Communication: Nanoscale electrostatic theory of epistructural fields at the protein-water interface

Ariel Fernández

Citation: J. Chem. Phys. **137**, 231101 (2012); doi: 10.1063/1.4772603
View online: http://dx.doi.org/10.1063/1.4772603
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v137/i23
Published by the American Institute of Physics.

Communication: Nanoscale electrostatic theory of epistructural fields at the protein-water interface

Ariel Fernández[a)]
Instituto Argentino de Matemática "Alberto P. Calderón," CONICET (National Research Council), Saavedra 15, Buenos Aires 1083, Argentina and Collegium Basilea, Institute for Advanced Study, Hochstrasse 51, CH 4053 Basel, Switzerland

(Received 19 October 2012; accepted 4 December 2012; published online 19 December 2012)

Nanoscale solvent confinement at the protein-water interface promotes dipole orientations that are not aligned with the internal electrostatic field of a protein, yielding what we term epistructural polarization. To quantify this effect, an equation is derived from first principles relating epistructural polarization with the magnitude of local distortions in water coordination causative of interfacial tension. The equation defines a nanoscale electrostatic model of water and enables an estimation of protein denaturation free energies and the inference of hot spots for protein associations. The theoretical results are validated vis-à-vis calorimetric data, revealing the destabilizing effect of epistructural polarization and its molecular origin. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4772603]

With intricate combinations of geometric and chemical features, protein-water interfaces enshrine the physical promoters of supra-molecular organization.[1–3] However, the understanding of such associations remains elusive partly because of the lack of an electrostatic model of interfacial water valid at nanoscale dimensions.[4,5] Such a model is needed to capture the effects of solvent confinement at protein-water interfaces, and to provide an electrostatics-based theory of interfacial tension.[2,3] This is the aim of this work.

We first derive an equation that captures the interplay between dielectric polarization and interfacial tension as water is partially confined within inhomogeneous nanoscale cavities on the surface of a folded soluble protein. We start with a general treatment, leaving aside the Debye assumption regarding the alignment of the dielectric polarization $\mathbf{P}$ with the internal electrostatic field $\mathbf{E}$.[6] The field $\mathbf{E}$ is defined by $\nabla\cdot(\varepsilon_o\mathbf{E}) = \rho + \gamma$, where ε_o is the vacuum permittivity, $\rho = \rho(\mathbf{r})$ is the charge distribution on the protein molecule, and $\gamma = \gamma(\mathbf{r}) = -\nabla\cdot\mathbf{P}$ is the net charge distribution for the statistical bath of solvent dipoles. The Poisson equation satisfied by the total field $\mathbf{F} = \mathbf{E} + {\varepsilon_o}^{-1}\mathbf{P}$ is

$$\nabla\cdot(\varepsilon_o\mathbf{F}) = \nabla\cdot(\varepsilon_o\mathbf{E}+\mathbf{P}) = \nabla\cdot(\varepsilon_o\mathbf{E}+(\mathbf{P}\cdot\mathbf{e})\mathbf{e}+\mathbf{P}^{\#}) = \rho, \tag{1}$$

where $\mathbf{P}$ is decomposed into a field-aligned component $(\mathbf{P}\cdot\mathbf{e})\mathbf{e}$ ($\mathbf{e} = \mathbf{E}/\|\mathbf{E}\|$) and an "epistructural" component $\mathbf{P}^{\#}$, orthogonal to $\mathbf{E}$. Since $\nabla\times\mathbf{F} = \mathbf{0}$, we define the electrostatic potential ϕ through: $\mathbf{F} = -\nabla\phi$. Thus, the electrostatic energy is

$$\begin{aligned} \mathrm{U} &= -(1/2)\varepsilon_o\int\phi\Delta\phi\,\mathrm{d}\mathbf{r} = (1/2)\varepsilon_o\int(\nabla\phi)\cdot(\nabla\phi)\mathrm{d}\mathbf{r} \\ &= (1/2)\varepsilon_o\int\mathbf{F}\cdot\mathbf{F}\,\mathrm{d}\mathbf{r} = \mathrm{U_D}+\mathrm{U}^{\#}, \end{aligned} \tag{2}$$

where the "Debye term" $\mathrm{U_D}$ incorporates the aligned polarization:

$$\mathrm{U_D} = (1/2)\varepsilon_o\int(\mathrm{E}+\varepsilon_o^{-1}\mathbf{P}\cdot\mathbf{e})^2\mathrm{d}\mathbf{r}, \quad \mathrm{E} = \|\mathbf{E}\|. \tag{3}$$

