The Importance of Solute−Solvent van der Waals Interactions with Interior Atoms of Biopolymers

Jed W. Pitera and Wilfred F. van Gunsteren*

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, Zürich, Switzerland

Received November 1, 2000
Revised Manuscript Received January 16, 2001

The balance between solute−solvent and solute−solute interactions defines the stability of proteins[1] and biomolecular complexes.[2] The large size of a protein means there are many interior solute atoms that do not contact the molecular surface. While these atoms are not in direct contact with solvent, they still make favorable van der Waals interactions with solvent atoms. To evaluate the significance of this attraction in macromolecular hydration we have carried out a number of thermodynamic integration (TI) free energy calculations in explicit solvent to compare the solvation free energies of "full" solutes with their "hollow" analogues (Figure 1). A calculation on the protein ubiquitin shows that the favorable solvation of interior atoms contributes −189 kJ/mol to the solvation free energy of a particular conformation. We went on to compare the solvation free energies of "full" model protein-like spheres and their "hollow" analogues (Figure 1A) or dimers of these species (Figure 1B). The pairs of solutes have identical solvent-accessible surface areas (SASA), but differ in solvation free energy by 49 to 282 kJ/mol. Furthermore, the dimer calculations show that the neglect of buried atom solvation can yield systematic relative binding free energy errors of 13 to 21 kJ/mol for small systems. The magnitude of this effect suggests that continuum solvation models based on a simple surface area (SA) dependence may be missing an important contribution to biopolymer solvation.

The explicit simulation of water accounts for the majority of the work involved in a modern protein simulation. As a result, much effort has been devoted to the development of continuum models of hydration that replace an explicit sampling over the solvent degrees of freedom with an analytical relation.[3] Typically, these models[4−7] separate the process of solvating a particular molecular conformation into a number of discrete steps that gradually transfer the solute from the gas phase into solution (Figure 2).[8] Many commonly used continuum models make use of a single linear SASA or molecular-SA dependent term to account for the free energy cost of inserting an uncharged van der Waals solute in solvent:

$$\Delta G_{\mathrm{solv}} = \Delta G_{\mathrm{eel}} + \gamma \mathrm{SA} \qquad (1)$$

where the constant of proportionality, γ, typically ranges from

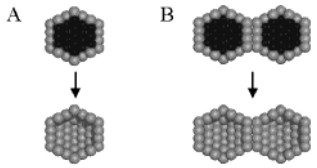

Figure 1. Model solutes compared in this work. First, the solvation free energy difference between ideal sphere-like solutes (top) and their "hollow" analogues (bottom) was calculated (A). A similar calculation compared dimers of ideal solutes (top) and their "hollow" analogues (bottom) (B). In both panels the interior atoms that disappear during the calculation are dark black and the rest gray. Solutes have been cut in half to display both interior and exterior atoms.

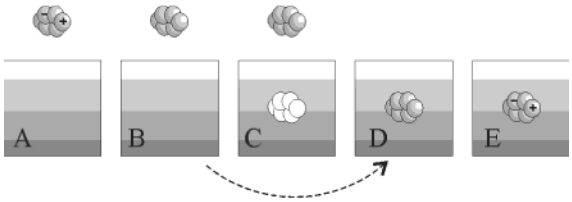

Figure 2. Typical stepwise processes used in continuum models of solvation. Starting with the solute in the gas phase and a neat volume of water (A), the free energy of discharging the solute is then calculated in the gas phase (B). Next, the work to create a solute-like hard-sphere cavity in the water is calculated (C). The uncharged (van der Waals) solute is then transferred from the gas phase into this cavity (D) and its charges are restored (E). C (cavitation) and D (van der Waals insertion) are often accounted for with the same linear surface area dependence (dashed arrow).

7.2[9] to 96.6[10] cal mol^{-1} Å^{-2}. This relation is based on the linear increase in hydration free energy of *n*-alkanes with chain length.[11] Unfortunately, there are no data available for macromolecular hydration free energies that would permit the parametrization of γ values for large solutes. Despite this, it has become popular to apply such models to calculate free energy differences associated with large conformational changes in biopolymers,[12−17] for example, the free energy difference between folded and unfolded states of a protein.[18] Continuum methods are also being used with increasing frequency in biomolecular simulations.[19−21] However, simulations of small alkanes have suggested that the single SA term represents the free energy cost of forming a hard-sphere cavity in water and may not capture the favorable van der Waals attraction between interior atoms of the solute and the solvent.[22,23]

* Corresponding author current address: Physikalische Chemie, ETH Zürich, ETH Zentrum, CH-8092 Zürich, Switzerland. Phone: +41-1-632-5501. Fax: +41-1-632-1039. E-mail: wfvgn@igc.phys.chem.ethz.ch.

