Comparison of volume and surface area nonpolar solvation free energy terms for implicit solvent simulations

Michael S. Lee, and Mark A. Olson

Citation: The Journal of Chemical Physics **139**, 044119 (2013);
View online: https://doi.org/10.1063/1.4816641
View Table of Contents: http://aip.scitation.org/toc/jcp/139/4
Published by the American Institute of Physics

Comparison of volume and surface area nonpolar solvation free energy terms for implicit solvent simulations

Michael S. Lee[1,2,a] and Mark A. Olson[2]
[1]*Computational Sciences Division, U.S. Army Research Laboratory, Aberdeen, Maryland 21005, USA*
[2]*Department of Cell Biology and Biochemistry, U.S. Army Medical Research Institute of Infectious Diseases, 1425 Porter St., Frederick, Maryland 21702, USA*

(Received 22 April 2013; accepted 11 July 2013; published online 30 July 2013;
publisher error corrected 1 August 2013)

Implicit solvent models for molecular dynamics simulations are often composed of polar and nonpolar terms. Typically, the nonpolar solvation free energy is approximated by the solvent-accessible-surface area times a constant factor. More sophisticated approaches incorporate an estimate of the attractive dispersion forces of the solvent and/or a solvent-accessible volume cavitation term. In this work, we confirm that a single volume-based nonpolar term most closely fits the dispersion and cavitation forces obtained from benchmark explicit solvent simulations of fixed protein conformations. Next, we incorporated the volume term into molecular dynamics simulations and find the term is not universally suitable for folding up small proteins. We surmise that while mean-field cavitation terms such as volume and SASA often tilt the energy landscape towards native-like folds, they also may sporadically introduce bottlenecks into the folding pathway that hinder the progression towards the native state. [http://dx.doi.org/10.1063/1.4816641]

INTRODUCTION

Implicit solvent models offer an efficient alternative to explicit solvent models for molecular dynamics applications, such as predicting protein folding thermodynamics, protein structure refinement, and protein-ligand docking.[1] Compared to explicit solvent, implicit solvent simulations have fewer degrees of freedom and, thus, can be more computationally efficient for sampling. In addition, in implicit solvent, the solute diffuses more quickly through various conformational states because friction from solvent molecules is absent. Furthermore, implicit solvent models respond instantaneously to changes in solute conformation which may accelerate sampling.

Most implicit solvent models for biomolecular simulations are composed of a polar/electrostatic term and a nonpolar term. The generalized Born (GB) method is often used to model the electrostatic solvation energy because it is an analytical approximation to the Poisson dielectric continuum method.[1] The nonpolar solvation contribution, on the other hand, is often modeled as a simple term proportional to the Lee-Richards solvent-accessible surface area (SASA).[2] Levy and co-workers introduced another nonpolar term which estimates the total attractive dispersion energy of water acting on solvent.[3,4] Wagoner and Baker suggested the addition of a third term proportional to the solute volume and fitted the three terms against nonpolar forces averaged over explicit solvent simulations.[5] While they found the best fit with all three terms, the volume term appears to be the dominant term. The somewhat unconventional conclusion that the volume term is important has been confirmed by others.[6] In this work, as others have done, we compute the solvation free energy forces associated with water molecules acting on fixed conformations of a solute. However, unlike previous studies, we test the compelling hypothesis that the molecular volume term should perform better compared to SASA for molecular dynamics simulations. Specifically, we test how well each term folds up small proteins in temperature-based replica exchange self-guided Langevin dynamics (ReX-SGLD) simulations. Our comparative analysis allows a relatively unbiased and simple test of energy function quality for native structure generation, detection, and thermodynamic observable estimation.

