Are Solvation Free Energies of Homogeneous Helical Peptides Additive?

René Staritzbichler,[†,‡] Wei Gu,[†] and Volkhard Helms*[,†]

Center for Bioinformatics, Saarland University, Saarbrücken, Germany, and Max-Planck-Institute for Biophysics, Frankfurt, Germany

Received: May 9, 2005; In Final Form: August 9, 2005

We investigated the additivity of the solvation free energy of amino acids in homogeneous helices of different length in water and in chloroform. Solvation free energies were computed by multiconfiguration thermodynamic integration involving extended molecular dynamics simulations and by applying the generalized-born surface area solvation model to static helix geometries. The investigation focused on homogeneous peptides composed of uncharged amino acids, where the backbone atoms are kept fixed in an ideal helical conformation. We found nonlinearity especially for short peptides, which does not allow a simple treatment of the interaction of amino acids with their surroundings. For homogeneous peptides longer than five residues, the results from both methods are in quite good agreement and solvation energies are to a good extent additive.

1. Introduction

It has been well recognized that solvation effects play a crucial role in almost every process in molecular biology, for example, in protein folding and the molecular recognition among proteins or for the aggregation of transmembrane helices.[1−8] All these processes are associated with the transfer of a solute, mostly amino acids or proteins, between a polar solvent with a high dielectric constant and a nonpolar medium. During transfer, a set of noncovalent contacts is formed or broken within the solute molecules and between solute and solvent molecules. The accurate description of solvation effects is therefore an essential part of any systematic approach aiming at contributing to the understanding of such processes.

Over the past decades, many experimental studies have addressed the solvation properties of amino acids as well as of peptides.[9−14] However, experimental techniques have not been able to determine solvation free energies of the charged amino acids, and it is hard or even impossible to control the backbone conformations of the peptides in experiment. Theoretical modeling of biological systems is thereby highly desirable to complement experimental studies.[3,15−18]

When applying computational methods for deriving solvation free energies of peptides or proteins, one constantly faces the dilemma of achieving both physical accuracy and computational efficiency. The most reliable theoretical methods available today are free energy calculations that have been thoroughly refined during the 1990s allowing for systematic studies of solvation properties. Two variants of these are free energy pertubation[19] and thermodynamic integration.[20] Recent studies on the hydration free energies of amino acid side-chain analogues[15−18] using multiconfiguration thermodynamic integration (MCTI)[20] with separation-shifted potential scaling[21,22] achieved very satisfactory agreement with experimental studies. However, this method requires an explicit representation of solvent molecules, and the results crucially depend on how complete the relevant parts of the conformational space were sampled. These requirements render the method computationally very expensive or even prohibitive when applied to large systems such as proteins.

Implicit solvent models reduce the explicit interactions between solute and solvent to a mean field property that only relies on the solute conformation.[3,23,24] Therefore, they are currently heavily used in areas comprising protein structure prediction,[25−30] protein folding,[4−8,31−35] and modeling protein−protein/ligand interaction.[4,36−39] All these implicit solvent models assume, either in part[23,24,40−43] or completely,[44−52] that solvation free energy contributions due to neighboring segments are additive. Whereas additivity is certainly not fulfilled for charged amino acids, this assumption is based on the idea that the interactions of polar and nonpolar side chains affecting the solvent structure are of the short-range nature. Supporting evidence comes from an experimental study of solubilities of the peptide backbone unit in various solvents.[14] There, backbone transfer free energies were found to be additive. On the other hand, a theoretical study of the formation of secondary structure observed nonadditivity for the free energies of the formation of short α-helices using the finite difference Poisson−Boltzmann method.[53] It appears that this implicit assumption of solvation free energies being additive has not comprehensively been tested so far.

Recently, parametric studies of implicit solvent models are focused on closely matching the data from experimental and explicit solvent simulations of small molecules.[43,54,55] However, if solvation free energies of neighboring segments are not additive, would it still be suitable to extend the parameters derived from data for small molecules to large systems or, the other way around, to apply models parametrized for large systems to small molecules?

What is an appropriate method to answer these questions? Experimental results suffer from shortcomings when decomposing results into sequence-dependent and conformation-dependent contributions. Another problem concerns the solubility of peptides that often requires the addition of blocking groups. Fortunately, such issues are less of a problem in theoretical studies. As mentioned before, the most reliable theoretical method available is multiconfiguration thermodynamic integration (MCTI). However, due to the large computational efforts

* To whom correspondence should be addressed. E-mail: volkhard.helms@bioinformatik.uni-saarland.de.
† Saarland University.
‡ Max-Planck-Insitute for Biophysics.

10.1021/jp052403x CCC: $30.25
Published on Web 09/20/2005

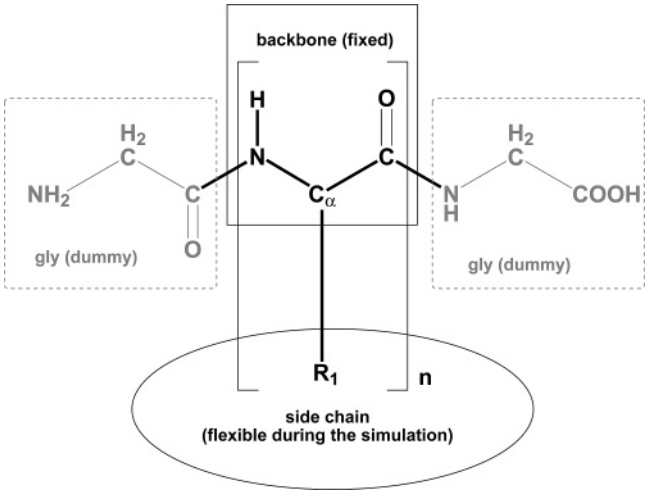

Figure 1. Structure of the system used for MD simulations and MCTI calculations. The atoms of the central residues are shown in bold and dummy atoms are colored in gray.

involved, applications of MCTI to the computation of solvation free energies were so far restricted to single amino acids. This study is the first attempt to tackle polypeptide systems up to nine residues in length. Therefore, an important test was to compare the results from MCTI calculations with GBSA, one of the most popular and efficient implicit solvent models.

