

Electrostatic Exploration of Biomolecular Interfaces: The Chemical Function of Interfacial Water

2

Abstract

This chapter explores the dielectric structure of interfacial water that envelopes a soluble protein. The most striking feature arising from this *epiststructural* analysis is the breakdown of the Debye *ansatz* that postulates the alignment of water polarization with the protein electrostatic field. The complexities of biological interfaces are shown to be in good measure due to this departure from the standard dielectric picture that has been historically extrapolated from the bulk interface. Accordingly, concepts like the permittivity coefficient are shown to be inadequate or at best insufficient to describe the interfacial dielectrics. The departure from bulk-like behavior is shown to enhance the physicochemical inhomogeneity of protein surfaces and enable their chemical functionality. The epiststructural analysis identifies a structural defect known as dehydron as the causative of anomalous polarization effects that bring about the breakdown of the Debye standard picture. The epiststructural analysis in the previous chapter revealed that interfacial tension is a central thermodynamic factor driving biomolecular events and may be stored as the electrostatic energy associated with the non-Debye component of water polarization. This chapter refines and completes this line of thought by showing that dehydrons locally induce interface basicity as a consequence of the promoted departure from Debye dielectrics. In this way, we delineate a chemical role for dehydrons as triggers of proton transfer events that enhance the nucleophilicity of enzymatically active groups nearby.

2.1 Interfacial Tension and Non-Debye Polarization of Interfacial Water

As noted in the previous chapter, the non-Debye orthogonal component $\overrightarrow{P^\#}$ of polarization is commensurate with the distortion in the structure of water due to subnanoscale confinement [1]. More precisely: $\overrightarrow{P^\#} = \xi \overrightarrow{\nabla} g$, yielding an

equivalence between electrostatic energy of orthogonal polarization and interfacial tension:

$$U^\# = \frac{1}{2} \epsilon_0^{-1} \int \left\| \overrightarrow{P^\#} \right\|^2 d\vec{r} = \frac{1}{2} \lambda \int \left\| \overrightarrow{\nabla} g \right\|^2 d\vec{r} \quad (2.1)$$

This equation asserts the electrostatic origin of interfacial tension, prompting the question: Where do the most significant contributions to non-Debye polarization take place? Since dehydrons create the biggest drops in g -value, we may intuitively assert that dehydrons are the culprits of the breakdown of the Debye picture and part of this chapter is devoted to prove this conjecture.

A measure of the local departure from the Debye scenario within the protein structure may be obtained for each residue generically denoted by n . This measure is furnished by a structure-dependent parameter ϑ_n that we term anomalous polarization fraction (APF) and define as

$$\vartheta_n = \langle U_n^\# / U_n \rangle \quad (2.2)$$

where the symbol “ $\langle \cdot \rangle$ ” denotes time average, and

$$U_n = \frac{1}{2} \epsilon_0 \int \left\| \overrightarrow{\nabla} \varphi \right\|^2 d\vec{r}, \quad U_n^\# = \frac{1}{2} \epsilon_0^{-1} \int \left\| \vec{P} - (\vec{P} \cdot \vec{e}) \vec{e} \right\|^2 d\vec{r}, \quad (2.3)$$

where, as indicated in Chap. 1, the electrostatic potential φ is defined through the relation: $-\overrightarrow{\nabla} \varphi = \vec{E} + \epsilon_0^{-1} \vec{P}$. Integration in Eq. (2.3) extends over a neighborhood around residue n defined as a sphere of radius r centered at its α -carbon. To capture the environment of all side chains we adopted $r = 6$ Å, the approximate diameter of tryptophan (W), the largest side chain. We also evaluated the APF using a larger radius ($r = 8$ Å) to determine the dilution of the anomalous polarization effect as the bulk solvent region is approached. The centering of the n -residue sphere at the α -carbon is justified a posteriori, as shown subsequently. We expected and confirmed that anomalous polarization would be mainly related to poor packing of the protein backbone, while the packing defects are identified by introducing backbone solvation domains consisting of spheres of radius 6 Å centered at the α -carbons. Thus, by centering the residue spheres at the α -carbons, we simultaneously interrogate the backbone and the side chains in search for anomalous dielectric patterns in interfacial water. A more obvious reason for choosing the α -carbon as opposed to other side chain carbon atom relates to the fact that we would otherwise need to treat glycine (G) as an exceptional case.

The parameter ϑ_n is computed at protein/water interfaces as a time average over a 5 ns period beyond equilibration of the protein structural backbone with the solvent. Thus, water polarization for soluble natural proteins with structures reported in the PDB (PDB) may be computed along molecular dynamics trajectories. Each 10 ns-trajectory is generated using as starting point the equilibrated

structural coordinates that result after thermalization of the PDB-reported structure immersed in a pre-equilibrated solvent bath. The referenced computational details [2–9] are provided in the caption for Fig. 2.1. Simulations are performed within an isobaric/isothermal ensemble (1 atm, 298 K). The optimized systems are pre-equilibrated for 500 ps. The resulting structures become the starting point for the 5 ns-thermalization trajectories. A total of 100 interfacial solvent configurations, one every 50 ps, are used to compute the time average of ϑ_n . To this end, we