THEORY

Solvation free energy is conventionally divided into two terms: polar and nonpolar (*np*),

$$\Delta G_{solv} = \Delta G_{polar} + \Delta G_{np}. \quad (1)$$

In a predominant formulation, the work of solvating a solute[7] proceeds in two stages. First, the nonpolar free energy is the work of immersing an uncharged solute into a bath of pure solvent. Second, the polar free energy is the work of turning on the electrostatic charge of the solute in its new environment. For molecular dynamics (MD) simulations, one of the most popular approximations to the polar solvation term is the GB model, which is itself an analytical estimate of the Poisson dielectric continuum model. For nonpolar solvation, the most widely used approximation is a term proportional to the solvent accessible-surface area (SASA). More precisely, the nonpolar solvation can be represented as the sum of several contributions (e.g., cavitation free energy and interactions of nonpolar solute atoms with water). Investigating the nonpolar component in more detail, Wagoner *et al.*[5] proposed that the nonpolar solvation energy can be represented

[a]Author to whom correspondence should be addressed. Electronic mail: michael.s.lee131.civ@mail.mil

as a sum of three distinct components

$$\Delta G_{np} \approx pV + \gamma S + \Delta G_{attr}^{vdW}, \quad (2)$$

where V is the solute volume, S is the solute surface, and ΔG_{attr}^{vdW} is the van der Waals (vdW) attraction term. In the macroscopic limit, the values of p and γ for liquid water are experimentally known. Pure water has a bulk surface tension value[8] of $\gamma_{\text{bulk}} = 103.6 \pm 0.5$ cal/mol/Å^2 at 25°C. Under standard conditions, the pressure, $p = 1$ atm, which in microscopic terms is very small; namely, $p = 0.0146$ cal/mol/Å^3, and can be arguably omitted. However, this paper and others[5,6] have fit this parameter to explicit solvent energies or forces and found that it can be quite large, at least in the microscopic regime. The term, ΔG_{attr}^{vdW}, was first introduced by Levy and co-workers[3,4] to model the dispersion interaction between the solvent and solute. In principle, this term can be obtained by integrating the vdW term between the solute and a constant density of solvent over the infinite solvent-accessible volume. In practice, this term can be efficiently estimated using the Born radii, B_i, that are already computed for the GB polar term.[9] We do not consider this term further in this work because we found its contribution to yield negligible improvements in correlation (<0.1%) to the parametric fits of explicit solvent force data. Therefore, even though the volume and SASA term are traditionally considered cavitation terms, they often incorporate attractive dispersion effects, as in this work.

Exact analytical computation of the SASA term is nontrivial and is described elsewhere.[10] Here, we devise a new, but relatively simple definition of solvent-accessible volume, V, of a solute that is continuous up to 2nd derivatives:

$$V = \iiint D(\mathbf{y})\, dx\, dy\, dz \approx \sum_{g_x, g_y, g_z} D(\mathbf{y}_{g_x, g_y, g_z})\, \Delta^3, \quad (3)$$

where g_x, g_y, and g_z, are cubic lattice indices, Δ is the lattice spacing,

$$\mathbf{y}_{g_x, g_y, g_z} = (x_{\min} + g_x \Delta)\vec{i} + (y_{\min} + g_y \Delta)\vec{j} + (z_{\min} + g_z \Delta)\vec{k}, \quad (4)$$

and (x_{min}, y_{min}, z_{min}) is the lower corner of the simulation space. The solvent-accessible volume function, D, is the union of atom-centered radial functions positioned at their respective atomic coordinates, $\mathbf{x}_i$,

$$D(\mathbf{y}) = 1 - \prod_i [1 - f(\|\mathbf{y} - \mathbf{x}_i\|)], \quad (5)$$

where the radial function, $f(r)$, is defined as a piecewise polynomial spline smooth up to second derivatives,

$$f(r) = \begin{cases} 1 & r < r_{start} + r_{vdW} \\ 1 + t^3(t(15 - 6t) - 10) & r_{start} + r_{vdW} < r < r_{stop} + r_{vdW}, \\ 0 & r > r_{stop} + r_{vdW} \end{cases} \quad (6)$$

$$t = \frac{r^2 - (r_{start} + r_{vdW})^2}{(r_{stop} + r_{vdW})^2 - (r_{start} + r_{vdW})^2} \quad (7)$$

and r_{start} and r_{stop} are fit to explicit solvent data but are roughly defined by the vdW radius, r_{vdW}, plus a probe radius. Note the squared distances in Eq. (7) mildly skew the radial function for different vdW radii, r_{vdw}, but remove the need for computationally intensive square roots. Computation of the volume function (and its atomic derivatives) is naïvely order $O(N^2)$, where N is the number of atoms, but can be reduced to $O(N)$ by using the same lookup grid already employed by GBMV2[11] to compute the Born radii.