As an extension of our previous work[15] and in order to combine this work with an ongoing project in our group, designing a residue scale force field for the structure prediction of transmembrane proteins,[56] we chose homogeneous α-helical peptides of different lengths as model systems for this study. This choice was motivated by the following considerations: (1) restricting the peptide backbone to a given conformation facilitates sampling during the simulations, (2) focusing on homogeneous α-helical peptides keeps the sequence-dependent contributions obvious and understandable, and (3) by comparing the results for different types of amino acid residues, one may attempt to dissect the backbone contributions from the side-chain contributions. We note, though, that backbone and side-chain contributions are commonly interdependent and a true separation is not possible in a strict sense.

2. Methods

Molecular Dynamics Simulations. This study addresses the solvation properties of homogeneous α-helices composed of uncharged amino acids. The coordinates of such α-helices with a length of five were modeled using the TINKER[57] package. Each peptide $(X)_n$ is flanked by two glycine residues of the form Gly−$(X)_n$−Gly (see Figure 1). The atoms of the two flanking glycine residues were treated as "dummy" atoms (see below under Free Energy Calculations). The systems are named GX5G (X refers to the one letter code of amino acids). For the cases of alanine and asparagine, we also investigated their homogeneous α-helices of lengths 2, 3, 4, 5, 6, and 9 (named GX*n*G, where *n* refers to the number of residues). The dihedral angles of the peptide backbone were set to the values of an ideal α-helix ($\phi = -58°$, $\psi = -47°$). For comparison, two five-residue-long homogeneous peptides were also modeled in an extended conformation ($\phi = -135°$, $\psi = 135°$) (named GA5GST and GN5GST). During the simulations, all backbone atoms were kept fixed in their starting geometry because we wanted to investigate the effect of the helical geometry. We note, of course, that an α-helix may not be the preferred conformation in solution for some of the investigated sequences. Molecular dynamics simulations were performed in both chloroform and water as the solvents.

All simulations were carried out using the NWChem 4.5 package[58] with the AMBER99 force field.[59] The atomic charges of the chloroform model were −0.3847 e for the carbon atom, 0.2659 e for the hydrogen atom, and 0.0396 e for the chlorine atoms, respectively. The molecules were solvated in cubic boxes of 4.0 nm side length, using chloroform or TIP3P water molecules,[60] respectively, with an initial minimum distance of at least 1.3 nm between the boundaries of the box and the nearest solute atom (excluding dummy atoms). All coordinate sets were first optimized by 500 steps of steepest-descent energy minimization. The solvent and modeled residues were then relaxed during a 1 ns molecular dynamics (MD) simulation at 300 K prior to the free energy calculation. The SHAKE procedure[61] was applied to constrain all bonds that contained hydrogen atoms. The time step of the simulations was 2 fs throughout. Nonbonded interactions were treated using a cutoff of 1.2 nm. The temperature and pressure were maintained by weak coupling to an external bath in all simulations.[62] For the simulations in chloroform, the pressure coupling time was set to 5.0 ps and the isothermal compressibility was set to 9.98×10^{-10} m^2 N^{-1}. For the simulations in water, the coupling time and compressibility were 0.5 ps and 4.53×10^{-10} m^2 N^{-1}, respectively.

Free Energy Calculations. The solvation free energies of all peptides were calculated according to the following thermodynamic cycle

$$\Delta G_{\text{solv, peptide}} = \Delta G_{\text{peptide}\rightarrow\text{dummy,vacuum}} - \Delta G_{\text{peptide}\rightarrow\text{dummy,solvent}} + \Delta G_{\text{solv,dummy}} \quad (1)$$

$\Delta G_{\text{peptide}\rightarrow\text{dummy,vacuum}}$ and $\Delta G_{\text{peptide}\rightarrow\text{dummy,solvent}}$ are the free energy differences for switching off the solute−solvent nonbonded interactions (van der Waals interactions and electrostatic interactions) while keeping their bonded interactions and atomic masses unchanged. $\Delta G_{\text{solv,dummy}}$, the free energy change for transferring dummy atoms from vacuum into solvent, is zero by definition. Solvation free energies calculated by such thermodynamic cycles are concentration independent.[63]

The free energy difference between two states of a system ($\lambda = 0$ and $\lambda = 1$) described by their Hamiltonians $H(\lambda = 0)$ and $H(\lambda = 1)$ can be obtained by thermodynamic integration (TI)[19,64]

$$\Delta G = \int_1^0 \left(\frac{\partial G(\lambda)}{\partial \lambda}\right)_\lambda \mathrm{d}\lambda = \int_1^0 \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \quad (2)$$

where λ is a control variable that determines the state of the system and the brackets denote taking the ensemble average at a particular value of λ. As the perturbation from $\lambda = 0$ to $\lambda = 1$ is performed in discrete steps, the integral is evaluated as a sum of ensemble averages in the MCTI method[20]

$$\Delta G = \sum_i \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda \Delta\lambda_i \quad (3)$$

where i is the index of different values of λ, and $\Delta\lambda_i$ is the difference between successive values of λ. An estimate for the statistical error of the ΔG can be computed by

$$E(\Delta G) = \sqrt{\sum_i E^2{}_i \Delta\lambda_i} \quad (4)$$

where $E_i = E(\langle \partial H(\lambda)/\partial\lambda \rangle_i)$ is the statistical error for each window

at a particular value of λ_i, see eq 3,[20] which is determined through a correlation analysis approach.[65]

The nonbonded interactions between the initial state and the final state are interpolated by a separation-shifted potential scaling[21] using $\delta = 0.075$ nm to avoid the well-known origin singularities. In our study, separate simulations were performed at 21 equally spaced points of λ from $\lambda = 0$ to $\lambda = 1$. At each point, the system was first equilibrated for 200 ps and data were collected during further 200 ps of simulation. The van der Waals and Columbic terms were turned off simultaneously. A similar protocol was previously used to compute solvation free energies of small model substances.[66] There, ΔG_{hydr} could be reliably computed with statistical errors of ≤ 1.5 kJ mol^{-1}. The protocol is also similar to the recent studies of Gu et al.,[15] Villa and Mark,[16] Shirts et al.,[17] and Deng and Roux[18] to compute ΔG_{hydr} for amino acid side-chain analogues. The convergence of the derivatives of the Hamiltonian with respect to λ was monitored for all individual windows and showed smooth behavior for all computed values (data not shown).