The orthogonal polarization determines the term $\mathrm{U}^{\#}$ given by

$$\mathrm{U}^{\#} = (1/2)\varepsilon_o^{-1}\int\left\|\mathbf{P}^{\#}\right\|^2\mathrm{d}\mathbf{r} = (1/2)\varepsilon_o^{-1}\int\|\mathbf{P}-\mathbf{P}\cdot\mathbf{e}\|^2\mathrm{d}\mathbf{r}. \tag{4}$$

From Eq. (1), we obtain the following equation for the orthogonal polarization divergence:

$$\nabla\cdot(\mathbf{P}^{\#}) = \rho - \nabla\cdot[\varepsilon_o\mathbf{E}+(\mathbf{P}\cdot\mathbf{e})\mathbf{e}]. \tag{5}$$

Due to nanoscale water confinement, the nonvanishing component $\mathbf{P}^{\#}$ is expected. Packing defects in the protein structure expose the polar backbone groups amide (>N–H) and carbonyl (>C=O) to the structure-disruptive effects of backbone hydration,[3,7] steering water dipoles into orientations not collinear with $\mathbf{E}$ (Fig. 1). This backbone solvation materializes through hydrogen bonds of confined water molecules with backbone polar groups. The confined water molecules relinquish some of their hydrogen bonding possibilities to adjacent water molecules[3] or to polar groups of the protein. The reduction in coordination represents a departure from the bulk water structure embodied in a tetrahedral lattice and the resulting polarization is termed *epistructural*, since it occurs around the protein structure.

To describe the relation between epistructural polarization and local water structure, we introduce a local descriptor of water hydrogen-bonding coordination. We define the scalar field $\Psi = \Psi(\mathbf{r})$ as a time-averaged number of hydrogen-bonds of a water molecule located at position $\mathbf{r}$ to other water molecules or to polar groups of the protein. We expect that departures from bulk-like coordination ($\Psi = 4$) will introduce significant polarization with no alignment to the field $\mathbf{E}$. Thus, local exposures of the protein backbone, with

a)E-mail: ariel@uchicago.edu.

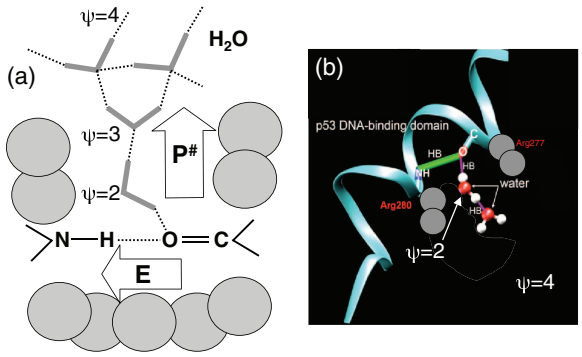

FIG. 1. Nanoscale cavities on the protein surface confine interfacial water, hampering dipole alignment with the internal field **E**, hence generating epistructural polarization ($\mathbf{P}^{\#}$). (a) Schematic depiction of the effect on interfacial tension and epistructural polarization. As nanoscale confinement introduces a reduction in the water coordination (Ψ) below bulk levels ($\Psi = 4$), the partial charges distributed in water molecular orbitals are not compensated intermolecularly (as in the regular tetrahedral lattice), a net dipole moment is created in the cavity, and the resulting polarization contribution ($\mathbf{P}^{\#}$) does not align with **E**. The panel describes a protein packing deficiency that exposes a backbone hydrogen bond to the solvent. (b) Illustration of the case represented in panel (a), provided by the hydrogen bond pairing residues Arginine 277 with Arginine 280 in the helical DNA-binding domain of the p53 protein.[7] Water penetrates the cavity and engages in hydrogen bonds with the backbone amide or carbonyl, but in so doing, the coordination falls below bulk levels ($\Psi < 4$). Due to the net lack of charge compensation, the local dipole moment resulting from solvent confinement is not aligned with **E**. The latter is generated by the backbone hydrogen-bond dipoles aligned along the helical axis.

protein-solvent hydrogen bonds not aligned with the field determined by intramolecular hydrogen bonds are likely to generate the **E**-orthogonal component $\mathbf{P}^{\#}$ (Fig. 1). In fact, explicit calculations of the vector field $\mathbf{P} = \mathbf{P}(\mathbf{r})$ ($\mathbf{r}$ = spatial position vector) for soluble proteins studied in this work reveal a significant contribution $U^{\#}$.

