Breakdown of the Debye polarization ansatz at protein-water interfaces

Ariel Fernández Stigliano

Citation: J. Chem. Phys. **138**, 225103 (2013); doi: 10.1063/1.4810867
View online: http://dx.doi.org/10.1063/1.4810867
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v138/i22
Published by the AIP Publishing LLC.

Breakdown of the Debye polarization *ansatz* at protein-water interfaces

Ariel Fernández Stigliano[a),b)]
Instituto Argentino de Matemática "Alberto P. Calderón," CONICET (National Research Council), Saavedra 15, Buenos Aires 1083, Argentina; Collegium Basilea, Institute for Advanced Study, Hochstrasse 51, CH 4053 Basel, Switzerland; and AF Innovation, Pharmaceutical Consultancy, Avenida del Libertador 1092, Piso 1, Buenos Aires 1112, Argentina

(Received 29 March 2013; accepted 24 May 2013; published online 14 June 2013)

The topographical and physico-chemical complexity of protein-water interfaces scales down to the sub-nanoscale range. At this level of confinement, we demonstrate that the dielectric structure of interfacial water entails a breakdown of the Debye *ansatz* that postulates the alignment of polarization with the protein electrostatic field. The tendencies to promote anomalous polarization are determined for each residue type and a particular kind of structural defect is shown to provide the predominant causal context. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4810867]

I. INTRODUCTION

At the time when Debye's "molecular theory" of dielectric response[1] emerged, the nanoscale structure of water, with its resilient tetrahedral lattice of intermolecular hydrogen bonds,[2–6] was largely unknown and the importance of hydrogen bonds as determinants of water structure was not yet recognized. With virtually no information on the nanoscale structure of water and merely assuming the existence of a polarizable "sea of dipoles," Peter Debye derived a theory applicable in contexts where the dielectric response could be treated as a bulk attribute. The theory postulates an alignment of water polarization $\mathbf{P}$ with the internal electrostatic field $\mathbf{E}$. This *ansatz* enables the computation of the effective permittivity ε of the dielectric medium by assuming that the proportionality between the vector fields $\mathbf{P}$ and $\mathbf{E}$ can be expressed without loss of generality as $\mathbf{P} = (\varepsilon - \varepsilon_o)\mathbf{E}$, where ε_o is the vacuum permittivity. The *ansatz* yields a simplified Poisson electrostatic equation $\nabla.(\varepsilon_o\mathbf{E} + \mathbf{P}) = \nabla.(\varepsilon\mathbf{E}) = \rho$ (ρ = fixed charge distribution) that holds for an electrostatic condenser with planar featureless parallel plates and for virtually all aqueous dielectric media provided it lacks nanoscale detail.

The success of Debye's simplification encouraged researchers to extrapolate his linear polarization assumption to biophysical contexts such as protein-water interfaces.[7–9] In these cases, the supramolecular structure of water and the nanoscale distortions relative to a bulk-like organization cannot be ignored. The theoretical results proved to be somewhat difficult to reconcile with experiment, as recent work suggests.[7,8] The crux of the problem is that confinement of water at the biological interface often creates dipole organizations that result from deprivation of hydrogen bonding opportunities, thereby creating uncompensated effective charges that are not a resultant of polarization along the direction of the internal field $\mathbf{E}$ of the protein.

In critically assessing the applicability of Debye's model to protein-water interfaces we first note that a single parameter, the effective permittivity coefficient ε, cannot quantify the dielectric response of water. Our computations reveal that there is a component of $\mathbf{P}$ generated by the nanoscale topography and chemical composition of the protein surface that offers resistance to the torque imposed by the field $\mathbf{E}$ on the water dipoles.

The biological interfaces confine water into patterns that cannot simultaneously minimize the disruption of the hydrogen-bond lattice and align the water dipoles with the electrostatic field. There are polarization components that do not align with the field and arise from tight local nanoscale organizations of water molecules deprived of hydrogen bonding opportunities.[6] This picture brings about a complex dielectric structure that cannot be subsumed into a single permittivity parameter. The nanoscale water structure must be incorporated into the electrostatic description of the resultant fields at interfaces that confine solvent to discrete levels.

