The Distribution of Electron Donor−Acceptor Potential on Protein Surfaces

Claudio Calonder,*,† Julian Talbot,‡ and Jeremy J. Ramsden†

Department of Biophysical Chemistry, Biozentrum, 4056 Basel, Switzerland, and Department of Chemistry and Biochemistry, Duquesne University, Pittsburgh, Pennsylvania 15282

Received: July 6, 2000; In Final Form: October 27, 2000

Maps of the surface distribution of the electron donor−acceptor (Lewis acid−base, AB) potential of proteins have been created by (i) identifying the surface residues, (ii) classifying them according to their AB characteristics, and (iii) projecting them onto a cylinder which is then cut and flattened. Net electron donor or acceptor potential densities are determined by an area-preserving gridding procedure, and the variation of these densities with grid size was investigated. Electron donors and acceptors are distributed at random on the surface of urease and cytochrome b_5, but not on human serum albumin. The results suggest that the characteristic length of the AB interaction which determines protein adsorption is greater than that implied by macroscopic surface tension measurements and, hence, that interfacial energies derived therefrom overestimate the hydration repulsion between a protein and a surface.

Three forces are involved in the interaction between two colloidal particles or between a particle and a surface:[1,2] (a) Lifshitz−van der Waals (LW), (b) electrostatic, and (c) Lewis acid−base (AB), also known as electron donor−acceptor, hydrophobic, hydrophilic, structural, etc. Experimental evidence corroborates theoretical predictions that in biological milieux, which are typically aqueous and close to neutral pH, with a moderately high monovalent salt concentration, the AB force dominates and may comprise 80−90% of the total interfacial energy (see, e.g., refs 3 and 4). In predicting the magnitudes of these forces, researchers have made frequent use of macroscopic single-substance surface tension parameters determined experimentally from surface tension measurements.[5] The use of this procedure would be unexceptionable were the AB potential distribution over the protein surface known to be uniform, but given that the protein is constituted from 20 different amino acids whose side groups (residues) vary widely, uniformity can by no means be presumed.

Previous attempts to quantify protein surface heterogeneity have fallen into two groups: (a) the work of Richards and others in quantifying the irregular surface topography[6] and (b) the mapping of the surface electrostatic potential. The distribution of charges on protein surfaces is often highly nonuniform, implying that under conditions where electrostatic interactions dominate, a ligand may be steered to approach a receptor in a particular orientation[7,8] and that a protein will be adsorbed to highly charged chromatographic materials in a preferred orientation.[9]

The purpose of this paper is first to develop a method for quantitatively determining the surface distribution of the AB potential and, second, to explore the implications of this distribution for the interfacial interaction energy of a protein with a uniform planar surface.

* Corresponding author: Dr. C. Calonder, Department of Chemical Engineering & Materials Science, Wayne State University, 5050 Anthony Wayne Drive, Detroit, MI 48202. Phone: 313-577-9357. FAX: 313-577-3810. E-mail: calonder@che.eng.wayne.edu.

† Biozentrum.

‡ Duquesne University.

TABLE 1: Amino Acids and Backbone Atoms Characterized According to Their Hydrogen Bonding Capacities[a]

amino acid residue	γ^+	γ^-	charge
Lys (K)	3	0	+
Arg (R)	5	0	+
Asp (D), Glu (E)	0	4	−
Trp (W)	1	0	
His (H), Tyr (Y)	1	1	
Asn (N), Gln(Q)	2	2	
Ser (S), Thr (T)	1	2	
backbone atom	γ^+	γ^-	**charge**
N	1	0	
O	0	2	

[a] Residues not listed are presumed to be incapable of forming hydrogen bonds, i.e. apolar.

Methods

Mapping Procedure. Each amino acid was classified by its side chain (residue) according to the maximum number of hydrogen bonds expected from simple chemical considerations[10] (Table 1); all charges are assumed to be sufficiently well-screened by the high ionic strength not to influence the hydrogen bonding propensities (it will be obvious that our treatment can be readily extended to relax this condition). All electron-donating (γ^-) and electron-accepting groups (γ^+) present in each amino acid residue were counted separately. The main-chain peptide groups invariably comprise electron acceptors (imine) and donors (oxygen), which were also included. The coordinates of the protein atoms were obtained from the Protein Data Bank (PDB) (Brookhaven National Lab., Brookhaven, NY). For simplicity, each polar amino acid residue was represented by the atoms responsible for the hydrogen bonding classification of the side chain. Apolar residues were represented by all their side-chain atoms.

