

## Variational mechanics of water at biological interfaces

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2011 J. Phys. A: Math. Theor. 44 292001

(<http://iopscience.iop.org/1751-8121/44/29/292001>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 201.235.214.217

The article was downloaded on 24/06/2011 at 20:13

Please note that [terms and conditions apply](#).

## FAST TRACK COMMUNICATION

# Variational mechanics of water at biological interfaces

**Ariel Fernández**

Instituto Argentino de Matemática 'Alberto Calderón', CONICET, Saavedra 15,  
1083 Buenos Aires, Argentina  
and  
Department of Computer Science, The University of Chicago, Chicago, IL 60637, USA

E-mail: [ariel@uchicago.edu](mailto:ariel@uchicago.edu)

Received 2 March 2011, in final form 4 June 2011

Published 20 June 2011

Online at [stacks.iop.org/JPhysA/44/292001](http://stacks.iop.org/JPhysA/44/292001)**Abstract**

A variational principle governing nanoscale water configurations at biological interfaces is derived. The underlying functional determines a scalar field  $g = g(\mathbf{r})$  indicating the expected hydrogen-bond coordination of water at position  $\mathbf{r}$ , and includes an elastic term to penalize local coordination losses. The minimization of interfacial free energy implies that  $g$  must obey a Poisson-type equation with Dirichlet boundary conditions. The equation is integrated and the specialization of the treatment to protein–water interfaces enables the determination of protein complexation propensities and the identification of protein binding sites and binding hot spots.

PACS numbers: 87.15.Aa, 61.20.Qg, 41.20.Cv

**1. Introduction**

In spite of significant progress [1–4], the physical underpinnings of biological–water interfaces remain elusive, hindering our understanding of biomolecular associations [5–9]. The partial confinement of interfacial water within inhomogeneous nanoscale cavities makes it difficult to define thermodynamic concepts such as interfacial tension, useful to describe homogeneous phase separation [10, 11]. In aqueous biological contexts, the interfacial physics is complex, as solutes associate forming complexes but do not precipitate. The only known phase separations for natural soluble proteins under physiological conditions are neurodegenerative fibrillogenic aggregates, involving misfolded prions [9].

This work addresses the problem of computing interfacial tension in biomolecular contexts and validates the results by identifying protein binding sites without resorting to information on homologous proteins or knowledge-based potentials [12]. Strikingly, the treatment singles out the soluble prion protein structure as possessing the highest protein–water (P–W) interfacial tension. Befitting a nanoscale description, the theoretical derivation introduces an elastic term that penalizes local changes in hydrogen-bonding coordination and allows for compensations

to this elastic penalization as water dipoles interact with pre-existing or induced electrostatic fields yielding a more or less restricted polarization depending on their local coordination environments.

## 2. Variational mechanics of water at the biological interface

To compute interfacial tension, we adopt the field  $g = g(\mathbf{r})$ , a descriptor that assigns to each position vector  $\mathbf{r}$  the expected value of hydrogen-bond coordination of a water molecule situated within a sphere centered at position  $\mathbf{r}$  with radius  $2.5 \text{ \AA}$  (the thickness of a single water layer [13, 14]). This radius spans a van der Waals radius plus half a hydrogen-bond distance. The  $g(\mathbf{r})$  value is computed as time average over solvent configurations determined by molecular dynamics over a 100 ns period after the protein structure is equilibrated with the solvent, thus allowing for breathing motions of exposed atoms [13].