II. WATER POLARIZATION UNDER NANOSCALE CONFINEMENT

We expect that nanoscale local organization will introduce a significant component of polarization, $\mathbf{P}^{\#}$, orthogonal to the protein internal field $\mathbf{E}$. This "anomalous" polarization is a predictable resultant of the partial hindrance in the alignment of water dipoles with the field $\mathbf{E}$ due to nanoscale confinement. The departure from the "linear dielectrics" picture is described by writing the Poisson equation as follows:[6]

$$\nabla.(\varepsilon_o\mathbf{E} + \mathbf{P}^{\|} + \mathbf{P}^{\#}) = \rho, \tag{1}$$

where, without loss of generality, $\mathbf{P}$ has been decomposed into a field-aligned component $\mathbf{P}^{\|} = (\mathbf{P}.\mathbf{e})\mathbf{e}$ ($e = \mathbf{E}/\|\mathbf{E}\|$) and a component $\mathbf{P}^{\#}$, orthogonal to $\mathbf{E}$. Accordingly, we obtain

$$\nabla.(\mathbf{P}^{\#}) = \rho - \nabla.[\varepsilon_o\mathbf{E} + \mathbf{P}^{\|}]. \tag{2}$$

a)Author to whom correspondence should be addressed. Electronic mail: ariel@afinnovation.com. Telephone: 54 11 4804 1711.
b)Formerly Ariel Fernández.

In Eq. (2) the term $\rho - \nabla.[\varepsilon_o \mathbf{E} + \mathbf{P}^{\|}]$ measures the departure from the Debye dielectrics since $\rho - \nabla.[\varepsilon_o \mathbf{E} + \mathbf{P}^{\|}] = 0$ when the Debye *ansatz* $\mathbf{P} = \mathbf{P}^{\|}$ is fulfilled.

Defining the electrostatic potential ϕ through: $\mathbf{E} + \varepsilon_o{}^{-1}\mathbf{P} = -\nabla\phi$, the total electrostatic energy becomes

$$U = -(1/2)\varepsilon_o \int \phi \Delta\phi \, d\mathbf{r} = (1/2)\varepsilon_o \int \|\nabla\phi\|^2 \, d\mathbf{r} = U_D + U^{\#}, \tag{3}$$

where $\Delta = \nabla^2$ is the Laplace operator, and

$$U_D = (1/2)\varepsilon_o \int \|\mathbf{E} + \varepsilon_o^{-1}\mathbf{P}^{\|}\|^2 \, dr \text{ (Debye component)}, \tag{4}$$

$$U^{\#} = (1/2)\varepsilon_o^{-1} \int \|\mathbf{P}^{\#}\|^2 \, d\mathbf{r} = (1/2)\varepsilon_o{}^{-1} \int \|\mathbf{P} - \mathbf{P}.\mathbf{e}\|^2 \, d\mathbf{r}. \tag{5}$$

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

$$\vartheta_n = \langle U^{\#}n/Un \rangle, \tag{6}$$

where the symbol "$\langle\,\rangle$" denotes time average, $U_n = (1/2)\varepsilon_o \int_{\Omega_n} \|\nabla\phi\|^2 d\mathbf{r}$, and $U^{\#}{}_n = (1/2)\varepsilon_o{}^{-1} \times \int_{\Omega_n} \|\mathbf{P}\text{- }\mathbf{P}.\mathbf{e}\|^2 d\mathbf{r}$, where Ω_n is 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 Ω_n at the α-carbon is justified *a posteriori*, as shown in Sec. III. 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.[4,6] 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.

III. RESULTS

A. Anomalous polarization at the protein-water interface: A study 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, the water polarization $\mathbf{P} = \mathbf{P}(\mathbf{r})$ for soluble natural proteins with structures reported in the Protein Data Bank (PDB) is computed along molecular dynamics trajectories (the field magnitude is estimated from the work associated with displacing a test charge). Each 5 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[10–17] are provided in the caption for Fig. 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 the quotient $U^{\#}{}_n$ / U_n. To this end, we recorded charge distribution $\rho(\mathbf{r}, t)$, internal field $\mathbf{E}(\mathbf{r}, t)$, and polarization $\mathbf{P}(\mathbf{r}, t)$ for each intermediate structure/solvent configuration.