As amino acids inside the protein are irrelevant for intermolecular interactions not involving conformational changes, only the external atoms of the molecule were taken into consideration. To separate them from the rest, we used the software package

10.1021/jp0024120 CCC: $20.00
Published on Web 12/15/2000

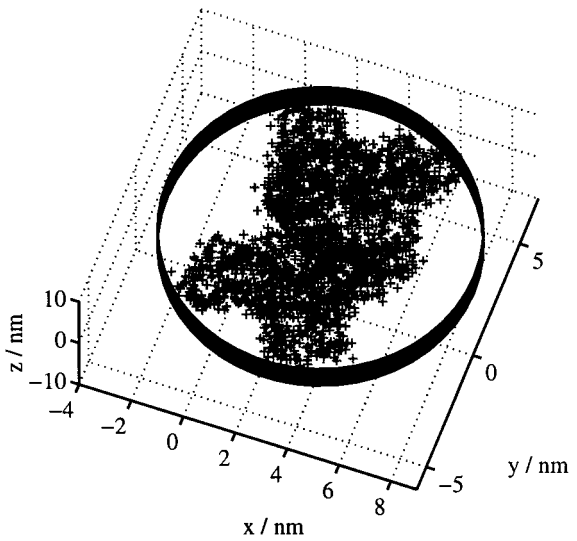

Figure 1. Positions of all atoms which determine the Lewis acid−base potential of the human serum albumin (HSA) surface (+), selected as described in the text, projected onto the paper, and viewed along an axis tilted 15° away from the axis of the marked cylinder which minimally encloses the protein. The cylinder is used to conformally project the atoms selected (+) for generating the AB potential map (Figure 2).

MSMS[11] to triangulate the solvent-excluded surface of a protein by rolling a probe sphere of radius 0.15 nm over it. The triangulation was carried out with a density of 1.0 vertices/Å^2, and atoms within a distance of 0.2 nm from the calculated surface were retained. If several atoms of a residue fulfilled this condition, just the one closest to the solvent-excluded surface was kept for the following mapping procedure.

After the location and classification of all atoms contributing to the AB distribution on the protein surface, a cylinder of the smallest radius able to enclose those atoms was positioned parallel to one of the three orthogonal axes of the Cartesian coordinate system used by the PDB (Figure 1). The positions of the atoms were projected onto the surface of the cylinder along lines extending radially from the chosen axis of revolution of the protein;i.e., a conformal cylindrical projection was used. The cylinder was then cut in the direction of its axis and flattened (Figure 2). The programs for carrying out these operations were written in Matlab version 5.3 (MathWorks, Natick, MA).

To properly compare the AB potentials of different protein areas with each other, the length distortions resulting from the cylindrical projection must be taken into account. Hence, a grid whose cells correspond to identically sized protein surface areas was superimposed on the AB map. To obtain such a grid, we projected the edges of the triangles produced by the MSMS package exactly as for the atoms. As these triangles are very small and of fixed size, equal numbers of triangle edges correspond to equal areas on the protein surface. The grid cell boundaries were therefore adjusted such that all cells contained equal numbers of edges. Each γ^+ or γ^- unit present in a grid rectangle was assigned a score of +1 or −1, respectively, and the apolar residues were assigned a score of zero. These scores were summed to obtain a renormalized AB potential for each cell, which was electron donating (γ^-, i.e., net score < 0), electron accepting (γ^+, i.e., net score > 0) or AB-neutral (net score = 0).[12] A useful parameter is the ratio R_a of the number of net γ^+ cells to the number of net γ^-, i.e.

$$R_a = \frac{N^+}{N^-} \tag{1}$$

which, as will be seen, depends on grid cell area a.