To make it mathematically tractable, the scalar field $\Psi = \Psi(\mathbf{r})$ is smoothened out by taking it to be the time average of hydrogen-bonding coordination of a water molecule within a sphere of radius 4 Å, centered at position $\mathbf{r}$. Then, to quantify the magnitude of epistructural polarization arising from a reduction in water coordination upon nanoscale confinement (Fig. 1), we introduce the ansatz $\mathbf{P}^{\#} \propto \nabla\Psi$, yielding

$$U^{\#} = (1/2)\varepsilon_o^{-1}\int \left\|\mathbf{P}^{\#}\right\|^2 d\mathbf{r} = (1/2)\boldsymbol{\lambda}\int \|\nabla\psi\|^2 d\mathbf{r} = U_{\psi}, \tag{6}$$

where λ is the proportionality factor. The elastic integrand $\frac{1}{2}\lambda\|\nabla\Psi\|^2$ implies that the epistructural polarization promotes interfacial tension, since $\frac{1}{2}\lambda\|\nabla\Psi\|^2$ accounts for tension-generating reductions in water coordination ($\|\nabla\Psi\| > 0$) and vanishes everywhere except at the protein-water interface. The parameter λ is obtained by an independent estimation of the interfacial free energy of a nonpolar sphere with radius θ and contrasting this result with the elastic integral in the macroscopic limit $\theta/1\text{nm} \to \infty$. We get $\lambda = 9.0$ mJ/m $= \lim_{\theta/1\text{nm}\to\infty} [\gamma(4\pi\theta^2)/\int \frac{1}{2}\|\nabla\Psi\|^2 d\mathbf{r}]$, where $\gamma = 72$ mJ/m^2 is the macroscopic surface tension of water at 298 K.[2]

From Eq. (6), it follows that $\mathbf{P}^{\#} = \xi\nabla\Psi$, where $\xi = (\lambda\varepsilon_o)^{1/2}$. Thus, Eq. (5) yields the following equation:

$$\xi\Delta\psi = \{\rho - \nabla\cdot[\varepsilon_o\mathbf{E} + (\mathbf{P}\cdot\mathbf{e})\mathbf{e}]\}, \tag{7}$$

where $\Delta = \nabla^2$ is the Laplace operator. The boundary condition in this case is $\Psi(\mathbf{r})\equiv 4$ (bulk value) for $\|\mathbf{r}\| = R$, with R sufficiently large so that the protein molecule is contained in the ball $\{\|\mathbf{r}\| < R', R' \ll R\text{-}5d\}$, with d = 2.7 Å ≈ thickness of a water layer.

By relating the coordination structure of water with epistructural polarization, Eq. (7) defines a nanoscale electrostatic model. The interfacial energy U_Ψ may be directly computed by integration of Eq. (7) and represents also a structure-destabilizing electrostatic energy, as implied by Eq. (6). *The structure destabilization arises from partial hindrance of the alignment of solvent dipoles with the internal field* **E**, *which is concomitant with reductions in water coordination, a promoter of interfacial tension.*

The interfacial energy U_Ψ is evaluated by computing the elastic integral in Eq. (6) as a time average over an interval beyond structure equilibration along a thermalization trajectory. The molecular dynamics trajectory is generated by immersion and thermalization in a pre-equilibrated solvent bath of a structure obtained from the protein data bank (PDB) (see the supplementary material).[16] Thus, the charge distribution $\rho(\mathbf{r}, t)$, internal field $\mathbf{E}(\mathbf{r}, t)$, and polarization $\mathbf{P}(\mathbf{r}, t)$ are recorded for structure snapshots along the equilibration trajectory to generate $U_\Psi = \langle(1/2)\lambda\int\|\nabla_\Psi\|^2\, d\mathbf{r}\rangle$ ($\langle\rangle$ = time average after equilibration) by numerical integration of Eq. (7). To assess the structure-destabilizing effects of epistructural polarization, we examined soluble monomeric proteins for which PDB-reported structure and thermodynamic data on thermal denaturation are available (see the supplementary material).[16] To establish a comparison with denaturation free-energy changes, we incorporate the entropic cost of solvent confinement at the interface

$$\Delta S_\psi = R\langle ln[\boldsymbol{\Pi}_j\Psi_j/4]\rangle, \tag{8}$$

where R is the universal gas constant and the dummy index j labels all water molecules along the equilibration