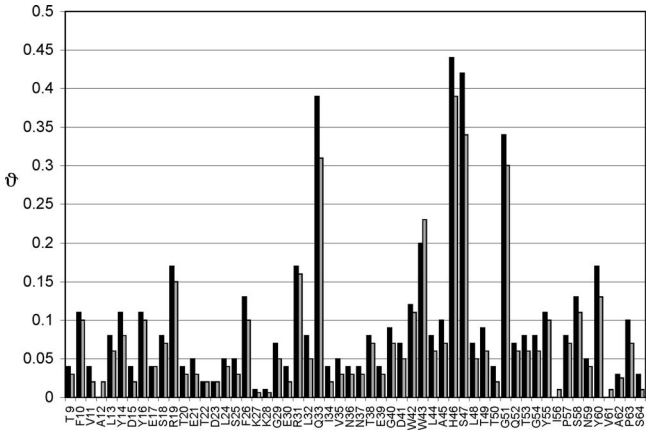

FIG. 1. Anomalous polarization fraction (ϑ) for each residue within the solvent-equilibrated folded protein chain for the soluble SH3 domain (PDB.1SRL). The PDB file numbering is followed in naming residues along the chain and the one-letter code for amino acids is adopted. Black bars represent protein surface interrogation with spheres of radius r = 6 Å centered at α-carbons, while grey 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 of a truncated octahedral cell of TIP3P water molecules that provide at least four water layers of solvent envelope.[10] Protein atoms are described with the parm99SB force field parameterization.[11] Water molecules extended at least 12 Å from the surface of the protein. Ewald sums[12] 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,[13] 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.[14] An AMBER package[15] was adopted for these MD simulations, with charges on the molecules assigned according to the BCC charge model using AM1 optimized geometries and potentials.[16,17] After protein/solvent equilibration (as defined in main text), the protein backbone coordinates are partially constrained according to the Shake scheme[13] and only side chains are allowed to explore conformation space, generating a gamut of local hydration patterns.

The structure/solvent system is considered equilibrated at time t_o if the RMSD of backbone atomic coordinates averaged over randomly chosen pairs of chain conformations within a time interval $[t_o, t_o + \tau]$ ($\tau \approx 1$ ns) is less than 1 Å. For all 9 proteins in this study, this criterion was fulfilled for t_o = 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. 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 structure. If we exclude the residues A12, I56, and V61 that are fully buried within the structure, it is clear that the positively charge lysines (K27, K28) have the lowest APFs. As described below, this is expected since the ammonium cation ($-NH_3^+$) 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 solvent-exposed backbone hydrogen bonds, known as *dehydrons* (marked in green in Fig. 2). Due to the nanoscale water confinement created by the packing defect that gives rise to the dehydron, a significant nonvanishing component $\mathbf{P}^{\#}$ is expected for residues paired by such hydrogen bonds.[6] These packing defects expose the backbone polar groups amide (>N–H) and carbonyl (>C=O) to structure-disruptive effects of backbone hydration with the net effect of steering water dipoles into orientations that are not collinear with **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 **E**. This is so since water molecules with reduced water coordination tend to preserve their hydrogen-bond pattern thereby becoming impervious to the torque $\mathbf{E(r)} \times \{\int[(\mathbf{r'}-\mathbf{r})\nabla.\mathbf{P(r')}]\, d\mathbf{r'}\}$ is imposed by **E** on the polarization-related dipole with moment $\{\int[(\mathbf{r'}-\mathbf{r})\nabla.\mathbf{P(r')}]\, d\mathbf{r'}\}$ (position **r** is generic). Thus, water polarization in this context is expected to contain and indeed contains (Fig. 1) a significant anomalous non-Debye contribution.