The above mapping procedure was applied to bovine liver cytochrome b_5 (soluble 88-residue fragment,[13] thereby neglecting

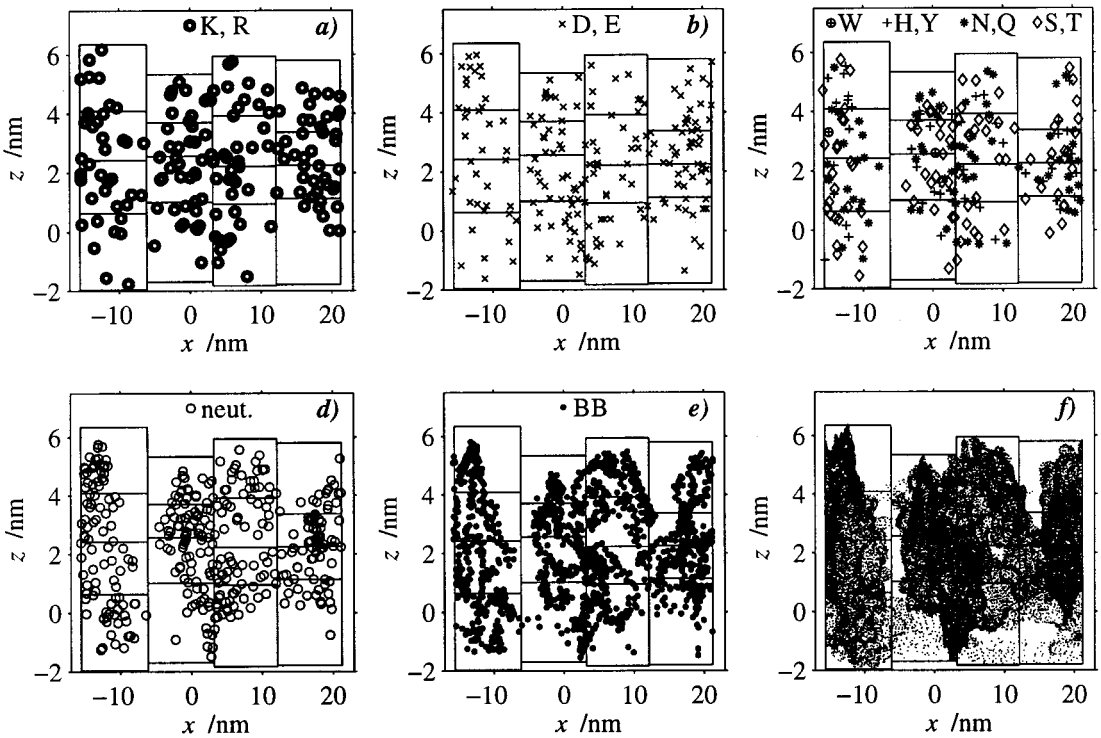

Figure 2. Mapped surface of HSA obtained as described in the text. For the sake of clarity the complete content of one map is displayed in six separate panels (a−f). All selected atoms are represented according to the classification given in Table 1. The abbreviation "neut" (panel d) denotes amino acids whose side chains are incapable of forming hydrogen bonds, and "BB" (e) denotes backbone atoms. Faint spots (f) represent the edges of the triangles with which the solvent-excluded surface is covered. A 4 × 4 grid has been superimposed on the map.

TABLE 2: Parameters Used to Simulate the Distribution of γ^+ and γ^- on the Protein Surface[a]

protein	A/nm^2	N_B	N_W	w_W
HEL	41.7	101	124	0.58
Cyt b_5	43.8	76	104	0.63
urease	191	460	593	0.61
HSA	262	838	1157	0.61

[a] The values shown are based on our maps containing classified and selected atoms which contribute to the AB potential.

the hydrophobic domain which anchors the protein to the membrane), *Klebsiella aerogenes* urease, human serum albumin (HSA), and hen egg white lysozyme (HEL). Projections onto cylinders parallel to each of the three Cartesian axes were made.

Data Simulation. The AB distribution found on protein surfaces was compared with the following situation: two species of particles—N_B "black" particles B representing electron acceptors and N_W "white" particles W representing electron donors—were randomly distributed in a plane of area A at average densities $\rho_B = N_B/A$ and $\rho_W = N_W/A$, respectively. The values for A, N_B, and N_W vary from one protein to another (Table 2) and had to be set correspondingly to allow a comparison between this simplified model and the results based on our maps. N_B stands for the number of protein atoms which only accept electrons ($\gamma^- = 0$ in Table 1). All other atoms present on the proteins surface and listed in Table 1 were counted as electron donors and therefore contributed to N_W. This is even true for amino acid residues containing equal numbers of electron donors and acceptors because on the basis of the real abundances of all γ^+ and γ^- atoms present on a protein surface, the black particles B have to be weighted by a factor w_B which, for all the proteins investigated so far, is much smaller than w_W (Table 2). As mentioned above, $N_B + N_W$ particles were randomly distributed in A, and the probability that a region of area $a < A$ contains a white majority; i.e.