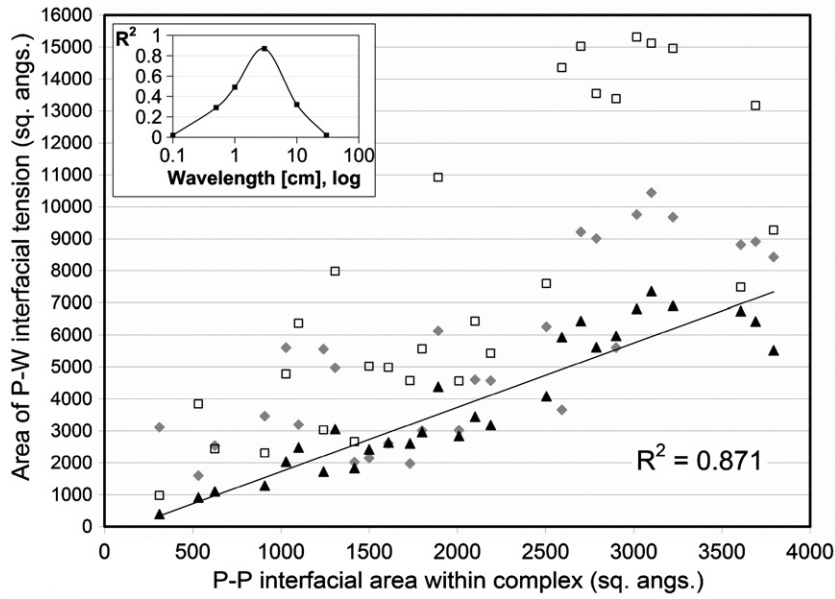
Within this representation, the minimization of interfacial tension requires a variational principle, whereby the extremal of the surface integral  $\Delta G_{\text{if}} = \int \Xi(g(\mathbf{r}), \nabla g(\mathbf{r})) \, d\mathbf{r}$  represents the free-energy cost of spanning the P–W interface. Thus,  $\Xi(g(\mathbf{r}), \nabla g(\mathbf{r})) \, d\mathbf{r}$  gives the free-energy cost of transferring water from bulk to volume  $d\mathbf{r}$  at position  $\mathbf{r}$ . A hydrogen bond is defined by the geometric constraints: O–O distance  $< 3.2 \text{ \AA}$  and O–H–O angle  $a_{\text{HB}}$  satisfying  $120^\circ \leq a_{\text{HB}} \leq 180^\circ$ . Compared with bulk water, interfacial water has reduced hydrogen-bonding opportunities ( $g < g_{\text{bulk}} = 4$ ) and may counterbalance such losses by interacting with polar groups on the protein surface. Thus, the term  $\Delta G_{\text{if}}$  incorporates unfavorable local decreases in  $g$  and favorable polarization contributions.

Exploiting the analogy with Lagrangian mechanics [10], we write  $\Delta G_{\text{if}} = \int \frac{1}{2} \{ \lambda |\nabla g|^2 - |\mathbf{P}(g(\mathbf{r}))|^2 \} \, d\mathbf{r}$ , where the elastic term  $\frac{1}{2} \lambda |\nabla g|^2$  accounts for tension-generating reductions in water coordination and the polarization  $\mathbf{P}(g(\mathbf{r}))$  accounts for dipole–electrostatic field interactions. Bulk water properties are incorporated since the scaling parameter  $\lambda$  is obtained from the interfacial tension of a large nonpolar sphere with radius  $\theta$  in the macroscopic limit  $\theta/1 \text{ nm} \rightarrow \infty$ . We obtain  $\lambda = 9.0 \text{ mJ m}^{-1} = \lim_{\theta/1 \text{ nm} \rightarrow \infty} \gamma(4\pi\theta^2) / \int \frac{1}{2} |\nabla g|^2 \, d\mathbf{r}$ , where  $\gamma = 72 \text{ mJ m}^{-2}$  is the macroscopic surface tension of water at 298 K, and  $\int |\nabla g|^2 \, d\mathbf{r} = \frac{1}{2} 4^2 \text{ m}^{-1} (4\pi\theta^2)$ , since  $\nabla g \neq 0$  only close to the interface and the associated  $g$ -jump is 4 units in magnitude. Thus, unfavorable P–W hydrophobic interactions are accounted for through the positive elastic contribution  $\int \frac{1}{2} \lambda |\nabla g|^2 \, d\mathbf{r}$  which extrapolates to macroscopic dimensions.

To determine the  $g$ -dependence of polarization  $\mathbf{P} = \mathbf{P}(\mathbf{r})$ , we adopt the Fourier-conjugate wavenumber space ( $\omega$ -space) and represent the dipole correlation kernel  $K_p(\omega)$  and the electrostatic field  $\mathbf{E} = \mathbf{E}(\mathbf{r})$  in this space. This representation is required to capture the entire dielectric loss spectrum occurring mostly within the microwave range ( $10^{-3} \text{ m} \leq \omega^{-1} \leq 0.3 \text{ m}$ ) [15, 16]. The biophysical experimental data analyzed required examining results at bulk wavelengths spanning several orders of magnitude and are best reproduced at bulk wavelength  $\omega^{-1} \approx 0.03 \text{ m}$  (figure 1, inset). At fixed wavelength,  $\mathbf{P}$  and  $\mathbf{E}$  are proportional [17, 18], but the proportionality constant is generally  $\omega$  dependent [18], which introduces a significant departure from the Poisson–Boltzmann electrostatics (see below). In  $\omega$ -space, we obtain