The MCTI calculations were performed under constant pressure conditions. Consequently, upon mutation of the peptide into a dummy molecule, the volume of the simulation box shrank. When free energies of solvation are computed, the values are concentration independent as noted previously.[63] As this study reports for the first time the application of MCTI to remove an entire peptide, the volume of the simulation box was checked during the MCTI calculation for GN9G, since GN9G has the largest solute volume in this study. In the first window, the volume of the simulation box is 64.0 ± 0.2 nm^3. In the last window of the MCTI calculation, in which the solute becomes invisible to the solvent molecules, the volume of the simulation box is 62.9 ± 0.2 nm^3. The observed volume difference of 1.1 ± 0.3 nm^3 is in reasonable agreement with the volume of the simulated asparagine 9-mer of 0.8 nm^3 (contact/reentrant volume, calculated using TINKER with a probe radius of 0.14 nm). The entropy changes due to modification of the water−peptide interactions are all taken into account by the MCTI method. As long as the peptide atoms are interacting with the solvent molecules, the volume occupied by the peptide is inaccessible to the solvent. When the peptide interactions are switched off, the volume of the simulation box shrinks by an amount comparable to the volume of the peptide as required by the condition of constant density. Therefore, the translational entropy of the bulk water molecules in the box remains unchanged.

For comparison, one of the most popular implicit solvent models, the generalized-born surface area model (GBSA)[23] implemented in the TINKER package,[57] was used to calculate the solvation free energies of homogeneous α-helices from a length of 1 up to 20 residues. All peptides are capped with ACE-($CH_3C{=}O$) at the N-terminus and $-NH_2$ at the C-terminus. The contributions of the capping groups are subtracted from the total solvation free energies.

3. Results

Solvation free energies of homogeneous helical peptides were computed from molecular dynamics simulations where during the simulation the interactions between solute and solvent are progressively switched off (see Methods). To derive the solvation free energies of the peptides in water or chloroform, respectively, thermodynamic cycles are constructed where the vacuum values are subtracted from those in water or in chloroform. Table 1 lists the values of all peptides with a length of five (both helical and extended conformations) in water and in chloroform as well as the ratio

TABLE 1: Solvation Free Energies of Five-Residue-Long Homogenous Peptides Calculated with MCTI

residue	$\Delta G^{(5)}_{H_2O}$ (kJ/mol)	ratio c	$\Delta G^{(5)}_{HCCl_3}$ (kJ/mol)	ratio c
GY5G	-200.5 ± 4.5	0.91	-189.7 ± 3.5	0.77
GW5G	-180.7 ± 4.8	0.85	-194.6 ± 3.7	0.71
GV5G	-100.3 ± 3.8	1.06	-117.9 ± 2.6	0.75
GC5G	-139.8 ± 3.4	0.93	-142.0 ± 2.5	1.04
GF5G	-115.7 ± 4.4	1.01	-161.3 ± 3.1	0.78
GG5G	-156.1 ± 2.8	1.27	-102.0 ± 1.9	1.05
GI5G	-92.7 ± 4.0	1.01	-125.8 ± 2.8	0.74
GL5G	-88.2 ± 4.0	1.08	-125.1 ± 2.9	0.78
GM5G	-111.1 ± 4.1	1.06	-147.0 ± 2.8	0.91
GN5G	-295.2 ± 3.6	0.89	-156.7 ± 2.8	0.61
GT5G	-140.4 ± 3.9	0.86	-116.9 ± 2.7	0.83
GS5G	-183.8 ± 3.6	0.95	-115.3 ± 2.4	0.80
GA5G	-128.4 ± 3.0	1.28	-106.3 ± 2.2	0.97
GA5GST	-113.1 ± 1.1	1.13	-110.2 ± 2.1	1.01
GN5GST	-262.2 ± 1.6	0.79	-152.4 ± 2.7	0.59

$$c = \frac{\Delta G^{(n)}}{n\Delta G^{(1)}} \tag{5}$$

This ratio is the solvation free energy of the whole peptide ($\Delta G^{(n)}$) divided by the solvation free energy of the single amino acids ($\Delta G^{(1)}$ taken from ref 15) multiplied by the number of residues n. In this case $n = 5$.

In a couple of cases, a rather surprising result is obtained: ratio c is greater than 1. This means that the solvation free energy of the entire peptide is larger than the sum of the single contributing values. Subsequently, we refer to this effect as "superunity".

If one considers applying simple residue scaled models to the solvation free energy of peptides, one may formulate

$$\Delta G^{(n)} = \sum_i S_i \cdot \Delta G_i^{(1)} \tag{6}$$

Here, $\Delta G_i^{(1)}$ should be the solvation free energy of residue i as a single residue and S_i is the ratio of the solvent accessible surface area (SASA) of this residue in the peptide context in a particular conformation (here, helical) relative to the SASA value of the isolated residue ($S_i \leq 1$). When eqs 5 and 6 are combined, it follows that ratio c must be less than 1 and more or less constant, which implies a linear behavior of the solvation free energy. We conclude that models that are purely based on SASA terms fully depend on the linearity of solvation free energies.