While calculating the absolute nonpolar free energy of a macromolecule via explicit solvent simulation would be very computationally demanding, it is relatively trivial to compute the mean forces and hence atomic forces due to the nonpolar free energy. In a simulation with explicitly-modeled solvent molecules, the nonpolar solvation free energy, G_{np}, can be written as[7]

$$e^{-\beta G_{np}(\mathbf{x})} = \frac{\iint e^{-\beta U_{solv-solv}(\mathbf{z})} e^{-\beta U_{solv-solute}(\mathbf{x},\mathbf{z})} d\mathbf{x} d\mathbf{z}}{\int e^{-\beta U_{solv-solv}(\mathbf{z})} d\mathbf{z}}, \quad (8)$$

where $\beta = 1/k_B T$, $\mathbf{x}$ and $\mathbf{z}$ are the solute and solvent degrees of freedom, respectively, $U_{solv\text{-}solv}$ is the solvent-solvent interaction energy, $U_{solv\text{-}solute}$ is the solvent-solute interaction energy, and k_B is Boltzmann's constant. Taking derivatives of both sides of Eq. (8) with respect to the solute coordinates, $\mathbf{x}$, shows that the nonpolar free energy force, $\frac{\partial G_{np}}{\partial \mathbf{x}}$, can be computed by ensemble-averaging the nonpolar potential force,[5]

$$\frac{\partial G_{np}}{\partial \mathbf{x}} = \left\langle \frac{\partial U_{solv-solute}}{\partial \mathbf{x}} \right\rangle_{q_{solute} = 0}. \quad (9)$$

In practice, nonpolar potential forces are calculated by setting the solute charges to zero during the simulation which, in effect, removes electrostatic interactions between the solute and solvent. We investigated the effect of trimming the tail of the distribution (i.e., removing the upper and lower segments of the force distribution for each Cartesian component of each atom) on the computed mean forces and the corresponding parameter fit in the Results and Discussion section.

METHODS

For each of 14 fixed protein conformations of the Trp-cage protein, we ran 4-ns NAMD simulations to compute the average forces on a solute due to explicit solvent (Eq. (9)).

The conformations were obtained from a previous study of implicit solvent ReX-SGLD simulations.[12] Nonbonded interactions were switched off to zero from 7 to 10 Å. Electrostatic interactions were computed by particle mesh Ewald. With an integration time step of 2 fs, the SETTLE algorithm was used to keep water bonds and angles rigid. Langevin dynamics was used to maintain constant temperature (300 K) and pressure (1 atm) with a piston oscillation time of 100 fs, decay time of 50 fs, and isotropic (cubic) cell fluctuation. We used CHARMM version c35b3[13] to evaluate the nonpolar solvation forces acting on the solute from the NAMD-generated trajectories using the same potential energy function.

The SASA term was computed using the analytical formulation in the *aspenr* module of CHARMM. The solvent-access volume energy and derivatives were implemented in CHARMM. The volume term with r_{start}, $r_{stop} = 0$ was validated against the *volume* function in the *coor* module. The grid resolution, Δ (from Eq. (3)), was set to 0.5 Å, as smaller values yielded diminishing gains in accuracy and higher computational cost.

The SASA and volume-based nonpolar solvation terms were applied to ReX implicit solvent folding simulations of chignolin (PDB ID: 1UAO),[14] tryptophan cage (Trp-cage, PDB ID: 1L2Y),[15] and tryptophan zipper (Trp-zip, PDB ID: 1LE0).[16]

For all of the folding simulations, windows were assigned temperatures by geometric spacing from 270 to 500 K. Chignolin simulations had eight windows and were run for 40 ns; Trp-cage and Trp-zip simulations had 16 windows and were run for 100 ns. The polar component of the solvation was modeled with GBMV2[17] using a smoothing parameter of $\beta = -12$.[18] Nonbonded interactions were switched off smoothly from 20 to 22 Å. The self-guided Langevin dynamics method[19] was used to enhance conformational sampling by up to a factor of two compared to regular Langevin dynamics or molecular dynamics.[20] The SGLD parameters were set as follows:[12] friction constant, $\beta = 1$ ps^{-1}; self-guided strength, $\lambda = 1$, and self-guided averaging time, $\tau_{\text{avg}} = 1$ ps.