$$\rho_W a w_W = N_W^a w_W > N_B^a w_B = \rho_B a w_B \tag{2}$$

had to be found. The probability P of having x particles in a certain area is given by a Poisson distribution

$$P(x;\mu) = \frac{\mu^x}{x!} e^{-\mu} \tag{3}$$

where μ stands for the mean of x. Applied to our situation, the probability that a region of area a has a white majority can now be written as

$$P_W(\rho_W,\rho_B,a) = \sum_{N_B^a=0}^{\infty} P(N_B^a;\rho_B a)\left(1 - \sum_{j=0}^{\Phi} P(j;\rho_W a)\right)$$

$$\Phi = N_B^a \frac{w_B}{w_W} \tag{4}$$

$P(N_B^a;\rho_B a)$ describes the probability of finding N_B^a black particles in a region of area a, and the second term stands for the probability that there are more than $N_B^a(w_B/w_W)$ white particles in the same area. Therefore, eq 4 describes the probability with which white particles fulfill the condition prescribed in eq 2. An upper limit of 40 instead of infinity was found to be accurate enough for the first summation term (by making a term-by-term comparison to an analytically solvable series, one can show that using the parameters in Table 2 P_W is underestimated by $<0.3\%$). The corresponding probability of a black majority in a can be determined analogously by

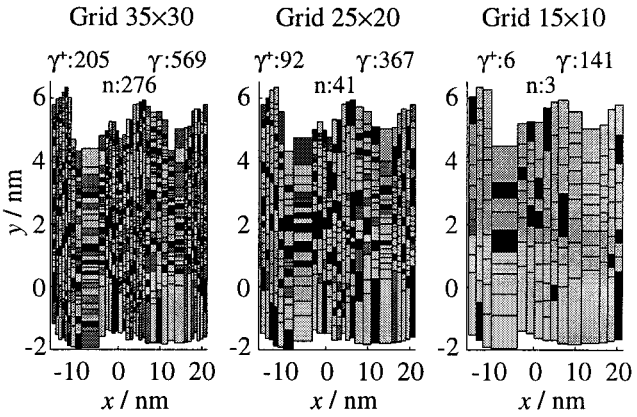

Figure 3. Cell-size-dependent distribution of the Lewis acid−base potential of the HSA protein surface projected into the plane: black, γ^+; dark gray, AB-neutral (n); and light gray, γ^-. The number of cells in each AB class (Table 1) are given.

$$P_B(\rho_W,\rho_B,a) = \sum_{N_W^a=0}^{\infty} P(N_W^a;\rho_W a)\left(1 - \sum_{j=0}^{\psi} P(j;\rho_B a)\right)$$

$$\psi = N_W^a \frac{w_W}{w_B} \tag{5}$$

The situation described above was also simulated numerically: two types of randomly distributed particles were overlaid by a grid of uniformly sized cells whose contents were summed as for the protein surface maps, thereby weighting the particles. All the relevant parameters (see Table 2) were the same as those described above. This procedure yielded the net AB characteristic for each cell (γ^- or γ^+), and a plot of the corresponding ratio vs cell size was compared with the quotient $P_W(\rho_W,\rho_B,a)/P_B(\rho_W,\rho_B,a)$ (eqs 4 and 5) for varying a. To study the effect of nonrandomly distributed particles, we used the same kind of simulation to generate accumulations of one particle type. To do so, a varying percentage of either white or black ones were randomly distributed in two regions of size $A/7$, while all remaining particles were randomly distributed in A.

To investigate the homogeneity of the AB potential on a protein surface, the area-dependent ratio of $P_W(\rho_W,\rho_B,a)$ to $P_B(\rho_W,\rho_B,a)$ was compared with the ratio of the grid cells in a map showing a renormalized γ^+ characteristic to those of type γ^-.

Results & Discussion

Figures 3 and 4 show some typical results for the example of HSA. Two general features emerged for all the proteins examined (Figure 3): (a) as cell size increases, all cells tend to become γ^-, and (b) no specific pattern of areas with γ^+, AB-neutral, or γ^- characteristics is apparent. Selecting only protruding protein surface areas[12] did not affect these general observations; in other words, there seems to be no correlation between surface topography and AB potential.