To investigate this effect in more detail for two selected systems, we performed MCTI calculations for homogeneous α-helical peptides of 2−6 and 9 residues in length. Alanine and asparagine were selected for the calculations as examples of nonpolar and polar residues. Due to their different sizes, we also expected different contributions from the residue backbone and the side chain. Water and chloroform were selected as the solvent environments to study the effects of different dielectric constants and different sizes of the solvent molecules. In our previous study on individual amino acid ΔG_{solv} values, the two solvents yielded results in very good agreement with experimental data. To compare with available implicit solvent models, solvation free energies of peptides with lengths of 1−20 residues were computed by the GBSA model as well.

MCTI: In Water. Figure 2 shows the results for poly-Ala (GA*n*G, *n* refers to the number of residues) and poly-Asn

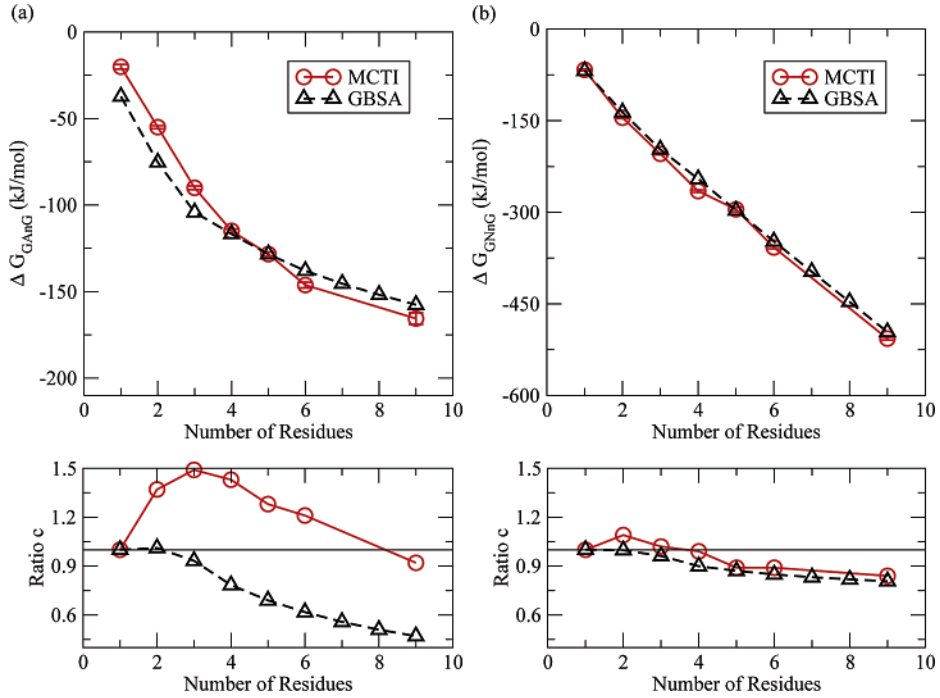

Figure 2. Solvation free energies for polyalanine peptides (left) and polyasparagine peptides (right) of different lengths in water from MCTI and GBSA calculations.

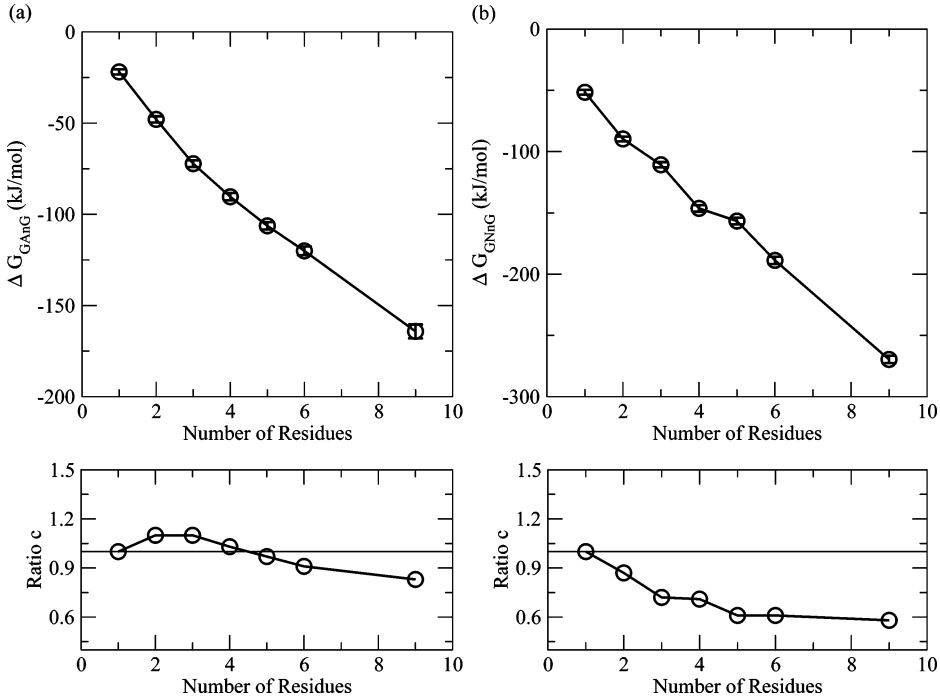

Figure 3. Solvation free energies for polyalanine peptides (left) and polyasparagine peptides (right) of different lengths in chloroform from MCTI calculations.

(GN*n*G) of different lengths in aqueous solution using MCTI. For comparison, corresponding results calculated by GBSA are shown as well.

Figure 2 reveals three noteworthy features. First, the curve shapes for GA*n*G and GN*n*G are different: a nonlinear behavior was found in the GA*n*G calculation, while the plot for GN*n*G shows a nearly linear behavior. Second, superunity ($c > 1$) is observed for both systems but is more strongly emphasized in GA*n*G. Third, the results from the two different approaches (MCTI and GBSA) show surprisingly good agreement, especially for GN*n*G. Sizable differences still exist for GA*n*G where ratio c remains greater than 1 up to the maximum length of 9 residues in the MCTI calculations, while in the GBSA calculations it reaches a value smaller than 1 for $n > 3$. For GN*n*G, the ratio shows a very similar trend in both calculations: it is below 1 for $n > 3$ and reaches a relatively constant value of 0.8−0.9.