$$F(\mathbf{P})(\omega) = K_p(\omega) F(\mathbf{E})(\omega), \quad (1)$$

where  $F$  denotes the 3D-Fourier transform  $F(\mathbf{f})(\mathbf{w}) = (2\pi)^{-3/2} \int e^{i\mathbf{w}\cdot\mathbf{r}} \mathbf{f}(\mathbf{r}) \, d\mathbf{r}$ , and the kernel  $K_p(\omega)$  is the Lorentzian  $K_p(\omega) = (\varepsilon_b - \varepsilon_o) / [1 + (\tau(\mathbf{r})c)^2 |\omega|^2]$ , with  $\tau(\mathbf{r})c$  the local dielectric relaxation length,  $c$  the speed of light,  $\varepsilon_b$  the bulk permittivity constant and  $\varepsilon_o$  the vacuum permittivity. For bulk water, we obtain  $\tau = \tau_b \approx 100 \text{ ps}$ , and the relaxation length



**Figure 1.** Protein–water interfacial tension promotes biomolecular associations. The total area of surfaces generating protein–water (P–W) interfacial tension in free subunits (uncomplexed state) plotted against the total protein–protein (P–P) interfacial area for 28 protein complexes (table 1). Datapoints with ordinates given by the areas  $A_1$ ,  $A_2$  and solvent-exposed surface area (SESA) are represented with diamonds, triangles and squares, respectively. The linear fit with correlation coefficient  $R^2 = 0.871$  for  $A_2$ -(P–P interface) points was obtained by linear regression. Inset: correlation coefficient  $R^2$  as a function of the bulk wavelength parameter  $\tau_{bc}$ . The correlations were examined at bulk wavelengths over various orders of magnitude spanning the microwave range (wavelengths  $10^{-3}$ – $0.3$  m). The data are best correlated at  $\tau = \tau_b = 100$  ps, with the wavelength  $\approx 0.03$  m, the adopted parameter value.

$\tau_{bc} = \omega_b^{-1} \approx 0.03$  m, the microwave wavelength yielding the best fit with experimental data (figure 1, inset containing sensitivity analysis). Through the Lorentzian, the frequency dependence of bulk permittivity is subsumed into the normal distribution factor  $[1 + (\tau(\mathbf{r})c)^2|\omega|^2]^{-1}$  with  $\tau(\mathbf{r}) \equiv \tau_b$ . Since  $\mathbf{P}(\mathbf{r})$  satisfies  $\nabla \cdot (\epsilon_0\mathbf{E} + \mathbf{P})(\mathbf{r}) = \rho(\mathbf{r})$  ( $\rho(\mathbf{r})$  is the charge density distribution) [17], equation (1) yields the following equation in  $\mathbf{r}$ -space:

$$\nabla \cdot \left[ \int F^{-1}(K)(\mathbf{r} - \mathbf{r}')\mathbf{E}(\mathbf{r}') d\mathbf{r}' \right] = \rho(\mathbf{r}), \quad (2)$$

with  $K(\omega) = \epsilon_0 + K_p(\omega)$ . The convolution  $\int F^{-1}(K)(\mathbf{r} - \mathbf{r}')\mathbf{E}(\mathbf{r}') d\mathbf{r}'$  captures the correlation of the dipoles with the electrostatic field. Equation (2) is not the Poisson–Boltzmann equation, which requires ad hoc proportionality between  $\mathbf{E}$  and  $\mathbf{P}$ .

Upon water confinement, the dielectric relaxation undergoes a frequency redshift arising from the reduction in hydrogen-bond partnerships that translates in a reduction in dipole orientation possibilities ( $g(\mathbf{r}) < 4$ ,  $\tau(\mathbf{r}) > \tau_b$ ) [18]. Thus, at position  $\mathbf{r}$ , the relaxation time is  $\tau = \tau_b \exp(B(g(\mathbf{r}))/k_B T)$ , where the kinetic barrier  $B(g(\mathbf{r})) = -k_B T \ln(g(\mathbf{r})/4)$  yields  $\tau(\mathbf{r}) = \tau_b(g(\mathbf{r})/4)^{-1}$ .