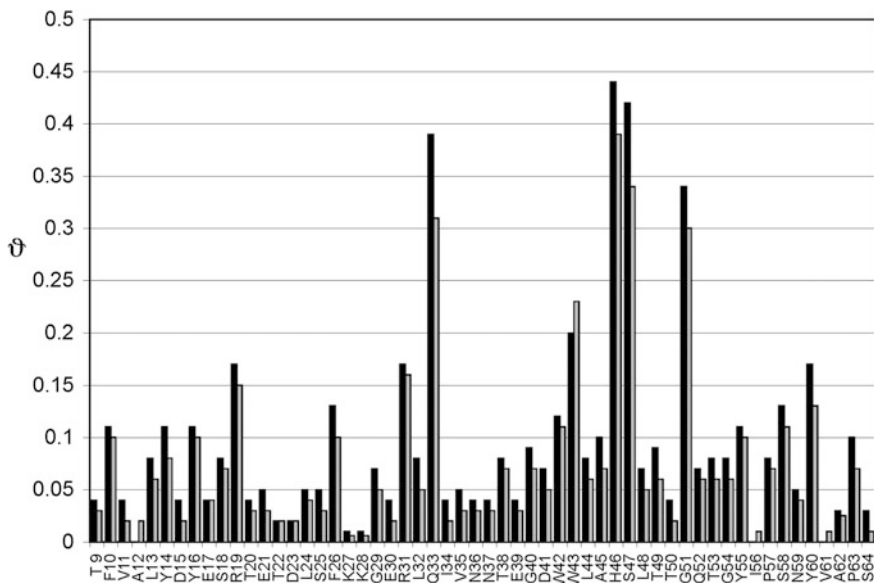


Fig. 2.1 Anomalous polarization fraction (ϑ) for each residue within the solvent-equilibrated folded protein chain for the soluble SH3 domain (PDB:1SRL). The PDB file numbering is followed in naming residues along the chain and the one-letter code for amino acids is adopted. *Black bars* represent protein surface interrogation with spheres of radius $r = 6$ Å centered at α -carbons, while *gray bars* were generated using radius $r = 8$ Å. The all-atom trajectories used to compute the time-averaged APF values thermalize the PDB structures in contact with a pre-equilibrated solvent bath consisting in a truncated octahedral cell of TIP3P water molecules that provide at least four water layers of solvent envelope [2]. Protein atoms are described with the parm99SB force field parameterization [3]. Water molecules extended at least 12 Å from the surface of the protein. Ewald sums [4] and an 8 Å-distance cutoff are used for treating long-range electrostatic interactions. A Shake scheme is employed to keep bonds involving hydrogen atoms at their equilibrium length [5] which allowed us to employ a 3 fs time step for the integration of Newton’s equations. Constant pressure of 1 atm and a temperature of 298 K are maintained using the Berendsen coupling scheme [6]. An AMBER package [7] was adopted for these MD simulations, with charges on the molecules assigned according to the BCC charge model using AM1 optimized geometries and potentials [8, 9]. After protein/solvent equilibration (as defined in main text), the protein backbone coordinates are partially constrained according to the Shake scheme [5] and only side chains are allowed to explore conformation space, generating a gamut of local hydration patterns. Reprinted from [Fernández Stigliano A (2013) Breakdown of the Debye polarization ansatz at protein–water interfaces. J Chem Phys 138:225103], copyright 2013 with permission from AIP Publishing LLC

recorded charge distribution $\rho(\vec{r}, t)$, internal field $\vec{E}(\vec{r}, t)$, and polarization $\vec{P}(\vec{r}, t)$ for each intermediate structure/solvent configuration.

The structure/solvent system is considered equilibrated at time t_0 if the RMSD of backbone atomic coordinates averaged over randomly chosen pairs of chain conformations within a time interval $[t_0, t_0 + \tau]$ ($\tau \approx 1$ ns) is less than 1 Å. For all nine proteins in this study (PDB entries 1SRL, 1ESR, 1A8O, 1PIT, 1QGB, 1ATA, 1Q7I, 1PI2, 2PNE), this criterion was fulfilled for $t_0 = 500$ ps. Solvent and side-chain conformations continue to vary significantly (i.e., RMSD > 2.25 Å) on the 1 ns timescale.

The APFs for individual residues for the natively folded SH3 domain (PDB.1SRL) are shown in Fig. 2.1. The context-dependence of APFs is evident since residues of the same type (i.e., serines S18 and S47, tryptophans W42 and W43) can have very different APFs depending on their location within the protein chain and therefore, within the structure. If we exclude the residues A12, I56, and V61 that are fully buried within the structure, it is clear that the positively charged lysines (K27, K28) have the lowest APFs. As described below, this is expected since the ammonium cation ($-\text{NH}_3^+$) in lysine has the highest charge concentration of all amino acids and hence it is the most capable of organizing solvent in accord with its highest hydration requirements.

Intriguingly, a structural context becomes the determinant factor for the APF high values, superseding individual residue propensities. Thus, the residues with the highest APFs, Q33, H46, S47, and G51, are the only ones paired by *dehydrons* (marked in green in Fig. 2.2a; the ribbon rendering in Fig. 2.2b is an aid to the eye).

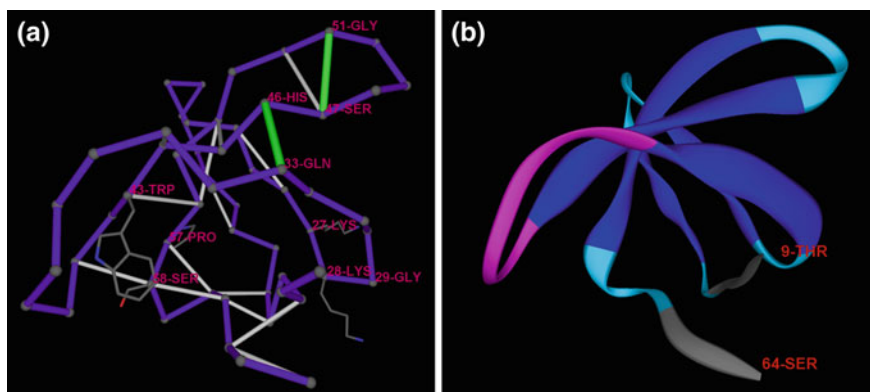


Fig. 2.2 **a** Dehydron pattern for the solvent-equilibrated structure of the soluble Src-SH3 domain (PDB.1SRL). The backbone is represented as virtual bonds joining the α -carbons of residues along the chain, with well-shielded backbone hydrogen bonds and dehydrons shown as segments sustained between the paired residues in *grey* and *green*, respectively. Dehydrons are determined from the PDB structural coordinates following the protocol indicated in Chap. 1. In accord with this protocol, the underwrapping of the backbone hydrogen bond due to a low number of surrounding nonpolar groups from the flanking side chains is considered to be a surrogate for the extent of solvent exposure. **b** Ribbon rendering of the structure of Src-SH3 domain

Due to the nanoscale water confinement created by the packing defect that gives rise to the dehydron, a significant nonvanishing component $\vec{P}^\#$ is expected for residues paired by such hydrogen bonds (cf. Chap. 1). These packing defects expose the backbone polar groups amide ($>\text{N-H}$) and carbonyl ($>\text{C=O}$) to structure-disruptive effects of backbone hydration with the net effect of steering water dipoles into orientations that are not collinear with \vec{E} . The confined water molecules relinquish some of their hydrogen-bonding possibilities in order to form hydrogen bonds with the backbone polar groups. This reduction in coordination represents a departure from the bulk water structure embodied in the tetrahedral lattice and the resulting water polarization becomes statistically independent of the internal field \vec{E} . This is so since water molecules with reduced water coordination tend to preserve their hydrogen bond pattern thereby becoming impervious to the torque

$$\vec{E}(\vec{r}) \times \vec{\mu}_p(\vec{r}) = \vec{E}(\vec{r}) \times \int (\vec{r}' - \vec{r}) \vec{\nabla} \cdot \vec{P}(\vec{r}') d\vec{r}' \quad (2.4)$$

imposed by $\vec{E} = \vec{E}(\vec{r})$ on the polarization-associated dipole with moment

$$\vec{\mu}_p(\vec{r}) = \int (\vec{r}' - \vec{r}) \vec{\nabla} \cdot \vec{P}(\vec{r}') d\vec{r}'.$$

Thus, interfacial water polarization in this context is expected to contain and indeed contains (Fig. 2.1) a significant anomalous non-Debye contribution.