In the two cases investigated for extended conformations (GA5GST and GN5GST, see Table 1), the superunity of the solvation free energies is more weakly emphasized than those in the cases of helical conformation. This indicates that the superunity is conformationally dependent in water solution.

MCTI: In Chloroform. Figure 3 shows the results for GA*n*G and GN*n*G in chloroform solution using MCTI.

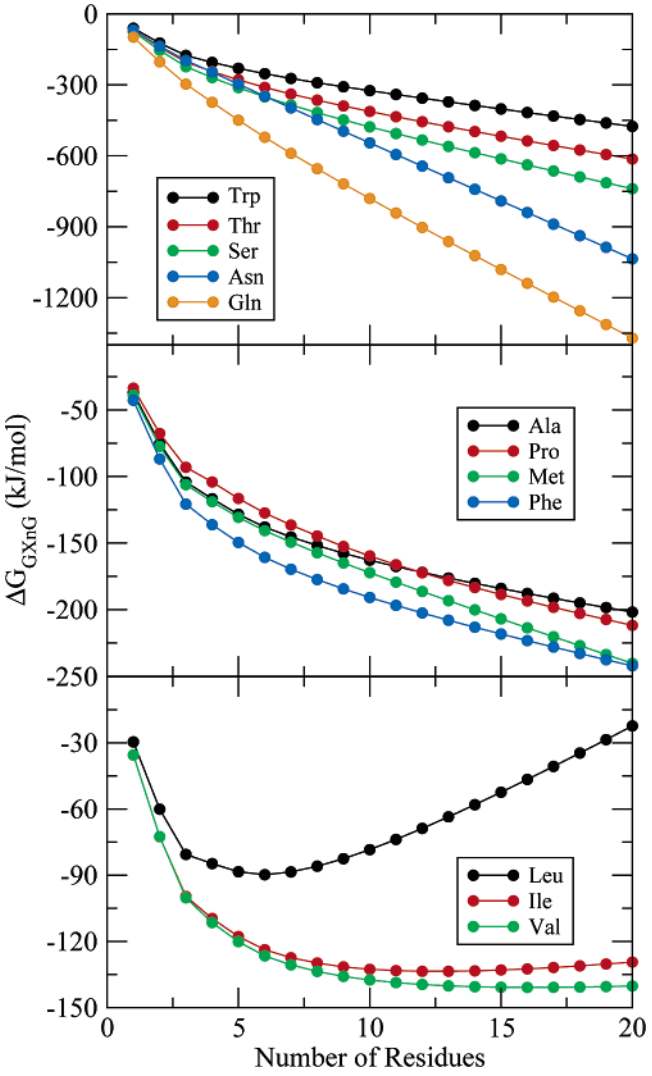

Figure 4. Solvation free energies for homogeneous α-helical peptides of different lengths calculated by the GBSA implicit solvent model.

The results are quite different from those in water as a nearly linear behavior was observed in both systems. Ratio *c* reaches relatively constant values in both systems. For GA*n*G calculations, it reaches ca. 0.8 for $n > 6$, and for GN*n*G, it approaches 0.6 for $n \geq 5$. Superunity was only found in the GA*n*G calculations for $n < 5$. In the GN*n*G calculation, no superunity was found. In this solvent, the results of GA5GST and GN5GST are quite close to those of GA5G and GN5G (see Table 1). The influences due to different backbone conformations are not as large as those in the water solution.

GBSA. The solvation free energies of homogeneous helical peptides from lengths of 1−20 residues are shown in Figure 4.

The peptides can be grouped into three classes according to the properties of their amino acids. The first class comprises the polar amino acids: Asn, Ser, Gln, Hid (Histidine with a proton at $N_{\delta 1/\pi}$), Hie (Histidine with a proton at $N_{\epsilon 2/\tau}$), Thr, Trp, and Tyr that show a very steep descent and reach values between −400 and −1400 kJ mol^{-1} for a 20-residue-long peptide. Linear behavior is clearly observed for all members of this class for $n > 5$.

The second class is formed by the nonpolar amino acids: Ala, Met, Phe, and Pro. They show a clearly nonlinear behavior for small peptides until $n > 10$ where they converge into a linear regime. Even though the amino acids in this class are nonpolar residues, the solvation free energies still decrease when the number of residues increases.

Ile, Val, and Leu, which contain aliphatic side chains, constitute the third class. The minimum solvation free energies reach −140 to −90 kJ mol^{-1} before reversing their slope. Furthermore, they converge to a linear behavior only for long peptides ($n \geq 10$ for Leu and $n \geq 15$ for Val and Ile). Leu clearly shows strongly opposing contributions. The linear part for long peptides shows that Leu is unfavorable in an aqueous environment. The negative slope for smaller peptides reflects the effect of unsaturated hydrogen bonds as well as other effects (see Discussion).

4. Discussion

Nonadditivity and Superunity. Both methods, the MCTI calculations and GBSA calculations, show nonlinearity for short peptides in most cases investigated. Basic considerations show that modeling helical peptides by adding one residue after the other will lead to some discontinuities in the solvation free energies. Figure 5 shows the backbones of peptides of different length ($n = 1-5$).

Up to a length of four residues there exists only next neighbors in the same turn, and their backbone peptide bonds do not form direct interactions (hydrogen bonds). Therefore, the contributions to the solvation free energy of each residue may be almost independent. From five residues on, there are additionally next-turn neighbors which will form intermolecular hydrogen bonds between backbone atoms. The interactions between the newly added NH group at position *n* and the surrounding solvent molecules are shielded by the C═O group of residue $n - 4$ because of this intermolecular hydrogen bond. As a consequence, the interactions between the C═O group at position $n - 4$ and solvent are shielded by the same hydrogen bond as well. This means that from four residues on, the number of "unsaturated backbone groups" (backbone N−H or C═O that are not involved in intrapeptide hydrogen bonds) does not increase when the helix is extended. Table 2 lists the number of unsaturated backbone groups as a function of *n*. This number remains constant when $n \geq 4$.