To obtain the  $g$ -dependent polarization term  $\mathbf{P}(\mathbf{r})$ , we introduce the electrostatic potential  $\Psi = -\Phi(\mathbf{r})$ , with  $\mathbf{E} = \nabla\Phi$ . Thus, equation (2) reads in  $\omega$ -space

$$F(\Phi)(\omega) = F(\rho)(\omega)/[|\omega|^2 K(\omega)], \quad (3)$$

where  $|\omega|^2$  is the Fourier conjugate of  $\nabla^2$ . Thus,  $\mathbf{E}$  is obtained from the inverse Fourier transform of equation (3):

$$\begin{aligned}\mathbf{E}(\mathbf{r}) &= \nabla F^{-1}\{F(\rho)(\omega)/[|\omega|^2 K(\omega)]\}(\mathbf{r}) \\ &= (2\pi)^{-3/2} \nabla \int d\omega e^{-i\omega \cdot \mathbf{r}} F(\rho)(\omega)/[|\omega|^2 K(\omega)].\end{aligned}\quad (4)$$

The negative potential  $\Phi(\mathbf{r})$  is determined for a generic charge distribution:

$$\rho(\mathbf{r}) = \sum_{m \in L} 4\pi q_m \delta(\mathbf{r} - \mathbf{r}_m), \quad (5)$$

with  $L$  the set of charges on the protein surface labeled by index  $m$ , which in  $\omega$ -space reads

$$F(\rho)(\omega) = (2\pi)^{-3/2} \sum_{m \in L} 4\pi q_m \exp(i\omega \cdot \mathbf{r}_m). \quad (6)$$

Thus, we obtain

$$\Phi(\mathbf{r}) = (2\pi)^{-3/2} \sum_{m \in L} \int d\omega e^{-i\omega \cdot (\mathbf{r} - \mathbf{r}_m)} 4\pi q_m / [|\omega|^2 K(\omega)]. \quad (7)$$

Now the  $g$ -dependent polarization term  $\mathbf{P}(\mathbf{r})$  is obtained from equation (7) and the definition  $\mathbf{E} = \nabla \Phi$ :

$$\begin{aligned}\mathbf{P}(\mathbf{r}) &= \int F^{-1}(K_p)(\mathbf{r} - \mathbf{r}') \mathbf{E}(\mathbf{r}') d\mathbf{r}' \\ &= (2\pi)^{-3} \sum_{m \in L} \int d\mathbf{r}' F^{-1}(K_p)(\mathbf{r} - \mathbf{r}') \nabla_{\mathbf{r}'} \int d\omega e^{-i\omega \cdot (\mathbf{r}' - \mathbf{r}_m)} 4\pi q_m / [|\omega|^2 K(\omega)].\end{aligned}\quad (8)$$

The field  $(g, \nabla g)$  is extremal of  $\Delta G_{\text{if}} = \int \Xi(g(\mathbf{r}), \nabla g(\mathbf{r})) d\mathbf{r} = \int \frac{1}{2} \{\lambda |\nabla g|^2 - |\mathbf{P}(g(\mathbf{r}))|^2\} d\mathbf{r}$  if the spatial version of the Euler–Lagrange equation [10] holds:

$$\nabla[\partial \Xi / \partial \nabla g] = \partial \Xi / \partial g, \quad (9)$$

with the Dirichlet boundary condition  $g(\mathbf{r}) \equiv 4$  for  $\|\mathbf{r}\| = R$ . This condition is defined with  $R$  sufficiently large so the protein molecule is contained in the ball  $\{\|\mathbf{r}\| < R', R' \ll R - 5d\}$ , with  $d = 2.5 \text{ \AA} \approx$  thickness of a water layer. Thus,  $g(\mathbf{r})$  satisfies the Poisson-type equation

$$\nabla^2 g = -h(\mathbf{r}), \quad h(\mathbf{r}) = (1/\lambda) \partial \mathbf{P} / \partial g, \quad \text{with } g(\mathbf{r}) \equiv 4 \quad \text{for } \|\mathbf{r}\| = R, \quad (10)$$

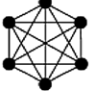

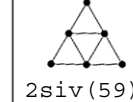
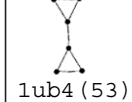

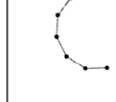
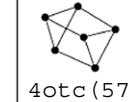




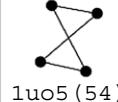

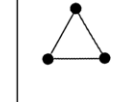

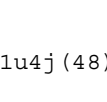
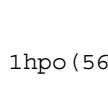

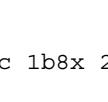
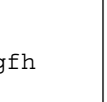

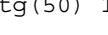

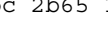
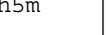

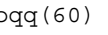
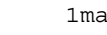
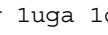


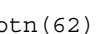


which is numerically integrated [19, 20] to yield  $g(\mathbf{r})$ .