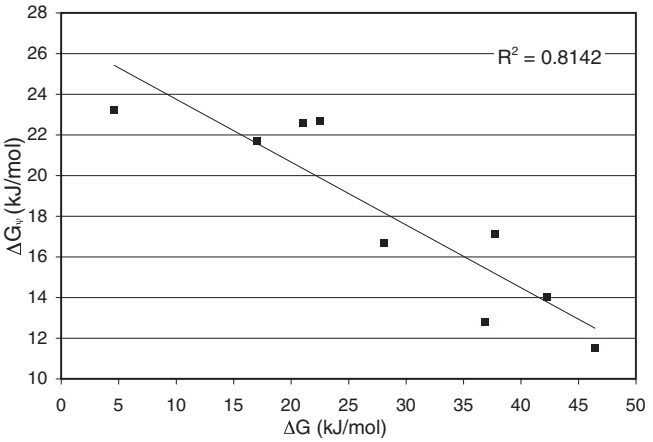

FIG. 2. Free-energy change (ΔG) for thermal denaturation *versus* interfacial free energy (ΔG_Ψ) for soluble monomeric proteins (see the supplementary material).[16]

trajectory. We found a significant negative correlation between the interfacial free energy $\Delta G_\Psi = U_\Psi - T\Delta S_\Psi$ of the PDB-reported structure and the free energy change (ΔG) associated with thermal denaturation (Fig. 2). Thus, denaturation is favored proportionally to the interfacial free energy, attesting to the structure-destabilizing role of epistructural polarization (cf. Eq. (6)). The epistructural polarization energy may be regarded as an indicator of the extent to which the "structure is at odds with the solvent," in the sense that polarization cannot fully align with the electrostatic field, while the solvent configuration responsible for this effect generates interfacial tension. The entropic term $-T\Delta S_\Psi$ reinforces the epistructural polarization effect since water molecules with $\Psi < 4$ contribute to decrease the interfacial entropy (cf. Eq. (8)).

We now establish the fact that binding hot spots are the sites that generate the most significant reductions in epistructural polarization upon protein-protein (P-P) association. Our approach focuses on interfacial solvent, whereas hot spots are usually characterized by examining the protein surface.[8] We analyze patterns of water exclusion upon complex formation, focusing on complexes for which individual residue contributions to the association free energy ΔG_a have been experimentally probed by alanine scanning of the P-P interface.[9–15] The site-directed substitution of a residue for alanine amounts to a truncation of the residue side chain at the β-carbon (expansion by one CH_3 group if the residue is glycine (G)). The increment in the association free energy $\Delta\Delta G_a$ is determined calorimetrically for the wild-type $\rightarrow$ mutant transformation, enabling identification of the residues that most contribute to the free energy of association.

A physical characterization of hot spots emerges by comparing the profiles resulting from integration of Eq. (7) for free protein subunits with the independent alanine scanning of the P-P interfaces within specific complexes. The hot spots are the residues that contribute upon association to reduce epistructural polarization or interfacial free energy by either displacing, or becoming deprived of "hot" ($\Psi < 4$) interfacial water.