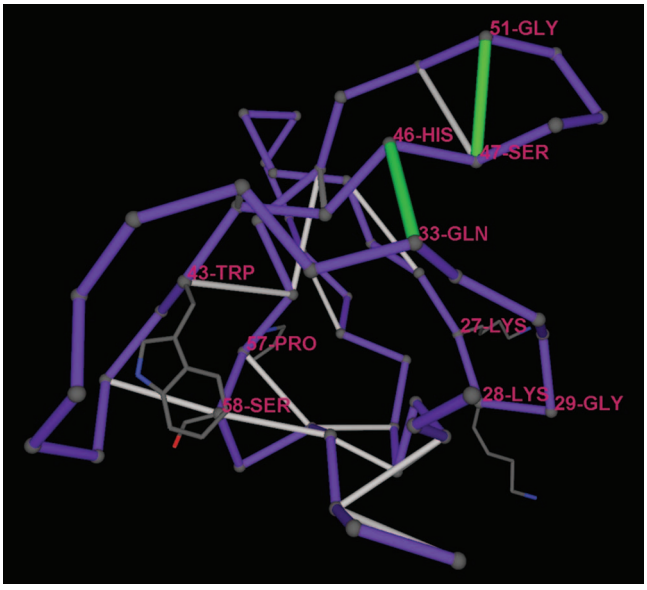

FIG. 2. Dehydron pattern for the solvent-equilibrated structure of the soluble SH3 domain (PDB.1SRL). The backbone is represented as virtual bonds joining the α-carbons of residues along the chain, with well-shielded backbone hydrogen bonds and dehydrons shown as segments sustained between the paired residues in grey and green, respectively. The dehydrons are determined from the PDB structural coordinates following the protocol indicated in Ref. 18. In accord with this protocol, the under-wrapping 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. Propensities of individual residues to promote anomalous polarization

In Fig. 3(a) we show the individual propensities of the 20 residue types to align interfacial water along the electrostatic field **E** by computing the APF of residues in 9 PDB-reported proteins (specified in caption for Fig. 4). The expected APF for each residue type is obtained by averaging the APFs for that residue type in all 9 proteins. The computation amounts to average over the structural contexts in the 9 proteins where the particular residue type occurs. Due to the dominance of dehydrons as structural determinants of APF (Figs. 1 and 4), superseding individual propensities (Fig. 1), we have excluded dehydron-paired residues from the calculations in Fig. 3.

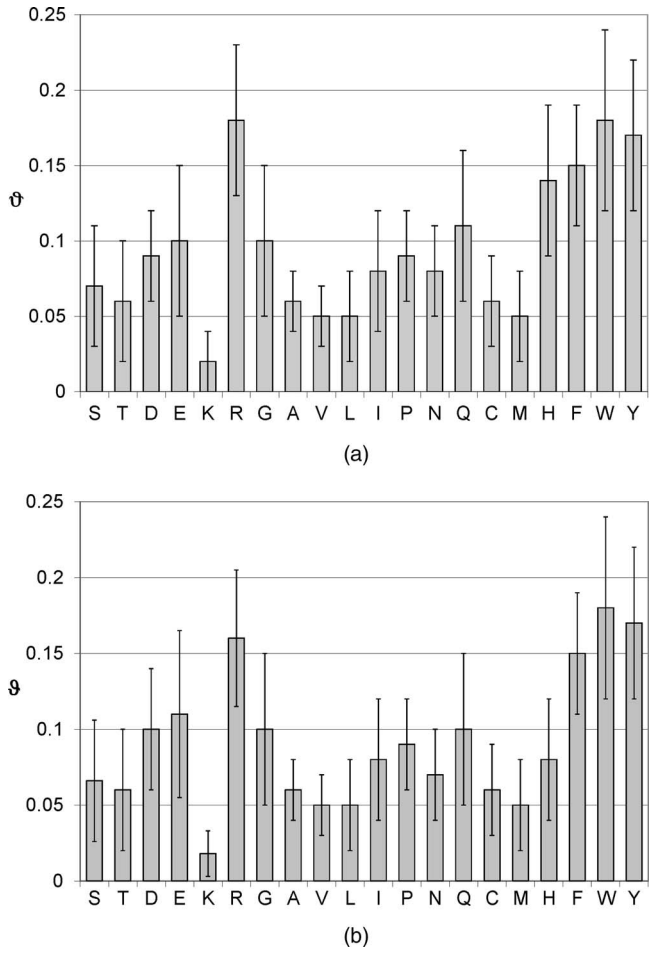

FIG. 3. APFs for each residue type averaged over all structural environments where the residue type occurs in 9 PDB-reported proteins described in the caption for Fig. 4. The radius r = 6 Å has been adopted and the error bars represent the dispersion in ϑ-values. (a) APFs at pH 7. (b) APFs at pH 5.5.