2.2 Non-Debye Polarization of the Aqueous Interface of a Soluble Protein

In Fig. 2.3a, we show the individual propensities of the 20 residue types to align interfacial water along the electrostatic field by computing the APF of residues in nine PDB-reported proteins (specified in caption for Fig. 2.4). The expected APF for each residue type is obtained by averaging the APFs for that residue type in all nine proteins. The computation amounts to average over the structural contexts in the nine proteins where the particular residue type occurs. Due to the dominance of dehydrons as structural determinants of APF (Figs. 2.1 and 2.4), superseding individual propensities (Fig. 2.1), we have excluded dehydron-paired residues from the calculations in Fig. 2.3a.

As a class, the aromatic residues (H, F, W, Y) have the highest APF values due to their water-organizing power and their role as significant disruptors of the tetrahedral water structure. Their delocalized π -electron quadrupole promotes interactions with partial positive charges in vicinal interfacial water molecules. Furthermore, the side chains of such residues cannot be clathrated (surrounded

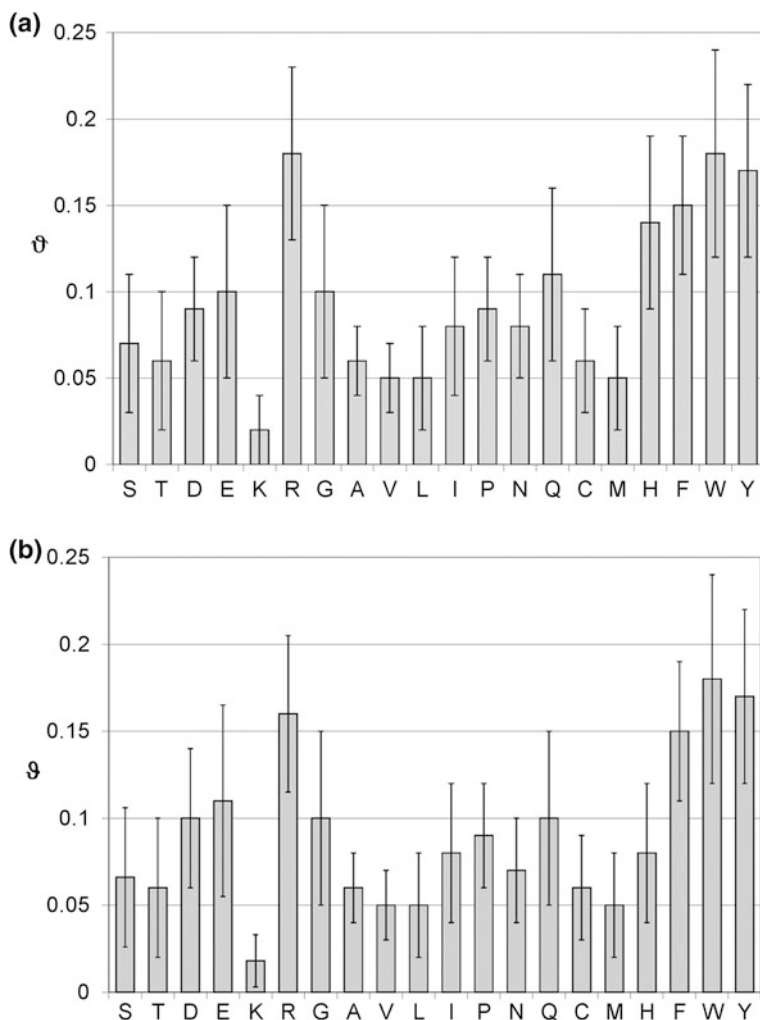


Fig. 2.3 APFs for each residue type averaged over all structural environments where the residue type occurs in nine PDB-reported proteins described in the caption for Fig. 2.4. The radius $r = 6 \text{ \AA}$ has been adopted and the error bars represent the dispersion in ϑ -values. **a** APFs at pH 7. **b** APFs at pH 5.5. Reprinted from [Fernández Stigliano A (2013) Breakdown of the Debye polarization ansatz at protein–water interfaces. *J Chem Phys* 138:225103], copyright 2013 with permission from AIP Publishing LLC

without disrupting the tetrahedral water lattice) as it is the case with nonpolar aliphatic side chains (L, V, I, A). Thus, the resilient nontetrahedral hydrogen bond pattern of vicinal water explains the superior APF-boosting activity of aromatic residues when compared with nonpolar aliphatic ones. The sharp contrast between the lowest APF-booster lysine (K) and the highest APF-booster arginine (R), both

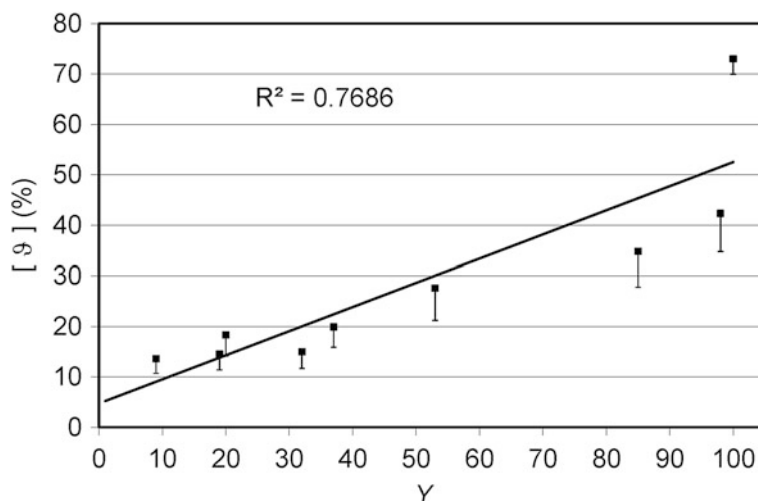


Fig. 2.4 Correlation between $\bar{\varrho}$ -value averaged over all residues in the protein ($[\bar{\varrho}]$), and the dehydron-to-backbone-hydrogen-bond ratio Y for the protein. Both parameters are given as percentages. The proteins studied identified by their respective PDB files and Y -ratios (in *brackets*) are: 1SRL (14.3), 1ESR (27.9), 1A8O (32.1), 1PIT (40.0), 1QGB (48.6), 1ATA (57.7), 1Q7I (70.1), 1PI2 (98.0), 2PNE (100.0). The average APF values indicated by the *filled squares* correspond to $r = 6$ Å. The *low error bar* indicates the net decrease in APF as bulk solvent is approached when adopting $r = 8$ Å. Reprinted from [Fernández Stigliano A (2013) Breakdown of the Debye polarization ansatz at protein–water interfaces. J Chem Phys 138:225103], copyright 2013 with permission from AIP Publishing LLC