### 3. Protein–water interfacial tension determining association propensities

The solvent-accessible envelope of the protein surface [21, 22] may be covered by a minimal set  $W$  of water-confining osculating spheres  $D_j$ ,  $j \in W$ . These spheres make first-order contact with the envelope obtained by sliding a water molecule along the surface, thus smoothening out its singularities [21]. As a minimal covering,  $W$  contains all interfacial water molecules and this property no longer holds if any osculating sphere is excluded from the set, even though the osculating spheres may intersect. Interfacial tension arises in  $D_j$  when  $\Delta G_j > 0$ , where  $\Delta G_j = \frac{1}{2} \int_{D_j} \{\lambda |\nabla g|^2 - |\mathbf{P}(g(\mathbf{r}))|^2\} d\mathbf{r}$  is the interfacial surface tension associated with the spanning contact region  $j$ .

To support the claim that  $\Delta G_{\text{if}}$  is a determinant of the propensity of the protein to associate, we examined free subunits in 28 protein complexes (table 1) with defined contact topologies [5]. The total area of the protein–protein (P–P) interfaces for each complex is computed after identification of the subunit residues engaged in intermolecular contacts. Separately,  $\Delta G_{\text{if}}$

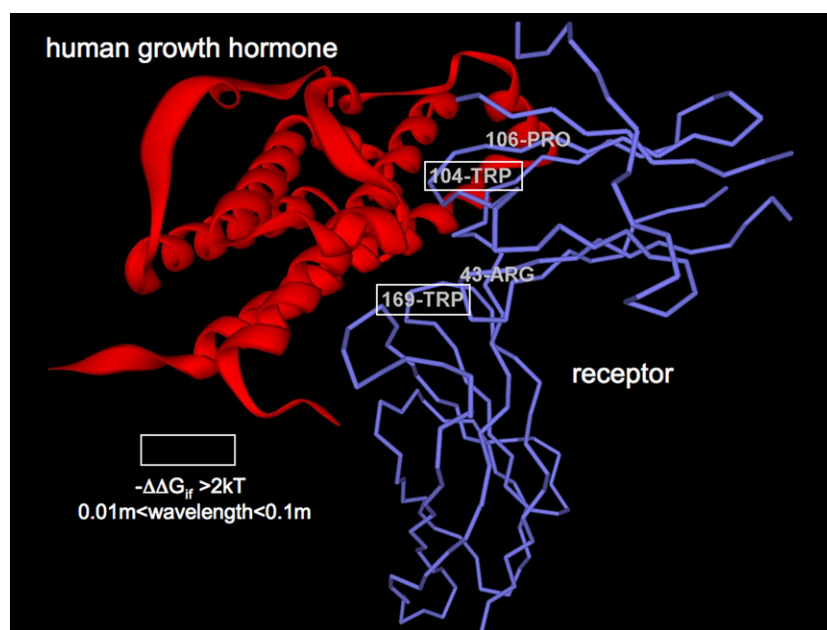
**Table 1.** Representative topologies of inter-subunit contacts for 28 protein complexes reported in the Protein Data Bank (PDB). The dataset is used to unravel the role of protein–water (P–W) interfacial tension patches as promoters of protein associations. In the topological representation of a complex, a ball represents a protein subunit and a stick between two balls represents a protein–protein (P–P) interface. The numbers in brackets denote % decrease in overall interfacial tension upon complexation starting with free subunits.

 2af7 (63)	 1fr3 (61)	 2siv (59)	 1ub4 (53)	 2a1b (51)	 2c5r (40)	 4otc (57)
 1qoh (54)	 1dvz (57) 1jtk (55)	 1hho (64)	 1nys (39)	 1uo5 (54) 1vwe (53)	 1fvu (50)	 1zij (52) 1ufy (51)
 1wud (42)	 1u4j (48)  1hpo (56)		 3orc (48)  1b8x (48)  2gfh (48)			
	 1htg (50)  1lve (47)		 1goc (50)  2b65 (50)  1h5m (50)			
	 2oqq (60)  1v2z (51)		 1may (50)  1uga (50)  1quz (50)			
	 1btn (62)  2ien (59)		 1ihb (57)  1p6z (72)			