Solvation free energies of organic molecules are commonly decomposed into a nonpolar term and a polar term. The nonpolar term, which includes the energy cost to form a cavity for the solute in the solvent and to establish van der Waals interactions between the solute and the solvent molecules, is in principle additive with respect to the number of peptide residues *n*. The polar term includes electrostatic interactions (monopole, dipole, and higher order multipoles). This term is most likely not additive per se. In a helical peptide, all dipoles of the backbone point in the same direction and therefore form an overall dipole along the helical axis.[67] This dipole will align water molecules in the solvation shell around the peptide. Concerning the scaling of this contribution with the peptide length *n*, the first residue induces orientational polarization of all solvent molecules inside a shell around the backbone. When the length of the helical peptide is increased, eventually all solvent molecules will be orientationally polarized within a cylinder around the helical peptide. The volume of this cylinder grows proportionally to *n*. Therefore, the electrostatic contribution of the solvation free energy should approach a linear dependence with peptide length for $n \geq 5-10$ while it may display nonlinearity for shorter peptides. For amino acids with nonpolar side chains, for example, class 2 and class 3 in the GBSA calculation, the contributions of their backbone groups are the dominant terms in water solution. The discontinuity of the solvation free energies

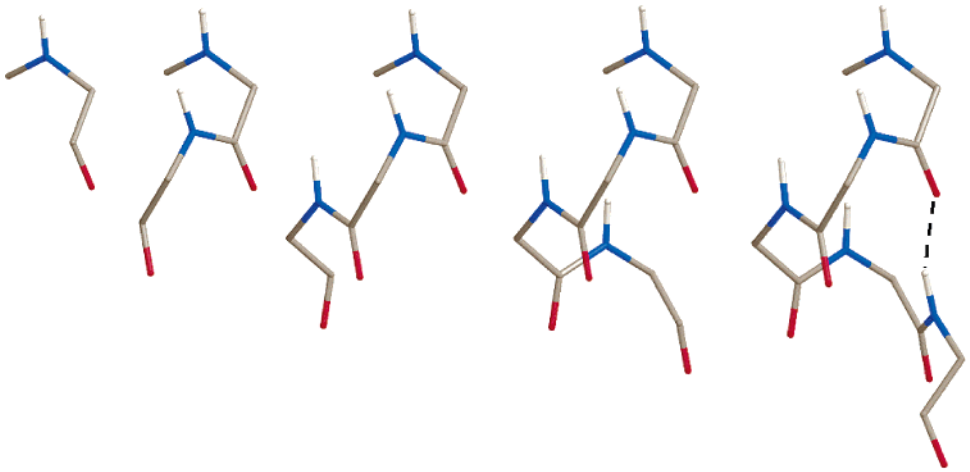

Figure 5. Backbone structures of α-helical peptides of lengths 1−5. Intermolecular hydrogen bonds are shown in dashed lines.

TABLE 2: Number of Unsaturated Backbone Groups ($n_{backbone}$) as a Function of the Number of Residues (n)

n	$n_{backbone}$
1	2
2	4
3	6
≥4	8

(see Figures 2 and 4) can thus be traced back to the discontinuity of the electrostatic contribution of the free backbone groups (see Table 2). Clearly, this behavior is more noticeable in water than in chloroform. Through the increase of the number of residues n, the contribution of the dipole term increases linearly whereas the overall contribution of the electrostatic term is limited by the number of unsaturated groups. Therefore, the solvation free energy as a function of n becomes linear and additive for longer peptides (e.g., $n \geq 10$). For amino acids with polar side chains or very large side chains, the electrostatic contribution of the side chains is comparable to that from the backbone groups. As the number of side chains and the solvent surface area grow approximately linearly with the number of residues, this term is nearly linear. In addition, polar or very large side chains shield the backbone from the solvent molecules, which somehow weakens the contribution of the backbone. Consequently, the nonadditive effect is less pronounced than that for nonpolar amino acids, and linearity or additivity may be observed for shorter peptides as well. Generally speaking, the shielding of the backbone by large side chains should, however, be more important for longer peptides than for shorter peptides. In other cases, for example, for Leu and Ile, where the solvation free energies increase after a certain inflection, the geometry of their aliphatic side chains leads to a stronger shielding of the backbone and the side-chain contribution becomes dominant. As real systems are composed of a mixture of different amino acid types, such deviations from linearity may partially compensate each other.

The same reasoning can also be applied to the effect of different solvent environments. Figure 3 shows that in chloroform, nearly linear behavior of peptide solvation free energies was observed for both nonpolar and polar amino acids. This is understandable because in a less polar or nonpolar solution, the importance of electrostatic interactions decreases. As a result, the nonpolar term, which is in principle additive, becomes dominant in this case.

The "superunity" effect observed for short peptides with small and nonpolar side chains (e.g., Ala and Gly) can be explained by this reasoning as well. In such peptides, the side-chain contribution is less important in magnitude. When a residue is added to the peptide, a part of the surrounding water molecules has already adapted to the overall dipole of the helix. This reduces the cost of aligning the nearby solvent molecules compared to solvating an individual residue. Therefore, it is more favorable to add an amino acid at the terminus of a short peptide than solvate the first amino acid of this peptide. When the peptide length increases beyond $n = 4$, the involvement of the backbone group in intrahelical hydrogen bonds reduces the contribution of the newly added residues. Thus, superunity does not exist anymore. We note that a 1.2 nm cutoff was applied to all nonbonded interactions in the MCTI simulations for technical reasons. Because the central peptide units will thus experience their electrostatic environment only within a limited range, the use of a cutoff may enhance additivity for long peptides. On the other hand, the MCTI calculations were only performed for systems up to nine residues long. The calculations of this study about additivity for long peptides are mainly based on the GBSA results, where no cutoff was applied.