Observation (a) is in accord with the general preponderance of γ^- in nature.[14] Figure 4 shows a histogram of frequency versus the amount of cells with γ^+, AB-neutral, and γ^- characteristics, which was determined by decreasing the grid

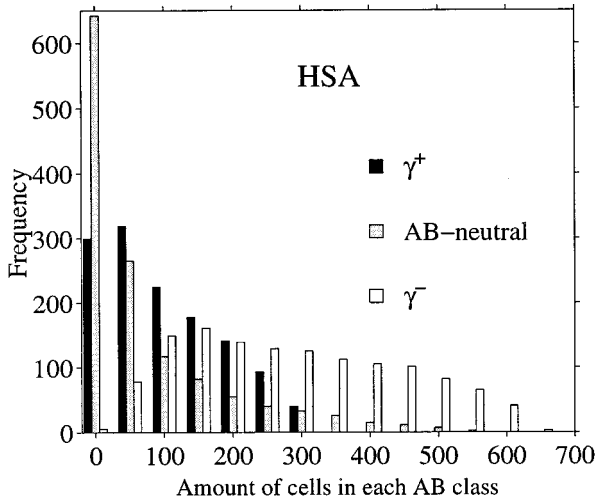

Figure 4. Histogram showing the cell-size-dependent frequency of γ^+, AB-neutral, and γ^- cell areas on the surface of HSA. The grid size varies from 4 × 4 to 39 × 39.

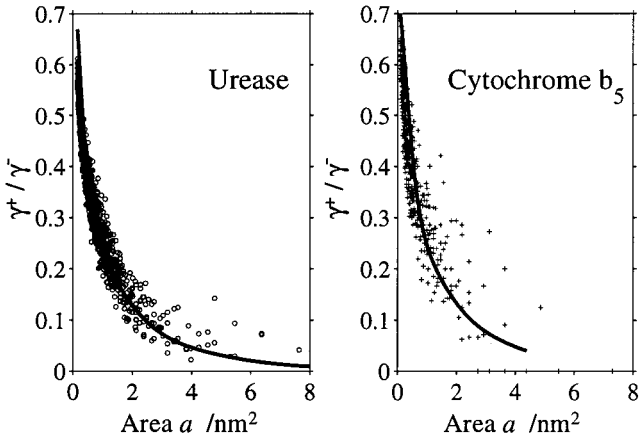

Figure 5. Ratio R_a (eq 1) for urease (○, left graph) and cytochrome b_5 (+, right graph) vs grid cell size, as obtained from the protein maps (all three orthogonal projections). Also shown (continuous lines) are calculated curves assuming that all electron donors and acceptors are randomly distributed on the protein surface (eqs 4 and 5).

cell size from $A/16$ to $A/1521$. To investigate the AB distribution on a protein surface in a more quantitative way, we calculated the ratio of the number of cells with a renormalized γ^+ characteristic to the number with γ^- for various cell sizes. The corresponding curves for four proteins (Figures 5 and 6) were compared (as outlined under the Methods section) with a system containing two types of randomly distributed particles. According to the simulation results, the AB distribution on the surface of urease and cytochrome b_5 (Figure 5) can be very well approximated by a random distribution of electron donors and acceptors. The output of the numerical simulation is in good agreement with the analytical predictions (eqs 4 and 5). Scattering in the data obtained from the protein maps is due to the fact that the shapes of the cells change somewhat from one projection to another. However, the results are essentially independent of the axis chosen for the cylindrical projection. This result is in sharp contrast to the generally nonuniform distribution of electrostatic charges.[9,15] In the cases of HSA and HEL (Figure 6), a random distribution of γ^+ and γ^- does not reflect the actual distribution found from the maps. The corresponding curve for HSA (Figure 6, left) is reasonably matched by 15% of the γ^- particles accumulated in two areas of size $A/7$, while the rest are distributed on A randomly. The results for HEL show the same overall tendency as seen with

