affinity is low.

medium-sized distances. The peak for low D values exhibited by all the curves is obviously expected, since it speaks of BHBs which are stable in the apo protein. However, the other population (high D values) signals the presence of CBHBs. More interestingly, we can learn that the protein binding site is clearly enriched in CBHBs, as indicated by the right peak that visibly develops for both criteria for the binding site BHBs. To discard that this effect were due to the nature of the binding-site BHBs as surface BHBs which are expected to be more labile than buried BHBs, we also present the curve for the BHBs located outside the binding site but that nonetheless reside at other regions of the protein surface (we call them "Surface" BHBs and identify them with a proximity criterion to water: a Surface BHB is one whose distance from any water molecule is less than 6 Å). As expected, from comparison with the curve for all the residues outside the binding site ("Rest" curve), the "Surface" BHBs are a bit more labile than non-surface or buried BHBs. However, the behavior is very different for the clear bimodality and the concurrent enrichment in CBHBs presented by the BHBs of the binding site.

When we compare the apo protein with the complex with its natural protein partner, we find that the binding process removes CBHBs. In fact, if we use a D-value larger than 0.5 in the DM (belonging to the right peak of fig. 3) as a threshold for CBHB we find that for the four apo proteins, more than 80% of their CBHBs disappear upon complex formation (they cease to be CBHBs within the complex). Figure 4 presents such study. From direct inspection we can learn that for MDM2, BCL and Zip proteins, all CBHBs present in the corresponding apo form are removed upon complex formation with the partner protein (for IL-2 this CBHB removal is only partial, since three CBHBs still remain as such in the corresponding protein-protein complex). A similar result is obtained when we compare the DMs of the target apo proteins with the DM for their corresponding protein-small molecule complex, albeit with a performance that is a bit suboptimal in terms of CBHB removal. These facts reveals CBHB-pattern determination as instrumental in defining ligand/drug targets. In particular, wherever a drug scaffold or lead compound performs a suboptimal CBHB quenching, this method might help as a reengineering tool for drug optimization. As such, it might be easily implemented within existing methods to incorporate a relevant dynamical dimension disregarded by structurally based approaches. In this sense, our method not only incorporates a dynamical analysis, but also puts the spotlight on a specific dynamical element essential for the binding process: the dynamics of BHBs.

5 Conclusions

Our work revealed the existence of "chameleonic" protein backbone hydrogen bonds (CBHBs) whose formation propensity during the dynamics significantly differs from their PDB state. We also found that backbone hydrogen bonds mainly consist of either CBHBs or stable ones, with a low population of intermediate states. Additionally, the relative abundance of CBHBs is considerably enhanced at protein binding sites where the protein chain is partially exposed to hydration, thus revealing their role as ligand targets. Upon association, the dehydration of the binding site is completed with the concurrent stabilization and CBHBs removal. Thus, the dynamic analysis of hydrogen bond propensity that determines the pattern of CBHBs might be useful to build protein binding-site predictors and to develop novel design concepts for drug design. In this sense, our approach would be complementary to existing methods [1–10] since it puts in the picture a novel relevant object, the CBHB, and might enable us to focus on backbone hydrogen bond stabilization as an operational principle to engineer better scaffolds.

Author contribution statement

All authors contributed equally to this paper.

The authors are deeply indebted to Prof. Ariel Fernández for enlightening conversations.

References

1. 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. R.A. Friesner *et al.*, J. Med. Chem. **49**, 6177 (2006).
5. J.L. Kulp III *et al.*, J. Am. Chem. Soc. **133**, 10740 (2011).
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. S.R. Accordino, J.A. Rodriguez-Fris, G.A. Appignanesi, A. Fernández, Eur. Phys. J. E **35**, 59 (2012).
9. S.R. Accordino, M.A. Morini, M.B. Sierra, J.A. Rodriguez Fris, G.A. Appignanesi, A. Fernández, Proteins: Struct., Funct., Bioinf. **80**, 1755 (2012).
10. M.B. Sierra, S.R. Accordino, J.A. Rodriguez-Fris, M.A. Morini, G.A. Appignanesi, A. Fernández Stigliano, Eur. Phys. J. E **36**, 62 (2013).
11. J.A. Wells, C.L. McClendon, Nature **450**, 1001 (2007).
12. L.T. Vassilev *et al.*, Science **303**, 844 (2004).
13. B.C. Raimundo *et al.*, J. Med. Chem. **47**, 3111 (2004).
14. M. Bruncko *et al.*, J. Med. Chem. **50**, 641 (2007).
15. T.S. Rush, J.A. Grant, L. Mosyak, A. Nicholls, J. Med. Chem. **48**, 1489 (2005).
16. D.A. Case *et al.* AMBER10, University of California, San-Francisco, CA (2008).
17. L.M. Alarcón, J.A. Rodriguez Fris, M.A. Morini, M.B. Sierra, S.A. Accordino, J.M. Montes de Oca, V.I. Pedroni, G.A. Appignanesi, in *Membrane Hydration*, edited by E.A. Disalvo (Springer Verlag, Berlin, 2015).
18. S.N. Jamadagni, R. Godawat, S. Garde, Annu. Rev. Chem. Biomol. Eng. **2**, 147 (2011).
19. S. Uhrinova, D. Uhrin, H. Powers, K. Watt, D. Zheleva, P. Fischer, C. McInnes, P.N. Barlow, J. Mol. Biol. **350**, 587 (2005).