Figure 3 validates this assertion, focusing on alanine scanning analysis for well-studied complexes, and contrasting the results with the coordination quality (Ψ-value) of the water molecules excluded upon P-P association. To identify

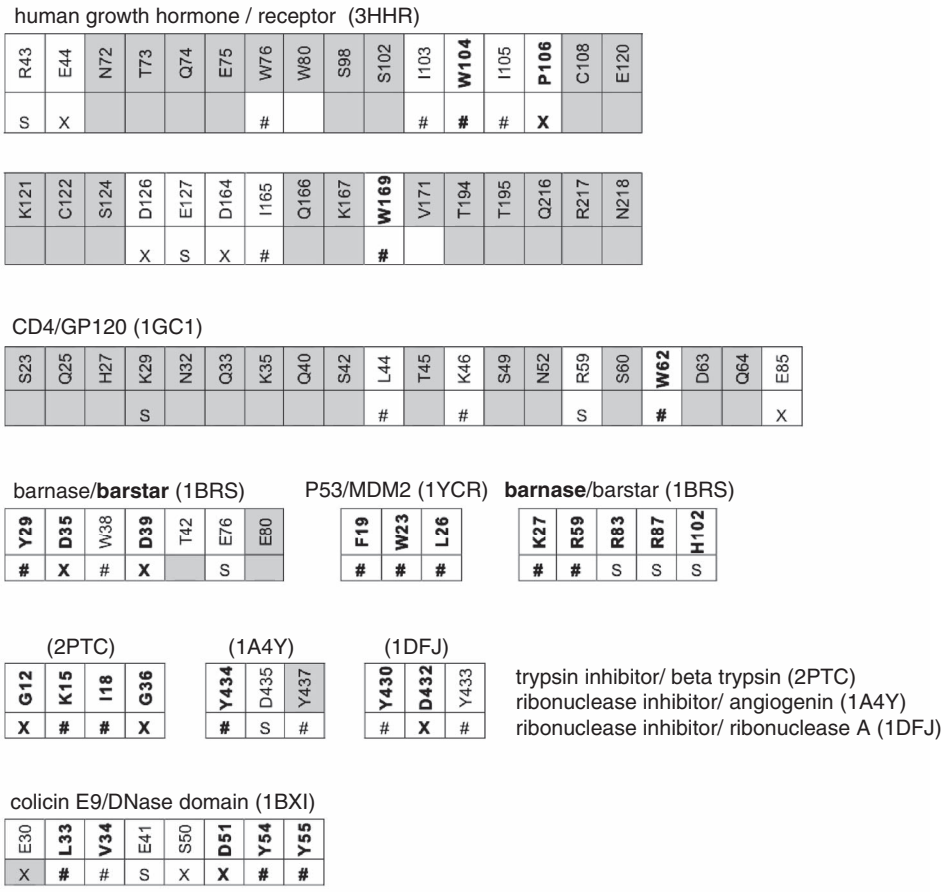

FIG. 3. Comparison between alanine scanning results and Ψ-values of residues in free subunits that become part of protein-protein (P-P) interfaces in complexes. The comparison involves the $\Delta\Delta G_a$ classifier (upper entries) and the Ψ-classifier (lower entries) for P-P interfacial residues in different protein complexes reported in the protein data bank (PDB). The $\Delta\Delta G_a$ classifier is constructed from calorimetric data,[10–16] while the Ψ-classifier is obtained by numerical integration of Eq. (7) for free protein subunits. Hot spot residues are marked in bold characters, regular font with white background, and regular font with grey background, corresponding, respectively, to the ranges $\Delta\Delta G_a \geq 3$ kcal/mol, 1 kcal/mol $\leq \Delta\Delta G_a < 3$ kcal/mol, and $\Delta\Delta G_a < 1$ kcal/mol (upper rows for each PP interface). According to the Ψ-classifier (lower rows), residues are marked in bold characters, regular font with white background and regular font with grey background if their vicinal water lies in the range $\Psi < 3$, $3 \leq \Psi < 3.5$, $3.5 \leq \Psi < 4$, respectively. The effect of the P-P association on the interfacial water in the hydration vicinity of a residue is represented by "X" if water is displaced, "#" if the residue displaces water intermolecularly across the P-P interface, "S" if an intermolecular salt bridge is formed across the PP-interface, and "blank space" if the hydration vicinity remains unaltered upon P-P association.

the location of hot ($\Psi < 4$) water molecules, we define the *hydration vicinity* of a residue by two 4 Å- radius spheres centered at the backbone amide nitrogen and carbonyl oxygen of the residue. A water molecule is said to be vicinal when the oxygen atom of the water molecule lies within the hydration vicinity of the residue.