Implications from Nonadditivity. Hydrophobic free energies are commonly derived from experimentally determined solubilities of small organic compounds such as hydrocarbons ("microscopic").[68−71] In prior parametrizations for the hydrophobic effect, a correlation was proposed between the hydrophobic free energy changes and the SASA.[72,73] The values derived for transfer from vacuum to water are typically in the order of 5−7 cal mol^{-1} Å^{-2}.[74] Can the values derived from small molecules easily be transferred to other scales? Our calculations identified different slopes for short peptides and long peptides that are due to mixed electrostatic and hydrophobic contributions. Deriving a parametrization of the hydrophobic effect would require a careful decomposition of these. However, the good agreement between MCTI and GBSA results for longer peptides indicates that the SASA term in GBSA must work quite well already, yielding a useful parametrization of the hydrophobic effect. Nonetheless, these results indicate that one should apply caution when transferring results that were derived for molecules of different sizes. The respective parametrization should be chosen based on experimental data for the same scale of the problem.

After accepting the consequences of nonadditivity of solvation free energies, it may seem that implicit solvent models based on additivity are problematic per se. Nevertheless, in many atomic scaled implicit solvent models, for example, GBSA, the solvation term is decomposed into polar and nonpolar terms which are individually treated at the atomic scale. Therefore,

these models are likely not affected by the phenomenon of nonadditivity. Only such models that fully rely on the SASA term may need some improvements. Here, a newly developed model of calculating nonlocal electrostatics interactions may be helpful.[75] As the implicit solvent models at the atomic scale are still quite expensive for large scale problems such as flexible protein−protein docking, assembly of transmembrane helices, or protein complexes, we believe that implicit solvent models at the residue scale should have very promising applications in those areas.

5. Conclusions

The conclusions of the present study are restricted to (i) fully homogeneous peptides composed of uncharged amino acids that (ii) are kept in a frozen backbone helical conformation and (iii) are fully solvated. On the basis of the investigated systems, we find:

1. Solvation free energies of peptides of various length were computed by the MCTI and GBSA methodologies. For five or more residues the results are in quite good agreement. This observation gives strong support for our strategy of computing ΔG_{hydr} for peptides up to nine residues from MCTI calculations. However, MCTI and GBSA still show sizable differences for short helices where MCTI should be quite accurate. Thus, it is important to consider molecular details of backbone hydration.

2. Nonadditivity is found by both methodologies for peptides shorter than five residues. On the other hand, according to GBSA calculations, additivity appears fulfilled for helices longer than 10 residues. This points toward using caution when transferring SASA parameters that are extracted on the basis of solubilities or partition coefficients of small molecules to large systems. Alternatively, it may be also problematic to use values that are derived from large systems to small molecules.

3. The design of simplified models, where helices are composed of residue beads and interactions are modeled additively, appears challenging.

Future work is needed that extends investigations of this type to heterogeneous sequences to see if additivity of solvation energies holds in a general sense.

Acknowledgment. NWchem Version 4.5, as developed and distributed by the Pacific Northwest National Laboratory, P.O. Box 999, Richland, WA 99352 and funded by the U. S. Department of Energy, was used for the calculation. This work was supported by Grant I177956 from the Volkswagen Foundation. We thank Dr. Jiang Zhu for valuable discussions on the GBSA model and Dr. Martin Zacharias, Dr. Andreas Hildebrandt, Dr. Michael Hutter, and Dr. Rainer Böckmann for a critical reading of the manuscript.