in the same class of positively charged residues with aliphatic (methylene) linkages, is also striking, yet expected. The ammonium cation ($-\text{NH}_3^+$) in lysine has the highest charge concentration of all amino acids; therefore, it strongly organizes hydration along electrostatic field lines, while the guanidinium cation ($[-\text{NH}=\text{C}(\text{NH}_2)_2]^+$) in arginine contains the most delocalized charge of all amino acids, hence the resulting local electrostatic field has the weakest water-organizing power.

From the above discussion, it becomes apparent that the polarization-steering power of individual residues is tightly related to the localization and concentration of their net charge. Thus, a pH dependence of the APF for an individual residue is expected in accord with the pK_a of the residue within the protein structure. The titration of a residue removes a net charge and thereby increases the APF by curbing the polarization-steering capabilities of the residue. This titration effect becomes apparent as we compare the expected APFs of individual residue types at neutral pH (Fig. 2.3a) and pH 5.5 (Fig. 2.3b). The pH window 5.5–7 is apparent when contrasting Fig. 2.3a, b contains only the $pK_a \approx 6.1$ of histidine (H), and hence this residue is predicted and shown to undergo the most dramatic gain in polarization steering (decrease in expected APF) as pH is decreased from 7 to 5.5. The effects of titration on other residue types could not be assessed in this study since their pK_a 's

dictate extreme pH values that would introduce denaturing conditions for the proteins studied.

A significant correlation is established for PDB-reported proteins (Fig. 2.4) between the average APF ($[g]$) overall residues in a protein and the protein ratio Y of dehydrons-to-backbone-hydrogen bonds. This correlation validates the assertion that dehydrons are the main structural motif promoting anomalous polarization. At $Y = 100\%$, the antifreeze protein from snow flea in PDB entry 2PNE [10], with its anomalously high APF-boosting activity, is a significant outlier. This enhanced effect can be understood based on the extreme solvent exposure of its dehydrons, promoting a local backbone-hydrated state that persists on a 100 ns timescale, compared with the ~ 1 ns lifetime of the hydrated state typical of the dehydrons present in the other proteins studied. Thus, the water-organizing power of the antifreeze protein is due to nanoscale confinement and supersedes the Debye polarization tendencies, introducing a major suprananoscale perturbation of the water structure, in accord with its purported function as a disruptor of the ice nucleation.

This connection between disruption of ice nucleation and anomalous polarization suggests mutational studies aimed at removing dehydrons by improving the backbone protection in antifreeze proteins. We predict that the removal of dehydrons by backbone-protective valine (V) substitutions of poor backbone protector residues (S, T, G) should significantly impair the antifreeze potency of the snow flea protein reported in PDB.2PNE.

2.3 Chemical Functionality of the Aqueous Interface: A Consequence of the Breakdown of Debye's Dielectric Picture

This section carries the analysis of the polarization structure of the aqueous interface one step further by showing that dehydrons not only promote protein associations but also functionalize interfacial water by inducing basicity at the interface. These packing defects confine interfacial water molecules turning them into proton acceptors. This result has profound ramifications for bioengineering and drug design as it implies that dehydrons are actually involved in chemical events, acting as stimulators of enzymatic activity.

To make notation more agile, we introduce aligned (Γ^{\parallel}) and orthogonal ($\Gamma^{\#}$) polarization-induced charges defined as

$$-\vec{\nabla} \cdot \vec{P}^{\parallel} = \Gamma^{\parallel}; -\vec{\nabla} \cdot \vec{P}^{\#} = \Gamma^{\#}; \Gamma^{\parallel} + \Gamma^{\#} = \Gamma \quad (2.5)$$

Departures from bulk water structure (spatially measured by $\nabla g \neq 0$) induce orthogonal polarization. Given the relation $\vec{P}^{\#} = \xi \vec{\nabla} g$, the Poisson equation $\vec{\nabla} \cdot \left(\vec{P}^{\#} \right) = \rho - \vec{\nabla} \cdot \left(\epsilon_0 \vec{E} + \vec{P}^{\parallel} \right)$ may be written in terms of the curvature $\nabla^2 g$ of the scalar field g :

$$\zeta \nabla^2 g = -\Gamma^\# \quad (2.6)$$

Equation (2.6) incorporates the nanoscale structure of water within an electrostatic relation, revealing that the curvature of the scalar field g is a measure of the departure from linear dielectrics.

We also introduce the “frustration” scalar field $\phi(\vec{r}) = 4 - g(\vec{r})$ that quantifies the extent of distortion from bulk-like water structure, with $\phi = 0$ representing no distortion (no hydrogen-bond opportunity has been frustrated). To obtain a partial differential equation for ϕ , we first note that at each position \vec{r} , the quotient $\Gamma^\#/\Gamma$ measures the local deviation from a Debye scenario where polarization fully aligns with the protein field (Fig. 2.5). Thus, we expect that a relation of the form $\frac{\Gamma^\#}{\Gamma} = c\phi$ must hold, where c is a proportionality constant. This relation is indeed valid with $c = 0.191$ as shown in Fig. 2.6a. The quotient $\Gamma^\#/\Gamma$ is computed at protein/water interfaces as a time average over a 10 ns period beyond equilibration of the protein structure with the solvent. Thus, the epistemic polarization \vec{P} for nine soluble natural proteins with structures reported in the Protein Data Bank (Table 2.1, Protocol in legend for Fig. 2.1) is computed along thermalization molecular dynamics