for each free protein subunit is computed by the numerical integration of equation (10) with charge and atomic radii (equation (5)) assigned using the program PDB2PQR [23]. We define filtered sets  $W_n = \{j \in W: \Delta G_j \geq nk_B T\}$  ( $n = 1, 2, \dots$ ) of contributors to the P–W interfacial tension in the free subunits ( $W_n \subset W_{n'}$  for  $n > n'$ ), with  $S_n$  the total P–W interface associated with  $W_n$  and  $A_n$  the surface area of  $S_n$ . In all 28 complexes examined, the total P–P interface distributed among the subunits within the complex is 100% contained in  $S_2$ , and it is only  $60 \pm 7\%$  contained in  $S_3$ .

These results suggest that interfacial solvent cavities spanned at a thermodynamic cost  $\Delta G \geq 2k_B T$  are the promoters of protein associations. This conclusion is corroborated by the tight correlation ( $R^2 = 0.87$ ) between  $A_2$  and the total P–P interfacial area of each complex (figure 1). This correlation becomes weaker for  $S_1$  with area  $A_1$  ( $R^2 = 0.62$ ) and becomes negligible ( $R^2 = 0.19$ ) when we consider the total solvent-exposed surface area (SESA) [21, 22] of free protein subunits. The correlation between the areas of surface patches that significantly destabilize the P–W interface ( $A_2$ ) and those of binding sites reveals that proteins associate to reduce P–W interfacial tension.

The interfacial tension  $\gamma_{if} = \Delta G_{if}/(\text{SESA})$  of soluble nonhomologous monomeric proteins with structures reported in the Protein Data Bank (PDB) has been computed for an exhaustive dataset of 11964 uncomplexed monomeric proteins or peptides and varies in the range

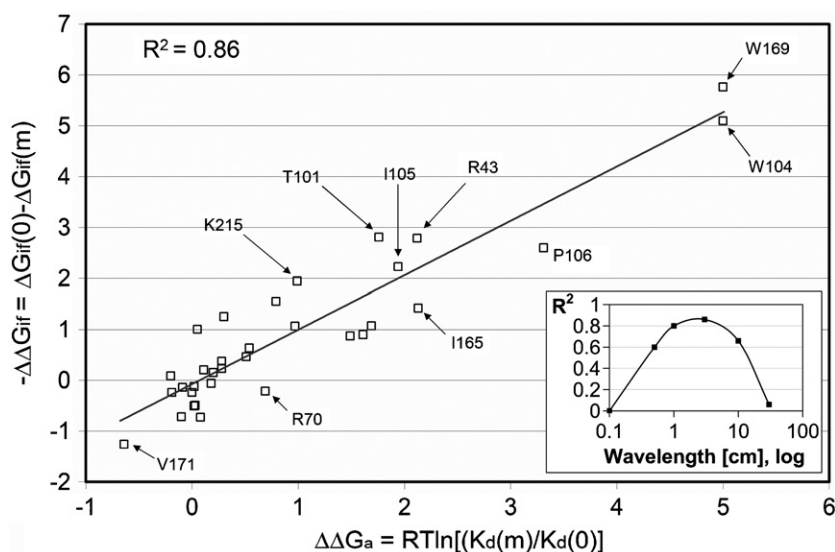


**Figure 2.** The protein–protein (P–P) interface between the human growth hormone (hGH, ribbon) and its receptor (backbone line) in the 1:1 complex [6].

(This figure is in colour only in the electronic version)

$1.25 \text{ mJ m}^{-2} \leq \gamma_{\text{if}} \leq 9.90 \text{ mJ m}^{-2}$ . The list of proteins or peptides with a sustainable structure is compiled in the supporting information of [24]. The computation follows exactly the same operational premises adopted for free protein subunits from the complexes indicated in table 1: the P–W interface is decomposed into a minimal covering of osculating spheres making first-order contact with the solvent-accessible envelope, and the free energy cost of spanning each interfacial sphere,  $\Delta G_j$ , is computed as described above with the field  $g$  determined for each protein structure by the numerical integration of equation (10). Strikingly, the maximum  $\gamma_{\text{if}}$  value corresponds to the cellular metastable fold of a prion protein (PDB.1qm0), capable of promoting phase separation through fibrillogenic aggregation [25].