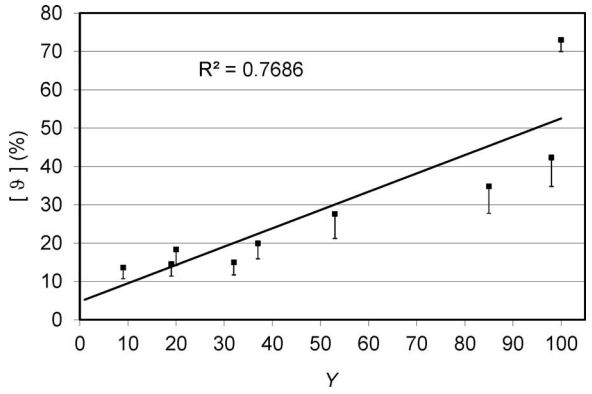

FIG. 4. Correlation between ϑ-value averaged over all residues in the protein ($[\vartheta]$), and the dehydron-to-backbone-hydrogen-bond ratio Y for the protein. Both parameters are given as percentages. The proteins studied identified by their respective PDB files and Y-ratios (in brackets) are: 1SRL (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.

As a class, the aromatic residues (H, F, W, Y) have the highest APF values due to their water-organizing power and more importantly, 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 without disrupting the tetrahedral water lattice) as it is the case with nonpolar aliphatic side chains (L, V, I, A). Thus, the resilient non-tetrahedral 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 in the same class of positively charged residues with aliphatic (methylene) linkages, is also striking, yet expected. The ammonium cation ($-NH_3^+$) in lysine has the highest charge concentration of all amino acids; therefore, it strongly organizes hydration along electrostatic field lines, while the guanidinium cation ($[-NH=C(NH_2)_2]^+$) 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 pKa 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. 3(a)) and pH 5.5 (Fig. 3(b)). The pH window 5.5–7 apparent when contrasting Figs. 3(a) and 3(b) contains only the pKa $\approx$ 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 pKas dictate extreme pH values that would introduce denaturing conditions for the proteins studied.

C. Anomalous polarization and structural defects in proteins

A significant correlation is established for PDB-reported proteins (Fig. 4) between the average APF ($[\vartheta]$) over all 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,[18] 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 $\sim$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 supra-nanoscale 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. Similar mutational studies have already been conducted for other antifreeze proteins like that of the winter flounder revealing comparable effects.[19]

D. Anomalous polarization and ephemeral hydration patterns

As highlighted in Figs. 1, 2, and 4, the dominant structural motif promoting anomalous polarization is the dehydron, a packing defect in proteins consisting of a solvent-exposed backbone hydrogen bond. A separate study revealed that the hydration patterns for residues paired by dehydrons are extremely ephemeral,[20] and thus dehydrons have been shown experimentally to be promoters of protein associations.[21] The lifetime of a hydration pattern is measured in picoseconds (ps) and is given by the average residence time τ_n of a water molecule in the neighborhood Ω_n of residue n defined above.[20] These considerations suggest an anticorrelation between the local parameters τ_n and ϑ_n, whereby a large anomalous polarization corresponds to a short residence time. A complete analysis of the 9 PDB-reported proteins indicated in the caption for Fig. 4 reveals that such correlation is upheld. The tight τ-ϑ linear anticorrelation ($R^2 = 0.86$) is displayed in Fig. 5 and can be rationalized by noting that anomalous polarization arises from nanoscale solvent confinement which deprives the water molecule of hydrogen-bond coordination