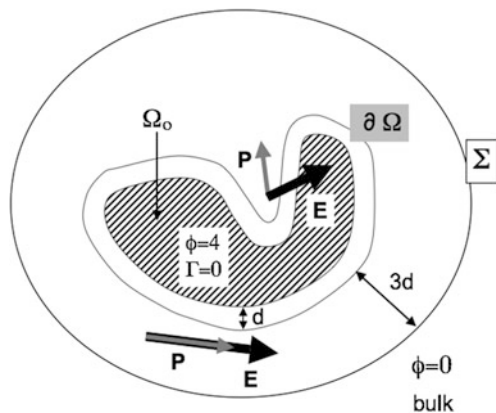


Fig. 2.5 Schematic depiction of the aqueous interface with boundary conditions that become constraints in determining the coarse-grained “frustration” descriptor $\phi = 4 - g$ of water structure. The interface is defined by the water-smear envelope $\partial\Omega$ of the solute. It is assumed that the solute surface has a topography endowed with nanoscale detail that may perturb the structure of interfacial water relative to bulk-like patterns. We get $\phi = 4$ (no water) in the “core” volume Ω_0 at distance $d = 2r = 8 \text{ \AA}$ from each point on $\partial\Omega$, while $\phi = 0$ for points at distance larger than $3d = 24 \text{ \AA}$ from $\partial\Omega$. The latter condition holds since $3d \gg 4$ water layers ($\sim 13 \text{ \AA}$) from the interface and hence in this region, water structure is assumed to have recovered its bulk-like pattern. The regions where the structure of interfacial water is relatively undistorted show an alignment between epistemic polarization \vec{P} and the electrostatic field \vec{E} , whereas regions of high structural distortion likely depart from the linear dielectrics picture, as reflected by a lack of alignment between the fields \vec{P} and \vec{E} . Reprinted from [1], copyright 2013 with permission from AIP Publishing LLC

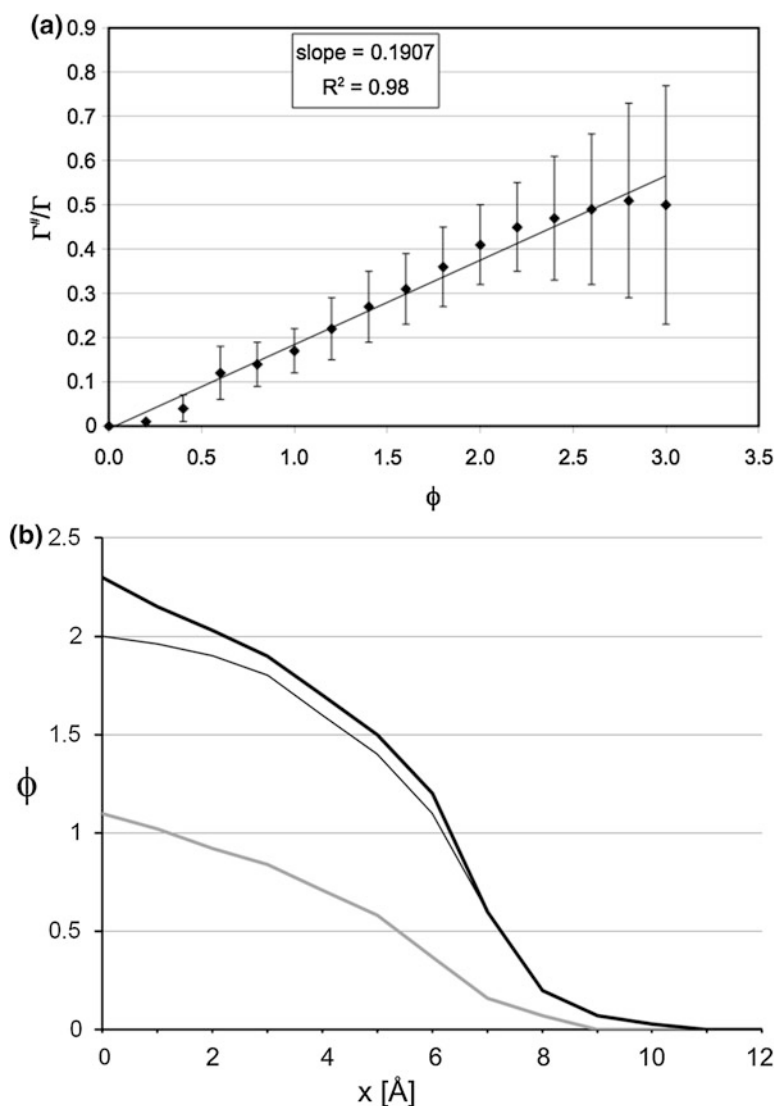


Fig. 2.6 **a** Local deviation from \bar{E} -aligned polarization measured as $\Gamma^\#/\Gamma$ and plotted against the structural function $\phi = 4 - g$. The quotient $\Gamma^\#/\Gamma$ is evaluated at protein/water interfaces as a time average over an interval beyond protein structure/solvent equilibration. To determine $\Gamma^\#/\Gamma$, the epistemic polarization \bar{P} for nine soluble natural proteins with structures reported in the Protein Data Bank (Table 2.1) is computed and averaged for each position in space along a thermalization molecular dynamics trajectory. The region spanning a distance of $3d = 24$ Å from $\partial\Omega$ (Fig. 2.5) is exhaustively interrogated by covering it with disjoint spheres of radius 4 Å centered at points $\bar{r}_n (n = 1, 2, \dots)$ to determine the set of water structure values $\phi(\bar{r}_n)$. The trend line and correlation coefficient were obtained by linear regression. **b** Behavior of $\phi(x)$ relative to the distance x to an interface. The interface is assumed nonpolar, consisting of a concave region of fixed curvature radius θ , with $x = 0 =$ center of curvature. The *flat surface* represents the macroscopic limit $\theta \gg 1$ nm, and the ϕ -values are shown in the *gray plot*. Other perturbations were obtained for $\theta = 3$ Å (*thin black plot*), and $\theta = 2.5$ Å (*thick black plot*). Reprinted from [1], copyright 2013 with permission from AIP Publishing LLC

Table 2.1 PDB accession codes and free energy changes (ΔG) for the thermal denaturation of soluble monomeric proteins with reported structures (Fernández A (2012) Communication: Nanoscale electrostatic theory of episturctural fields at the protein–water interface. J Chem Phys 137:231101, and references therein)

PDB accession code	ΔG (kJ/mol)	T (C)
1BSQ	46.46	40.00
1RTB	42.28	25.00
4LYZ	37.76	26.85
1CX1	22.52	24.85
1QG5	36.84	40.00
2AIT	28.05	25.00
3SSI	17.04	20.00
1HIC	21.01	25.00
1PMC	4.60	20.00