Hot spots in P–P associations were identified by scanning the interface through the site-specific substitution of amino acids for alanine (effectively a side-chain truncation) [6]. An extensively scanned interface corresponds to the human growth hormone (hGH)-receptor complex (figure 2). To validate our molecular marker, the hot spots obtained by alanine scanning were contrasted with an *in silico* side-chain truncation procedure (side-chain groups beyond  $\beta$ -carbon are removed) that determines a reduction in interfacial tension,  $\Delta\Delta G_{\text{if}} = \Delta G_{\text{if}}(\text{m}) - \Delta G_{\text{if}}(0)$ , caused by a site-directed mutation ( $0 = \text{wild type} \rightarrow \text{m} = \text{mutation}$ ). The side-chain truncation at the  $\beta$ -carbon is equivalent to the alanine substitution in the experimental dissection of the interfacial hot spots. A strong correlation exists between the terms  $-\Delta\Delta G_{\text{if}}$  and the association free-energy difference  $\Delta\Delta G_{\text{a}}$  experimentally obtained from the substitution of interfacial residues (figure 3). Thus, inferred tension patches coincide with hot spots at complex interfaces identified by mutational scanning.



**Figure 3.** Comparison with experimental alanine scanning of the protein–protein (P–P) interface of the human growth hormone (hGH) receptor validates the interfacial tension computation. Correlation between the association free-energy differences  $\Delta\Delta G_a$  between mutant (m) and wild type (0) for the alanine substitution of each residue from the hGH receptor at the P–P interface and the interfacial free-energy difference between wild type and mutant for the free protein subunit (uncomplexed hGH receptor). The association free energy is computed as  $\Delta G_a = -RT \ln K_a = RT \ln(K_d)$ , where  $K_a$  and  $K_d$  are respectively the association and dissociation equilibrium constants. Inset: correlation coefficient  $R^2$  as a function of the bulk wavelength parameter  $\tau_b$ . The dependence was examined over several orders of magnitude spanning the microwave range (wavelengths  $10^{-3}$ – $0.3$  m). Optimal correlation is achieved at  $\tau = \tau_b = 100$  ps, wavelength  $\approx 0.03$  m, the adopted parameter value.

#### 4. Conclusions

The aqueous biological interface is physically heterogeneous and geometrically confined at multiple scales. For these reasons, no generic thermodynamic concept such as interfacial tension, broadly used in the context of homogeneous phase separation, can be directly determined without a surface integration procedure. This work provides the theoretical framework to enable such integration. The biological interfacial tension is unique in that biological solutes may form obligatory, ephemeral or adventitious complexes but never precipitate under physiological conditions. The only distinct phase separation encountered in biology appears to be the amyloidogenic aggregation of prions [9, 25], which as noted in this work are endowed with a soluble structure generating the highest protein–water (P–W) interfacial tension.

Since P–W interfaces are highly heterogeneous both physically and geometrically, broad classifications of proteins into hydrophilic, amphiphilic and hydrophobic classes are not very helpful to understand biomolecular interactivity and the fact that complexation (and not precipitation) is always the evolutionary solution to the need for functional regulation and cooperativity. In broad terms we could say that, for identical local curvature, a hydrophilic patch is more likely to reduce interfacial tension than a hydrophobic patch since the latter curtails the hydrogen bonding opportunities of interfacial water ( $g < 4$ ). However, interfacial tension is also defined by local curvature: a hydrophilic



patch introducing favorable polarization may be partially hindered from hydration due to local geometric constraints and thus, the local interfacial  $g$ -value may be lower than the bulk value  $g = 4$ , contributing to the interfacial tension.

The P–W interfacial tension may be modulated by the cellular machinery to alter complexation propensities, and hence the degree of functional cooperativity and regulation, in a purposely choreographed manner. Thus, glycosylation and phosphorylation are two ubiquitous post-translational modifications of proteins that increase hydrophilicity by the transesterification of solvent-accessible specific side chains (serine, threonine and tyrosine). These processes generally decrease the P–W interfacial tension by increasing the number of favorable interactions with interfacial water.

The results presented in figures 1–3 uphold the variational treatment of biological water and highlight the relevance of P–W interfacial tension as a determinant of protein associations. The polarization and elastic terms are defined respectively by a sub-microscopic ( $\tau_b$ ) and a macroscopic ( $\lambda$ ) parameter. While the latter is fixed, the former is adjustable, since dielectric loss spreads over the microwave range; yet the adopted value  $\tau_b = 100$  ps (wavelength  $\approx 0.03$  m) yields the optimum prediction of binding sites (figures 1 and 3).