Supporting Information Available: Table S1 lists the numerical data shown in Figures 2 and 3. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Honig, B.; Yang, A. S. *Adv. Protein Chem.* **1995**, *46*, 27−58.
(2) Dill, K. A. *Biochemistry* **1990**, *29*, 7133−7155.
(3) Feig, M.; Brooks, C. L. *Curr. Opin. Struct. Biol.* **2004**, *14*, 217−224.
(4) Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L. *J. Med. Chem.* **2004**, *47*, 3032−3047.
(5) Gnanakaran, S.; Nymeyer, H.; Portman, J.; Sanbonmatsu, K. Y.; Garcia, A. E. *Curr. Opin. Struct. Biol.* **2003**, *13*, 168−174.
(6) Pitera, J. W.; Swope, W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7587−7592.
(7) Ohkubo, Y. Z.; Brooks, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13916−13921.
(8) Nymeyer, H.; Garcia, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13934−13939.
(9) Fauchere, J. L.; Pliska, V. *Eur. J. Med. Chem.* **1983**, *18*, 369−375.
(10) Kim, A.; Szoka, F. C. *Pharm. Res.* **1992**, *9*, 504−514.
(11) Radzicka, A.; Wolfenden, R. *Biochemistry* **1988**, *27*, 1664−1670.
(12) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849−855.
(13) Wimley, W. C.; Creamer, T. P.; White, S. H. *Biochemistry* **1996**, *35*, 5109−5124.
(14) Auton, M.; Bolen, D. W. *Biochemistry* **2004**, *43*, 1329−1342.
(15) Gu, W.; Rahi, S. J.; Helms, V. *J. Phys. Chem. B* **2004**, *108*, 5806−5814.
(16) Villa, A.; Mark, A. E. *J. Comput. Chem.* **2002**, *23*, 548−553.
(17) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119*, 5740−5761.
(18) Deng, Y. Q.; Roux, B. *J. Phys. Chem. B* **2004**, *108*, 16567−16576.
(19) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.
(20) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1991**, *95*, 1175−1188.
(21) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025−9031.
(22) Beutler, T. C.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *101*, 1417−1422.
(23) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(24) Constanciel, R.; Contreras, R. *Theor. Chim. Acta* **1984**, *65*, 1−11.
(25) Dominy, B. N.; Brooks, C. L. *J. Comput. Chem.* **2002**, *23*, 147−160.
(26) Felts, A. K.; Gallicchio, E.; Wallqvist, A.; Levy, R. M. *Proteins: Struct., Funct,. Genet.* **2002**, *48*, 404−422.
(27) Feig, M.; Brooks, C. L. *Proteins: Struct., Funct., Genet.* **2002**, *49*, 232−245.
(28) Zhu, J.; Zhu, Q. Q.; Shi, Y. Y.; Liu, H. Y. *Proteins: Struct., Funct., Genet.* **2003**, *52*, 598−608.
(29) Forrest, L. R.; Woolf, T. B. *Proteins: Struct., Funct., Genet.* **2003**, *52*, 492−509.
(30) Fiser, A.; Feig, M.; Brooks, C. L.; Sali, A. *Acc. Chem. Res.* **2002**, *35*, 413−421.
(31) He, J. B.; Zhang, Z. Y.; Shi, Y. Y.; Liu, H. Y. *J. Chem. Phys.* **2003**, *119*, 4005−4017.
(32) Zagrovic, B.; Sorin, E. J.; Pande, V. *J. Mol. Biol.* **2001**, *313*, 151−169.
(33) Zhou, R. H. *Proteins: Struct., Funct., Genet.* **2003**, *53*, 148−161.
(34) Suenaga, A. *J. Mol. Struct. (THEOCHEM)* **2003**, *634*, 235−241.
(35) Liu, Y. X.; Beveridge, D. L. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 128−146.
(36) Donnini, S.; Juffer, A. H. *J. Comput. Chem.* **2004**, *25*, 393−411.
(37) Lazaridis, T. *Proteins: Struct., Funct., Genet.* **2003**, *52*, 176−192.
(38) Mardis, K. L.; Luo, R.; Gilson, M. K. *J. Mol. Biol.* **2001**, *309*, 507−517.
(39) Gohlke, H.; Case, D. A. *J. Comput. Chem.* **2003**, *25*, 238−250.
(40) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005−3014.
(41) Lee, M. S.; Feig, M.; Salsbury, F. R.; Brooks, C. L. *J. Comput. Chem.* **2003**, *24*, 1348−1356.
(42) Im, W.; Lee, M. S.; Brooks, C. L. *J. Comput. Chem.* **2003**, *24*, 1691−1702.
(43) Zhu, J. A.; Shi, Y. Y.; Liu, H. Y. *J. Phys. Chem. B* **2002**, *106*, 4844−4853.
(44) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. *J. Comput. Chem.* **2002**, *23*, 517−529.
(45) Wesson, L.; Eisenberg, D. *Protein Sci.* **1992**, *1*, 227−235.
(46) Ooi, T.; Oobatake, M.; Nemethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086−3090.
(47) Ferrara, P.; Apostolakis, J.; Caisch, A. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 24−33.
(48) Weiser, J.; Shenkin, P. S.; Still, W. C. *Biopolymers* **1999**, *50*, 373−380.
(49) Hou, T. J.; Qiao, X. B.; Zhang, W.; Xu, X. J. *J. Phys. Chem. B* **2002**, *106*, 11295−11304.
(50) Zhou, H. Y.; Zhou, Y. Q. *Proteins: Struct., Funct., Genet.* **2002**, *49*, 483−492.
(51) Lomize, A. L.; Riebarkh, M. Y.; Pogozheva, I. D. *Protein Sci.* **2002**, *11*, 1984−2000.
(52) Levy, R. M.; Zhang, L. Y.; Gallicchio, E.; Felts, A. K. *J. Am. Chem. Soc.* **2003**, *125*, 9523−9530.
(53) Yang, A. S.; Honig, B. *J. Mol. Biol.* **1995**, *252*, 351−365.
(54) Nina, M.; Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 5239−5248.

(55) Banavali, N. K.; Roux, B. *J. Phys. Chem. B* **2002**, *106*, 11026−11035.
(56) Park, Y.; Elsner, M.; Staritzbichler, R.; Helms, V. *Proteins: Struct., Funct., Bioinfo.* **2004**, *57*, 577−585.
(57) Ren, P. Y.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933−5947.
(58) Kendall, R. A.; Apra, E.; Bernholdt, D. E.; Bylaska, E. J.; Dupuis, M.; Fann, G. I.; Harrison, R. J.; Ju, J. L.; Nichols, J. A.; Nieplocha, J.; Straatsma, T. P.; Windus, T. L.; Wong, A. T. *Comput. Phys. Commun.* **2000**, *128*, 260−283.
(59) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 2309−2309.
(60) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(61) Ryckaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327−341.
(62) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(63) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.
(64) van Gunsteren, W. F. *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*; ESCOM: Leiden, The Netherlands, 1989.
(65) Straatsma, T. P.; Berendsen, H. J. C.; Stam, A. *J. Mol. Phys.* **1986**, *57*, 89−95.
(66) Helms, V.; Wade, R. C. *J. Comput. Chem.* **1997**, *18*, 449−462.
(67) Hol, W. G.; Halie, L. M.; Sander, C. *Nature* **1981**, *294*, 532−536.
(68) Chothia, C. *J. Mol. Biol.* **1976**, *105*, 1−14.
(69) Hermann, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 4144−4145.
(70) Reynolds, J. A.; Gilbert, D. B.; Tanford, C. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 2925−2927.
(71) Eisenberg, D.; Mclachlan, A. D. *Nature* **1986**, *319*, 199−203.
(72) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379−.
(73) Sharp, K. A.; Nicholls, A.; Fine, R. F.; Honig, B. *Science* **1991**, *252*, 106−109.
(74) Simonson, T.; Brunger, A. T. *J. Phys. Chem.* **1994**, *98*, 4683−4694.
(75) Hildebrandt, A.; Blossey, R.; Rjasanow, S.; Kohlbacher, O.; P, L. H. *Phys. Rev. Lett.* **2004**, *93*, 108104.