The temperature (T) for thermal denaturation is indicated

trajectories. Each 10 ns-trajectory is generated using as starting point the equilibrated structural coordinates that result after thermalization of the PDB-reported structure immersed in a pre-equilibrated solvent bath. Simulations are performed within an isobaric/isothermal ensemble (1 atm, 298 K). A total of 100 interfacial solvent configurations, one per 100 ps along a 10 ns-thermalization trajectory, are used to compute the epistemic polarization quotient as a time average using the relation

$$\frac{\Gamma^{\#}}{\Gamma} = \frac{\rho - \vec{\nabla} \cdot [\epsilon_0 \vec{E} + P^{\parallel}]}{\rho - \vec{\nabla} \cdot \epsilon_0 \vec{E}} \quad (2.7)$$

To this end, we recorded charge distribution, internal field, and polarization from the 100 snapshots that partition the 10 ns period in identical intervals.

The structure/solvent system is considered equilibrated at time t_0 if the RMSD of backbone atomic coordinates averaged over randomly chosen pairs of chain conformations within a time interval $[t_0, t_0 + \tau]$ ($\tau \approx 1$ ns) is less than 1 Å. For all nine proteins in this study, this criterion was fulfilled for $t_0 = 500$ ps. Solvent and side-chain conformations continue to vary significantly (i.e. RMSD > 2.25 Å) on the 1 ns timescale.

By introducing the relation $\frac{\Gamma^{\#}}{\Gamma} = c\phi$, and defining the constant $k = c/\xi$, (2.6) becomes a linear homogeneous differential equation in $\phi(\vec{r})$ of the Schrödinger type:

$$-\nabla^2 \phi + k\Gamma \phi = 0 \quad (2.8)$$

The boundary conditions are described in Fig. 2.5 and are determined by the water-smearred envelope $\partial\Omega$ of the solute–protein interface. Thus, $\phi = 4$ (no water) in the “core” volume Ω_0 at distance $d = 2r = 8$ Å from each point on $\partial\Omega$, while $\phi = 4$ for points at distance larger than $3d = 24$ Å from $\partial\Omega$. The latter condition holds since $3d \gg 4$ water layers (~ 13 Å) from the interface and hence in this

region, water structure is assumed to have recovered its bulk-like tetrahedral pattern.

The behavior of the frustration field $\phi = \phi(\vec{r})$ relative to the distance x to an interface is indicative of the propagation in space of the distortion of the water structure and is shown in Fig. 2.6b. The results were obtained by numerical integration of (2.8) for structural perturbations generated by confinement of water at the interface. To eliminate confounding factors, the interface was assumed nonpolar and physicochemically featureless, consisting of a concave region of fixed curvature radius q , capable of partially confining water molecules, with $x = 0$ representing the center of curvature of the surface. The flat surface represents the macroscopic limit $\theta \gg 1$ nm, and the interfacial ϕ -value is expectedly close to 1 ($g \sim 3$), as revealed by the gray plot in Fig. 2.6b. Other perturbations were obtained for $\theta = 3$ Å (thin black plot), and $\theta = 2.5$ Å (thick black plot), generating $\phi(0)$ -values 2.01 and 2.30, respectively. In all cases, the matrix distortion decays to zero for $x > 11$ Å (less than 4 water layers). No water molecule enters the cavity if doing so implies that the molecule retains on average less than 1.6 hydrogen bonds, making the range $\theta < 2.37$ Å, a forbidden region in real terms.

Equation (2.8) is the central result of this section and governs the interplay between epistemic polarization and the nanoscale structure of interfacial water.

We now compute the energy increment ΔU_ϕ associated with spanning a protein/water interface. This energy is in fact an elastic contribution stored in the distortion of water structure as shown in Chap. 1, with $\nabla\phi(\vec{r})$ measuring the local structural distortion at position \vec{r} and the elastic integrand of $\int \frac{1}{2} \lambda \left\| \vec{\nabla}\phi \right\|^2 d\vec{r}$ ($\lambda = 9.0$ mJ/m) quantifying the energetic contribution of spanning a differential region $d\vec{r}$ centered at point \vec{r} . Using Gauss' divergence theorem, we obtain:

$$\Delta U_\phi = \int \frac{1}{2} \lambda \left\| \vec{\nabla}\phi \right\|^2 d\vec{r} = -\frac{1}{2} \lambda \int \phi \nabla^2 \phi d\vec{r} \quad (2.9)$$

Using (2.8), we can substitute the integrand in the r.h.s of (2.9) obtaining the alternative expression for the interfacial elastic energy:

$$\Delta U_\phi = -\frac{1}{2} c(\lambda/\epsilon_0)^{1/2} \int \Gamma \phi^2 d\vec{r} \quad (2.10)$$

Since $\Delta U_\phi = \int \frac{1}{2} \lambda \left\| \vec{\nabla}\phi \right\|^2 d\vec{r} \geq 0$ ($\lambda > 0$), the r.h.s of (2.10) is a positive term, and thus (2.9) and (2.10) imply:

$$\int \Gamma \phi^2 d\vec{r} \leq 0. \quad (2.11)$$

Equation (2.11) implies that $\Gamma \leq 0$, that is, the polarization-induced charge is negative around dehydrons known to promote interfacial tension. This is a fundamental result of broad applicability and it establishes the following

Theorem *Interfacial water molecules tend to orient and organize in subnanoscale cavities leaving negative charges uncompensated ($\Gamma < 0$) when deprived of hydrogen-bonding opportunities ($\phi > 0$).*

This is a crucial result as it delineates the chemical basicity of dehydrons, or more properly, the dehydron-induced basicity of interfacial water. The theorem reveals that packing defects play a crucial role in enzyme catalysis and in biochemical events in general by functionalizing nucleophilic protein groups through dehydron-promoting proton accepting events. This type of catalytic stimulation requires that dehydrons be spatially close to catalytic nucleophilic groups and that the proton transfer event induced by the dehydron have a reasonable probability. These striking aspects of the functionalization of biomolecular interfaces will be explored and ultimately established in Chap. 7.