The variational principle derived incorporates an elastic term to penalize local losses in water hydrogen-bond coordination and was shown to govern nanoscale configurations of water at biological interfaces. In consonance with a nanoscale description, the treatment allows for compensations to the elastic penalization as water dipoles interact with the electrostatic field in a more or less restricted manner depending on their individual coordination environments. The extremal of the functional that underlies the variational principle determines the expected hydrogen-bond coordination  $g(\mathbf{r})$  of water at a given position  $\mathbf{r}$  in space. This scalar field  $g$  is easily determined numerically once it is shown that it obeys a Poisson-type equation with Dirichlet boundary conditions.

These findings may impact the field of biomaterial self-assembling, since molecular fabrication is determined by the structural compatibility of subunits that confers noncovalent interactions [26]. These organizing interactions are promoted by the propensity to reduce P–W interfacial tension. While the charge distribution will be altered upon protein association and the interfacial tension will change accordingly, such interactions can be in principle engineered by introducing patches of high P–W interfacial tension in the basic subunits.

## Acknowledgments

This work was partially supported by funds from the Institute of Biophysical Dynamics and the Computer Science Department at the University of Chicago and by the National Research Council of Argentina (CONICET). The author is grateful to Professor Ridgway Scott for valuable discussions.

## References

- [1] Ashbaugh H S and Pratt L R 2006 *Rev. Mod. Phys.* **78** 159
- [2] Despa F 2005 *Ann. New York Acad. Sci.* **1066** 1
- [3] Gerstein M and Chothia C 1996 *Proc. Natl Acad. Sci. USA* **93** 10167
- [4] Giovambattista N, Lopez C F, Rosky P and Debenedetti P G 2008 *Proc. Natl Acad. Sci. USA* **105** 2274
- [5] Levy E D, Pereira-Leal J B, Chothia C and Teichmann S A 2006 *PLoS Comput. Biol.* **2** e155
- [6] Clackson T, Ultsch M H, Wells J A and de Vos A M 1998 *J. Mol. Biol.* **277** 1111
- [7] Frauenfelder H *et al* 2009 *Proc. Natl Acad. Sci. USA* **106** 5129
- [8] Fenimore P W, Frauenfelder H, McMahon B H and Young R D 2004 *Proc. Natl Acad. Sci. USA* **101** 14408
- [9] Fernández A 2010 *Transformative Concepts for Drug Design: Target Wrapping* (Heidelberg: Springer)

- [10] Rowlinson J S and Widom B 1982 *Molecular Theory of Capillarity* (New York: Oxford University Press)
- [11] Stillinger F H 1973 *J. Solut. Chem.* **2** 141
- [12] Capra J A, Laskowski R A, Thornton J M, Singh M and Funkhouser T A 2009 *PLoS Comput. Biol.* **5** e1000585
- [13] Schulz E, Frechero M, Appignanesi G and Fernández A 2010 *PLoS One* **5** e12844
- [14] Kumar P, Buldyrev S V and Stanley H E 2009 *Proc. Natl Acad. Sci. USA* **106** 22130
- [15] Hasted J B 1972 *Water: A Comprehensive Treatise* vol 1 ed F Franks (New York: Plenum) pp 255–309
- [16] Buchner R, Barthel J and Stauber J 1999 *Chem. Phys. Lett.* **306** 57
- [17] Debye P 1929 *Polar Molecules* (New York: Dover)
- [18] Scott R, Boland M, Rogale K and Fernández A 2004 *J. Phys. A: Math. Gen.* **37** 9791
- [19] Johnson C 1987 *Numerical Solution of Partial Differential Equations by the Finite Element Method* (Cambridge: Cambridge University Press)
- [20] Persson P O and Strang G 2004 *SIAM Rev.* **46** 329
- [21] Street A G and Mayo S L 1998 *Fold. Des.* **3** 253
- [22] Zhang N, Zeng C and Wingreen N S 2004 *Proteins Struct. Funct. Bioinform.* **57** 565
- [23] Dolinsky T J, Nielsen J E, McCammon J A and Baker N A 2004 *Nucleic Acids Res.* **32** W665
- [24] Fernández A and Berry R S 2010 *J. Proteome Res.* **9** 2643
- [25] Zahn R *et al* 2000 *Proc. Natl Acad. Sci. USA* **97** 145
- [26] Zhang S 2003 *Nat. Biotechnol.* **21** 1171