To validate our electrostatics-based computation, the results generated by the integration of Eq. (7) on the free subunits of complexes reported in the protein data bank (PDB) were contrasted with the alanine-scanning results on the P-P interfaces of the same complexes. Thus, we label residues according to ranges of $\Delta\Delta G_a$ for wild-type $\rightarrow$ mutant transformation and also according to the coordination quality of the interfacial water they displace or are deprived of upon protein-protein association. Hot-spot residues are classified according to the ranges $\Delta\Delta G_a \geq 3$ kcal/mol, 1 kcal/mol $\leq \Delta\Delta G_a < 3$ kcal/mol and $\Delta\Delta G_a < 1$ kcal/mol (upper rows for each P-P interface, Fig. 3). According to a second classifier (lower rows, Fig. 3), residues are grouped according to the coordination ranges for vicinal water: $\Psi < 3$, $3 \leq \Psi < 3.5$, $3.5 \leq \Psi < 4$. The P-P interfaces for complexes with available alanine scanning data were examined, classifying residues according to the $\Delta\Delta G_a$ and independently according to the Ψ-parameter obtained by integration of Eq. (7) on the free protein subunits. The following complexes were examined (PDB entries are given in brackets): human growth hormone/hGH receptor (3HHR),[9] HIV-1-CD4/GP120 (1GC1),[10] barnase and barstar in barnase/barstar complex (1BRS),[11] P53/MDM2 (1YCR),[12] trypsin inhibitor/beta-trypsin (2PTC),[13] ribonuclease inhibitor/angiogenin (1A4Y),[14] ribonuclease inhibitor/ribonuclease A (1DFJ),[14] and colicin E9 immuno-protein/colicin E9 DNase domain (1BXI).[15]

A statistically significant correlation (P-value $< 10^{-7}$) exists between the $\Delta\Delta G_a$ classifier and the Ψ-classifier of interfacial residues (Fig. 3). We may assert that protein association is driven by a displacement of hot interfacial water that promotes a reduction in interfacial energy with the concomitant reduction of epistructural polarization.

This work introduced and validated a theory for water dielectrics under nanoscale confinement that captures the electrostatic effects resulting from a breakdown of the Debye ansatz while extending the concept of interfacial tension to nanoscale dimensions by associating distortions in water structure with epistructural polarization fields surrounding soluble proteins.

The author acknowledges the valuable help of Dr. Jianping Chen with solvent structure computations.

[1]N. Giovambattista, C. F. Lopez, P. Rossky, and P. G. Debenedetti, Proc. Natl. Acad. Sci. U.S.A. **105**, 2274 (2008).
[2]A. Fernández, Proc. R. Soc. London, Ser. A **467**, 559 (2011).
[3]A. Fernández, Phys. Rev. Lett. **108**, 188102 (2012).
[4]F. Despa, Ann. N. Y. Acad. Sci. **1066**, 1 (2006).
[5]Y.-K. Cheng and P. Rossky, Nature (London) **392**, 696 (1998).
[6]P. Debye, *Polar Molecules* (Dover, New York, 1929).
[7]A. Fernández, *Transformative Concepts for Drug Design: Target Wrapping* (Springer-Verlag, Heidelberg, 2010).
[8]W. DeLano, Curr. Opin. Struct. Biol. **12**, 14 (2002).
[9]T. Clackson and J. A. Wells, Science **267**, 383 (1995).
[10]A. Ashkenazi, L. Presta, S. Marsters, J. Camarato, K. Rosenthal, B. Fendly, and D. Capon, Proc. Natl. Acad. Sci. U.S.A. **87**, 7150 (1990).
[11]G. Schreiber and A. R. Fersht, J. Mol. Biol. **248**, 478 (1995).
[12]V. Bottger, A. Bottger, C. Garcia-Echeverria, Y. Ramos, A. van der Eb, A. G. Jochemsen, and D. P. Lane, J. Mol. Biol. **269**, 744 (1997).
[13]M. J. Castro and S. Anderson, Biochemistry **35**, 11435 (1996).
[14]C. Chen and R. Shapiro, Proc. Natl. Acad. Sci. U.S.A. **94**, 1761 (1997).
[15]U. C. Kuhlmann, A. Pommer, G. R. Moore, R. James, and C. Kleanthous, J. Mol. Biol. **301**, 1163 (2000).
[16]See supplementary material at http://dx.doi.org/10.1063/1.4772603 for details on molecular dynamics simulations and calorimetric parameters for Fig. 2.