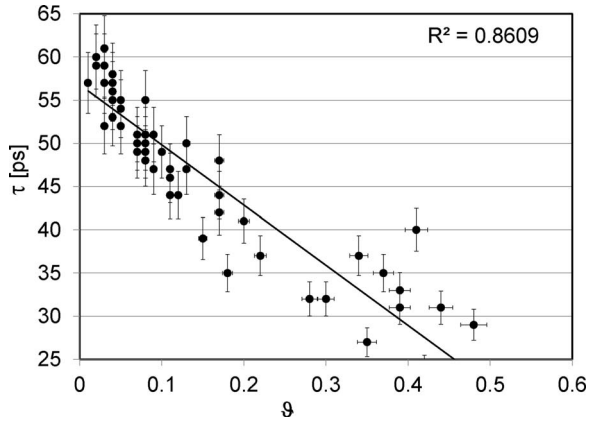

FIG. 5. Anticorrelation between water residence time τ and APF ϑ for all residues in the 9 PDB-reported proteins indicated in the caption for Fig. 4. The datapoints marked by circles correspond to the residues in PDB.1A8O. The error bars along both coordinates indicate Gaussian dispersions of datapoints corresponding to the remaining 8 proteins.

possibilities when compared with bulk levels. The latter coordination deprivation enhances the kinetic energy, thereby decreasing the residence time.

IV. CONCLUDING REMARKS

In this work, we demonstrated that the dielectric structure of interfacial water entails a breakdown of the Debye *ansatz* that postulates the alignment of water polarization with the protein electrostatic field. By solvent-equilibrating soluble proteins with reported structure, we identified the dehydron as the main type of structural defect that causes the anomalous polarization effect.

ACKNOWLEDGMENTS

The author is indebted to Professor Ridgway Scott for insightful discussions and to Felipe Belinky and Dr. Jianping Chen for their help with the solvent dynamics calculations. This work was partially supported by CONICET, the National Research Council of Argentina and by the Collegium Basilea, the Institute for Advanced Study of Basel, Switzerland.

[1]P. Debye, *Polar Molecules* (Dover Publications, New York, 1929).
[2]N. Giovambattista, C. F. Lopez, P. Rossky, and P. G. Debenedetti, Proc. Natl. Acad. Sci. U.S.A. **105**, 2274 (2008).
[3]Y.-K. Cheng and P. Rossky, Nature (London) **392**, 696 (1998).
[4]A. Fernández, Phys. Rev. Lett. **108**, 188102 (2012).
[5]F. Despa, Ann. N.Y. Acad. Sci. **1066**, 1 (2006).
[6]A. Fernández, J. Chem. Phys. **137**, 231101 (2012).
[7]G. Sigalov, P. Scheffel, and A. Onufried, J. Chem. Phys. **122**, 094511 (2005).
[8]S. Tanizaki and F. Feig, J. Chem. Phys. **122**, 124706 (2005).
[9]A. V. Morozov, T. Kortemme, and D. Baker, J. Phys. Chem. B **107**, 2075 (2003).
[10]W. L. Jorgensen, J. Chandrasekhar, J. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[11]V. Hornak, R. Abel, A. Okur, B. Strockbine, A. E. Roitberg, and C. L. Simmerling, Proteins: Struct., Funct., Bioinf. **65**, 712 (2006).
[12]T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[13]J. P. Ryckaert, G. Ciccotti, and H. J. Berendsen, J. Comput. Phys. **23**, 327 (1977).
[14]H. J. Berendsen, J. P. Postma, W. F. van Gunsteren, A. Di Nola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).
[15]J. Wang, R. M. Wolf, J. W. Caldwell, and P. A. Kollamn, J. Comput. Chem. **25**, 1157 (2004).
[16]A. Jakalian, B. L. Bush, D. B. Jack, and C. I. Bayly, J. Comput. Chem. **21**, 132 (2000).
[17]A. Jakalian, D. B. Jack, and C. I. Bayly, J. Comput. Chem. **23**, 1623 (2002).
[18]A. Fernández and R. S. Berry, J. Proteome Res. **9**, 2643 (2010).
[19]K. Sharp, Proc. Natl. Acad. Sci. U.S.A. **108**, 7281 (2011).
[20]A. Fernández, J. Chen, and A. Crespo, J. Chem. Phys. **126**, 245103 (2007).
[21]A. Fernández and R. L. Scott, Phys. Rev. Lett. **91**, 018102 (2003).