The interfacial energy stored in the anomalous polarization or, equivalently, in the distortion of water structure, is readily evaluated using (2.10). This result has been contrasted against thermodynamic data on the spanning of aqueous interfaces with nanoscale detail. A suitable testing ground is provided by the aqueous interfaces for soluble monomeric proteins with a stable fold characterized by structural and thermodynamic information (Table 2.1) [1]. The reversible work performed on the system to span the protein–water interface is destabilizing of the native fold, thus facilitating thermal denaturation, and hence it should anticorrelate with the free energy change for protein denaturation, as it is indeed the case (cf. [1]). To compare interfacial thermodynamics with thermodynamic data on protein denaturation (Table 2.1), we introduce the entropic cost of solvent confinement at the interface $\Delta S_\phi = k_B \ln \left[\prod_{j=1}^L g_j / 4 \right]$, where k_B = Boltzmann constant, g_j = time-averaged number of hydrogen bonds for the j th—water molecule (L = total number of water molecules), and the dummy index j labels molecules within $3d = 24$ Å from the solvent-smeared envelope of the protein (Fig. 2.5). Note that $-T\Delta S_\phi \geq 0$ and reinforces the trend defined by ΔU_ϕ . The reversible work $\Delta G_\phi = \Delta U_\phi - T\Delta S_\phi \geq 0$ performed on the system to span the protein–water interface is destabilizing of the native fold, thus facilitating thermal denaturation.

To quantitatively assess the folding-destabilizing effects of spanning the interface of the folded protein, we examined the same soluble monomeric proteins used to generate the data in Fig. 2.6 (Table 2.1). In thermodynamic terms, protein denaturation is facilitated proportionally to the reversible work required to span the interface, attesting to the folding-destabilizing effect of interfacial tension arising from the structural distortion of surrounding water. Thus, the computed reversible work for creating the interface measures the extent to which the “protein structure is at odds with the structure of surrounding water” since it quantifies the distortion of

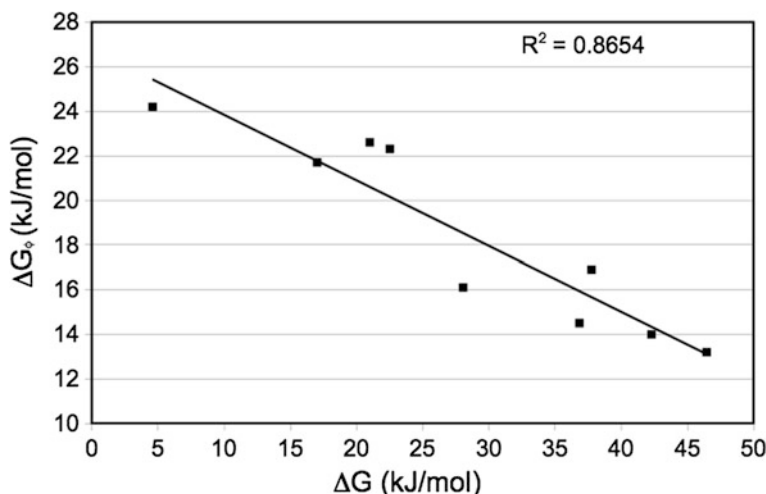


Fig. 2.7 Anticorrelation between reversible work required to span the protein–water interface, ΔG_ϕ , and free energy change for thermal denaturation, ΔG , for the nine monomeric soluble proteins indicated in Table 2.1. Reprinted from [1], copyright 2013 with permission from AIP Publishing LLC

water structure around the protein. The tight anticorrelation between interfacial free energy and the stability of the protein structure (Fig. 2.7) provides experimental support to the underlying Eq. (2.8) since it reveals that protein destabilization is commensurate with the thermodynamic cost of creating its interface with water, computed using (2.10). This observation prompts us to formulate the principle of minimal epistemic distortion (MED) that should govern conformational changes in the solute that generate concomitant changes in the interface. The validity of the MED principle in the context of protein folding is established in Chap. 3 that provides a semiempirical solution to the protein folding problem.

2.4 A New Twist on Enzyme Catalysis: Nanoscale Packing Defects as Catalytic Stimulators

Enzyme catalysis is often viewed as a closed chapter where the core issues have already been dealt with [11, 12]. Yet, several open problems still stand on the way of progress in mechanistic understanding [13–15], and the design and optimization of enzyme catalysts [15, 16] and drug-based enzyme inhibitors [16–18]. Especially, opaque yet germane to these problems is the role of interfacial water in enzymatic reactions [13, 19, 20]. Because the protein aqueous interface is essentially sculpted by the protein structure [19, 20], the problem may be said to belong to the field of *epiststructural biology*, as argued in Chap. 1. In this realm, as we have already noted,

one structural feature of proteins stands out: the so-called *dehydron*, a packing defect that creates interfacial tension and thereby promotes protein associations that exclude surrounding water [20, 21]. Thus, the recently established fact that catalytic sites in enzymes are actually “decorated” with dehydrons [13] proves to be tale-telling and gives a significant spin to related biotechnologies. These observations will be properly delineated, expanded, and validated in Chap. 7.

As implied by the theorem formulated in Sect. 2.3 (cf. [14]), besides promoting dehydration, dehydrons are also likely to be endowed with a biochemical role that may prove to be exquisitely complementary: they turn local interfacial water into a chemical base, a proton acceptor. Thus, if a catalytic group (hydroxyl in Ser, Thr or Tyr, thiol in Cys, amide in His) performs a nucleophilic attack on a substrate, the dehydrons nearby enhance its catalytic potential through a chemical functionalization of vicinal water that promotes deprotonation of the catalytic group.

The dual functional and interactive role of dehydrons combined with the fact that they are ubiquitously found at catalytic sites suggest a dual participation in catalysis: *first, dehydrons prepare the solvent for enzyme activity and, once the enzymatic nucleophilicity is enhanced and the solvent turns into a better leaving group (as hydronium), dehydrons promote enzyme-substrate association in consonance with their dehydration propensity*. This duality of functionality and interactivity makes dehydrons both enablers and stimulators of enzyme catalysis, an observation with paramount biotechnological implications, especially in regards to what we may term “epistucture-based enzyme design”.

In more rigorous terms, as dehydrons activate nearby catalytic groups to perform a chemical (nucleophilic) attack on the substrate, causing transesterification, they turn the local water into hydronium (H_3O^+ , a product of proton acceptance). In turn, the hydronium is easily removed from the active since it requires further hydration, thereby enabling enzyme–substrate association. This association process entails the exogenous “wrapping” of the dehydron, which is tantamount to the intermolecular correction of the structural defect [21]. Thus, *the dehydron may be regarded as a two-stroke molecular engine that agonizes and enables enzyme catalysis*, as described in Fig. 2.8.