(1) Dill, K. A.; Bromberg, S.; Yue, K. Z.; Fiebig, K. M.; Yee, D. P.; Thomas, P. M.; Chan, H. S. *Protein Sci*. **1995**, *4*, 561−602.
(2) Fersht, A. *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding;* W. H. Freeman and Co.: New York, 1998.
(3) Leach, A. *Molecular Modeling: principles and applications*; Addison-Wesley Longman Ltd.: Essex, 1996.
(4) Eisenberg, D.; McLachlan, D. *Nature* **1986**, *319*, 199−203.
(5) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc*. **1990**, *112*, 6127−6129.
(6) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem*. **1994**, *98*, 1978−1988.
(7) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161−2000.
(8) Gilson, M. K.; Honig, B. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 7−18.
(9) Hasel, W.; Henderson, T. F.; Still, W. C. *Tetrahedron Comput. Methodol.* **1988**, *1*, 103−116.
(10) Postma, J. P. M.; Berendsen, H. J. C.; Haak, J. R. *Faraday Symp. Chem. Soc.* **1982**, *17*, 55−67.
(11) Herman, R. B. *J. Phys. Chem*. **1972**, *76*, 2754−2759.
(12) Novotny, J.; Bruccoleri, R. E.; Saul, F. A. *Biochemistry* **1989**, *28*, 4735−4749.
(13) Froloff, N.; Windemuth, A.; Honig, B. *Protein Sci*. **1999**, *6*, 1293−1301.
(14) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc*. **1998**, *120*, 9401−9409.
(15) Rapp, C. S.; Friesner, R. A. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 173−183.
(16) Hünenberger, P. H.; Helms, V.; Narayana, N.; Taylor, S. S.; McCammon, J. A. *Biochemistry* **1999**, *38*, 2358−2366.
(17) Reyes, C. M.; Kollman, P. A. *J. Mol. Biol*. **2000**, *297*, 1145−1158.
(18) Lee, M. R.; Duan, Y.; Kollman, P. A. *Proteins: Struct., Funct., Genet.* **2000**, *39*, 309−316.
(19) Williams, D. J.; Hall, K. B. *Biophys. J*. **1999**, *76*, 3192−3205.
(20) Dominy, B. N.; Brooks, C. L. *J. Phys. Chem. B* **1999**, *103*, 3765−3773.
(21) Tsui, V.; Case, D. A. *J. Am. Chem. Soc.* **2000**, *122*, 2489−2498.
(22) Ashbaugh, H.; Kaler, E.; Pauliatis, M. *J. Am. Chem. Soc*. **1999**, *121*, 9243−9244.
(23) Gallicchio, E.; Kubo, M. M.; Levy, R. M. *J. Phys. Chem. B* **2000**, *104*, 6271−6285.

10.1021/ja0057474 CCC: $20.00
Published on Web 03/09/2001

Since large solutes have many buried atoms while *n*-alkanes have none, it is important to quantitate the influence of buried atoms on hydration free energies before using small molecule-derived γ values for macromolecules. The issue can be avoided by combining a SA-based cavitation term with an explicit modeling of solute−solvent van der Waals interactions, as in the explicit solvent/implicit solvent (ES/IS) model and others.[24,25] Multiple atom-dependent γ values[4,7] may also help, but again raise the question of parametrization.

The GROMOS biomolecular simulation package[26] was used for all simulations reported in this paper along with the GROMOS96 molecular mechanics force field.[27] Details of the system setup and simulations performed are included as Supporting Information. The first calculation was carried out on the X-ray crystal conformation of ubiquitin (1ubq[28]). A GROMOS96 molecular mechanics model of the protein was built with the correct van der Waals parameters but all partial charges set to zero. The free energy difference between this single uncharged conformation of ubiquitin and its "hollow" analogue with noninteracting interior atoms was calculated with TI. Interior atoms were defined as atoms that do not contact the molecular surface.