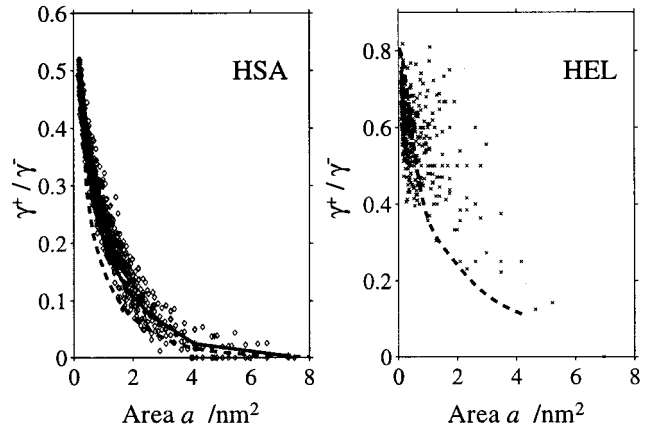

Figure 6. Ratio of γ^+ to γ^- (eq 1) for HSA (◇, left graph) and HEL (×, right graph) as a function of a. Dashed lines describe the corresponding ratio if electron donors and acceptors are randomly distributed on the protein surface (eqs 4 and 5). For HSA ($A = 262$ nm^2), the result of a simulation based on 15% nonrandomly distributed γ^- particles is displayed as well (continuous line).

the other proteins, but the data are less clear and therefore difficult to interpret quantitatively.

γ^+ and γ^- can be obtained from determinations of the contact angles of various liquids on layered protein films, and results are available for HSA (78% hydrated) and hydrated HEL, for which the γ^+/γ^- ratios are 0.12 and ~0.08 respectively.[1]

Although water molecules reorganize themselves in the vicinity of a protein, they will do so such that just above a protein surface they will still reflect the underlying AB distribution, since (for example) a donor group will hydrogen bond to the solvent acceptor moiety, leaving the solvent donor moiety outermost. Hence, surface tension measurements should properly reflect the underlying AB distribution. By comparison with our plot (Figure 6, left), the HSA value corresponds to a cell area of 2.1 nm^2. It has long been a puzzle why HSA adsorbs strongly to glass and other predominantly γ^- metal oxides.[16,17] Using the published γ^+ and γ^- single substance surface tensions for HSA and assuming the protein to be a sphere, we found the calculated interaction energy U[1,2,18] to be +23 mJ/nm^2, i.e., repulsive. A ratio γ^+ over γ^- of 0.36, corresponding to $a = 0.5$ nm^2, gives $U = 0$. Hence, the puzzle could be resolved by taking the characteristic length for adsorption to be smaller than that implied by surface tension measurements (see ref 19 for some discussion of the latter).

References and Notes

(1) van Oss, C. J. *Forces Interfaciales en Milieux Aqueux*; Masson: Paris, 1996.

(2) Cacace, M. G.; Landau, E. M.; Ramsden, J. J. *Q. Rev. Biophys.* **1997**, *30* (3), 241.

(3) Ramsden, J. J. *Colloids Surf. B* **1999**, *14*, 77.

(4) Ramsden, J. J.; Vergères, G. *Arch. Biochem. Biophys.* **1999**, *371*, 241.

(5) van Oss, C. J. In *Protein Interactions*; Visser, H., Ed.; VCH Verlag: Weinheim, Germany, 1992; pp 25−55.

(6) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379.

(7) Koppenol, W. H.; Margoliash, E. *J. Biol. Chem.* **1982**, *257*, 4426.

(8) Kozack, R. E.; d'Mello, M. J.; Subramaniam, S. *Biophys. J.* **1995**, *68*, 807.

(9) Roush, D. J.; Willson, R. C. *Biophys. J.* **1994**, *66*, 1290.

(10) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 99.

(11) Sanner, M. F.; Olson, A. J.; Spehner, J. *Proc. 11th ACM Symp. Comput. Geom.* **1995**, C6.

(12) Topographical information could be readily added to the maps by letting a probe sphere of radius 1.0 nm roll over the protein surface. This created a second surface touching the first one at protruding structures, which could be relevant for modulating protein−surface interactions.

(13) Durley, R. C. E.; Mathews, F. S. *Acta Crystallogr., Sect. D* **1996**, *52*, 65.

(14) van Oss, C. J.; Giese, R. F.; Wu, W. *J. Adhesion* **1997**, *63*, 71.

(15) Haynes, C. A.; Norde, W. *Colloids Surf. B* **1994**, *2*, 517.

(16) Kurrat, R.; Prenosil, J. E.; Ramsden, J. J. *J. Colloid Interface Sci.* **1997**, *185*, 1.

(17) Kurrat, R.; Ramsden, J. J.; Prenosil, J. E. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 587.

(18) Ref 3 may be consulted for more details. The electrostatic contribution was assumed to be negligible.

(19) de Gennes, P. G. *Rev. Mod. Phys.* **1985**, *57*, 827.