This discovery heralds the advent of novel biomolecular design based on “dehydron enablers-stimulators” that may be created or removed through engineered mutations directed to fine-tune the protein structure. This finding makes it possible to activate or silence a catalytic site in a protein enzyme by, respectively, creating or annihilating a nearby dehydron through a change in the chemical composition of the protein. On the other hand, novel drug-based enzyme inhibitors will emerge as dehydron enablers–agonists are targeted through engineered protein–drug associations [18], as described in Chap. 10.

The newly established participation of dehydrons in enzymatic reactions will likely invite an extensive revision of the biochemical mechanistic literature, while novel molecular designs inspired by “epistuctural catalytic stimulation” are expected to herald a new era in the optimization of enzyme catalysts [15, 22] and pharmaceuticals.

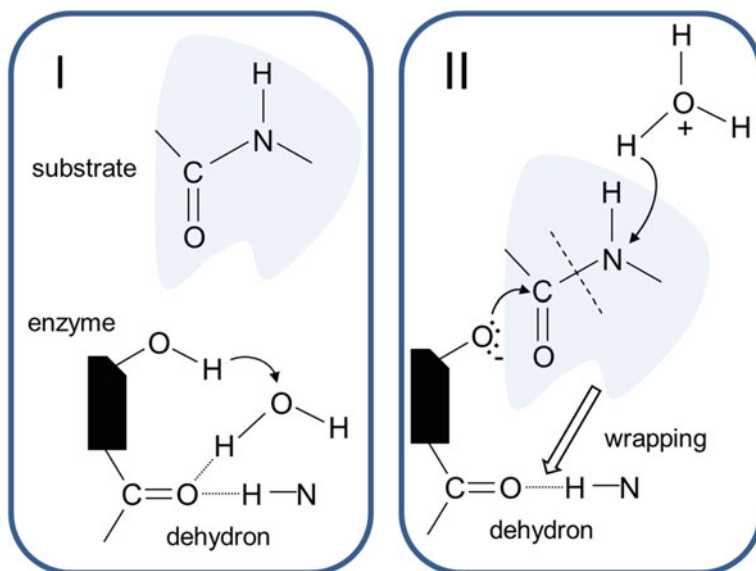


Fig. 2.8 Dehydron as two-stroke molecular engine sustaining enzyme catalysis by (I) functionalizing nano-confined water around the catalytic group and (II) promoting enzyme–substrate association, thereby correcting the nanoscale packing defect

Problems

2.1. Using Gauss' divergence theorem, prove the following relation:

$$\Delta U_\phi = \int \frac{1}{2} \lambda \left\| \vec{\nabla} \phi \right\|^2 d\vec{r} = -\frac{1}{2} \lambda \int \phi \nabla^2 \phi d\vec{r}$$

- 2.2. Water promotes the sealing of nanoscale packing defects in proteins. This is hinted by the fact that during protein folding, a minimization of the departure from the Debye alignment is observed. How does the solvent promote the structure sealing? Assume a flexible chain that seeks to minimize interfacial free energy.
- 2.3. Using the protocol provided in this chapter, thermalize the monomeric soluble proteins with PDB-reported structure identified by entries 1SRL, 1ESR, 1A8O, 1PIT, 1QGB, 1ATA, 1Q7I, 1PI2, 2PNE, and validate the following empirical relation indicated in Chap. 1 concerning the density fluctuations:

$$P(N = 0) = e^{-\frac{\langle N^2 \rangle}{2\sigma^2}} = 1 - g/4, \quad \sigma^2 = \left\langle (N - \langle N \rangle)^2 \right\rangle.$$

References

1. Fernández A (2013) The principle of minimal epistemic distortion of the water matrix and its steering role in protein folding. *J Chem Phys* 139:085101
2. Jorgensen WL, Chandrasekhar J, Madura J, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
3. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct Funct Bioinf* 65:712–725
4. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N.Log(N) method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092
5. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical-integration of cartesian equations of motion of a system with constraints: molecular dynamics of N-alkanes. *J Comput Phys* 23:327–341
6. Berendsen HJ, Postma JP, van Gunsteren WF, Di Nola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690
7. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general Amber force field. *J Comput Chem* 25:1157–1174
8. Jakalian A, Bush BL, Jack DB, Bayly CI (2000) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J Comput Chem* 21:132–146
9. Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* 23:1623–1641
10. Fernández A, Berry RS (2009) Golden rule for buttressing vulnerable soluble proteins. *J Proteome Res* 9:2643–2648
11. Frushicheva MP, Mills MJ, Schopf P, Singh MK, Prasad RB, Warshel A (2014) Computer aided enzyme design and catalytic concepts. *Curr Opin Chem Biol* 21C:56–62
12. Fersht A (1998) Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding. Freeman, New York
13. Fernández A (2014) Communication: chemical functionality of interfacial water enveloping nanoscale structural defects in proteins. *J Chem Phys* 140:221102
14. Frauenfelder H (2008) What determines the speed limit on enzyme catalysis? *Nat Chem Biol* 4:21–22
15. Korendovych IV, DeGrado WF (2014) Catalytic efficiency of designed catalytic proteins. *Curr Opin Struct Biol* 27C:113–121
16. Fernández A, Rogale K, Scott RL, Scheraga HA (2004) Inhibitor design by wrapping packing defects in HIV-1 proteins. *Proc Natl Acad Sci USA* 101:11640–11645
17. Walsh R, Martin E, Darvesh S (2011) Limitations of conventional inhibitor classifications. *Integr Biol* 3:1197–1201
18. Fernández A (2010) Transformative concepts for drug design: target wrapping. Springer, Berlin
19. Kapcha LH, Rossky PJ (2014) A simple atomic-level hydrophobicity scale reveals protein interfacial structure. *J Mol Biol* 426:484–498
20. Fernández A (2012) Epistructural tension promotes protein associations. *Phys Rev Lett* 108:188102
21. Fernández A, Lynch M (2011) Non-adaptive origins of interactome complexity. *Nature* 474:502–505
22. Privett HK, Kiss G, Lee TM, Blomberg R, Chica RA, Thomas LM, Hilvert D, Houk KN, Mayo SL (2012) Iterative approach to computational enzyme design. *Proc Natl Acad Sci USA* 109:3790–3795

Biomolecular Interfaces

Interactions, Functions and Drug Design

Fernandez Stigliano, A.

2015, XIX, 372 p. 145 illus., 59 illus. in color., Hardcover

ISBN: 978-3-319-16849-4