In addition to ubiquitin, we studied a series of model protein-like solutes. Coordinates for single sphere-like model solutes ("methylene crystals") were constructed by placing atoms on a bcc array with a spacing (0.25 nm) appropriate for a protein-like density (0.72 g/cm^3).[29] Solute atoms were modeled as uncharged methylene groups (CH_2). Again, our TI calculations determined the relative free energies of the solid ("full") solute and its corresponding "hollow" counterpart by calculating the work of converting all interior atoms of the solute to completely noninteracting dummy particles. Figure 1A depicts the two end-states of the calculation. We also calculated the free energy difference between "full" and "hollow" states of simple dimers of the 0.9 and 1.2 nm radius model solutes. In this case the interior atoms for the free energy calculation were defined to be the same as those for the model solute monomers (Figure 1B).

The results of our TI calculations are presented in Table 1. Our data for ubiquitin ($\Delta\Delta G_{solv}$(full→hollow) = +189 kJ/mol) show that buried atoms make a favorable contribution to the solvation free energy, a result that is verified by the model solute calculations. This significant effect cannot be captured with a simple surface area dependence, as SASAs for the "full" and "hollow" solutes are identical (ΔSASA < 0.01 nm^2). The solvation free energy difference between "full" and "hollow" species is linearly correlated with both the SASA ($R^2 = 0.93$) and solvent excluded volume ($R^2 = 0.99$).

Table 1. System Parameters and Relative Solvation Free Energies

solute	radius (nm)	no. of atoms (total/interior)	SASA (nm^2)	solvent-excluded vol (nm^3)	ΔG_{solv} (full→hollow) (kJ/mol)
1ubq		761/266	37.157	10.3421	188.6
sphere	0.6	99/33	8.848	2.0678	48.7
sphere	0.9	175/71	11.877	3.4182	77.8
sphere	1.2	457/239	22.317	8.5477	152.0
dimer	0.9/0.9	350/142	22.115	6.8166	142.0
dimer	1.2/1.2	914/478	41.826	17.0567	282.3

Our dimer calculations verify that interior atoms can have a significant effect on relative binding free energies which is not reproduced by a simple SA based model. The relative binding free energies of "full" and "hollow" solutes can be calculated by subtracting twice the relative solvation free energies of "full" and "hollow" monomers from the relative solvation free energy of the "full" and "hollow" dimers. The result is the relative binding free energy of "full" versus "hollow" species, $\Delta\Delta G_{bind}$(full→hollow). From our TI calculations, $\Delta\Delta G_{bind}$(full→hollow) is +13.6 kJ/mol for the 0.9/0.9 nm dimer and +21.7 kJ/mol for the 1.2/1.2 nm dimer. Since there is no SA difference between "full" and "hollow" solutes, a SA based continuum model instead yields 0 kJ/mol for either species. Clearly, the SA model overestimates the relative stability of the "hollow" dimers. While we have chosen an extreme comparison to illustrate our point, it is worthwhile to note that our model dimerization involves burial of fewer atoms than typical protein−protein dimerization events.[30]

As in the ES/IS model,[24] the solute−solvent van der Waals energies collected during an explicit solvent simulation could be used to calculate the van der Waals contribution to solvation. Contrary to that model, however, we find that the ratio of free energy to solute−solvent van der Waals energy is not 1 but roughly 0.8 ± 0.1. Single point continuum solvation calculations cannot be corrected in this manner. Instead, one might consider adding a term to the standard SA based function. A combined function of the solute SA or volume and density may be required to accurately model the contribution of buried atoms. Alternatively, different γ values could be used for solutes of different densities. Based on our data, γ should be 14 cal mol^{-1} Å^{-2} greater for the "hollow" solutes than for the "full" species.

These simulations show that the favorable van der Waals dispersion between interior atoms of the solute and the solvent, insignificant for small molecules, is an important effect in the solvation of large solutes such as biopolymers. Its contribution is significant (−49 to −282 kJ/mol) and may not be captured in a simple surface-area dependent term. As a result, one must be careful with standard "continuum electrostatics plus simple surface area dependence" models of solvation when they are applied to large solutes such as proteins or other biopolymers, particularly when comparing macromolecular conformations where the number of buried and solvent-exposed atoms differ significantly.

Supporting Information Available: The details of the calculations (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0057474

(24) Vorbojev, Y. N.; Almagro, J. C.; Hermans, J. *Proteins: Struct., Funct., Genet.* **1998**, *32*, 399−413.
(25) Carlson, H. A.; Jorgensen, W. L. *J. Phys. Chem.* **1995**, *99*, 10667−10673.
(26) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.
(27) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag AG an der ETH Zürich: Zürich, 1996.
(28) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(29) Creighton, T. E. *Proteins*; Freeman: New York, 1984.
(30) Jones, S.; Thornton, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13−20.