Two common uses for implicit solvation are protein structure prediction/refinement and calculation of folding thermodynamics. To assess structure prediction suitability, we evaluated three metrics: (1) C_α root-mean-squared-deviation (RMSD) to native of lowest energy structure, (2) average C_α RMSD of the 10 lowest energy structures, and (3) lowest C_α RMSD in the set. We also plotted energy vs. C_α RMSD for several sets of structures.

Regarding folding thermodynamics, we investigated whether a given set of parameters folded the protein to the native basin and, conversely, whether the protein was too stable. We evaluated the melting temperature, T_m, of the simulated protein using two methods: (1) temperature at which heat capacity is a maximum and (2) temperature at which the folding free energy, ΔG_{fold}, equals 0. To calculate both quantities, we use the weighted histogram analysis method.[21] Heat capacity, C_v, is calculated as

$$C_v = \frac{\langle U^2 \rangle - \langle U \rangle^2}{k_B T^2}, \tag{10}$$

where U is the potential energy. Free energy of folding, $\Delta\Delta G_{fold}$, is calculated as

$$\Delta\Delta G_{fold}(T) = -k_B T \ln \frac{\rho_{fold}(T)}{\rho_{unfold}(T)} = -k_B T \ln \frac{\int_{RMSD<RMSD_{fold}} \rho(T) dV}{\int_{RMSD>RMSD_{fold}} \rho(T) dV}, \tag{11}$$

where $RMSD_{fold}$ is an arbitrary dividing line between folded (native-like) and unfolded states. In this work, it was set to 3 Å.

RESULTS AND DISCUSSION

Volume and surface area terms were fit to ensemble-averaged vdW forces due to explicit solvent molecules acting on the solute in Table I and Figure 1. Notably, the volume term provides a better fit to the explicit solvent data compared to the surface area term. Given that the fitted surface area and volume coefficients (p and γ, respectively) did not perform as expected in molecular dynamics simulations (discussed below), we experimented with smaller coefficients and justified their usage by analyzing the distribution of instantaneous vdW forces for arbitrary atomic components. We found some distributions to be symmetric, others to be one-sided as seen in Fig. 2. Therefore, we sorted instantaneous forces for each atom and its Cartesian components and then employed filtering of various fractions at each tail to obtain filtered means. Parameters were separately fitted to the unfiltered and filtered mean as seen in Table I. Trimming the distribution at both ends reduces the coefficients down to a lower bound value (median) of $p = 17$ cal/mol/Å^3 and $\gamma = 24$ cal/mol/Å^2. We determined a reasonable compromise was a 15% reduction of each tail which led to $p = 30$ cal/mol/Å^3 and $\gamma = 40$ cal/mol/Å^2. The marked difference between the mean and median can be attributed to the non-symmetric nature of the many of the component force distributions (Figure 2). Interestingly, the unfiltered values of p are in the ballpark of values ($p = 55 \pm 2$ cal/mol/Å^3) obtained by the study of Wagoner *et al.*[5,22] where the explicit solvent nonpolar forces were evaluated for a 130-residue protein test system. This similarity suggests that the p and γ parameters are solute size-independent for small macromolecules.

TABLE I. Nonpolar solvation single parameter fits to unfiltered and filtered mean.

Term	No filter	Tail filter (15%)	Median
Surface tension, γ (cal/mol/Å^2)	70	40	24
Probe radius, r_{probe} (Å)	0.3	0.7	0.8
Correlation coefficient	0.72	0.65	0.52
Pressure, p (cal/mol/Å^3)	60	30	15
r_{start} (Å)	0.6	0.6	1.0
r_{stop} (Å)	1.2	1.2	1.5
Correlation coefficient	0.91	0.80	0.61

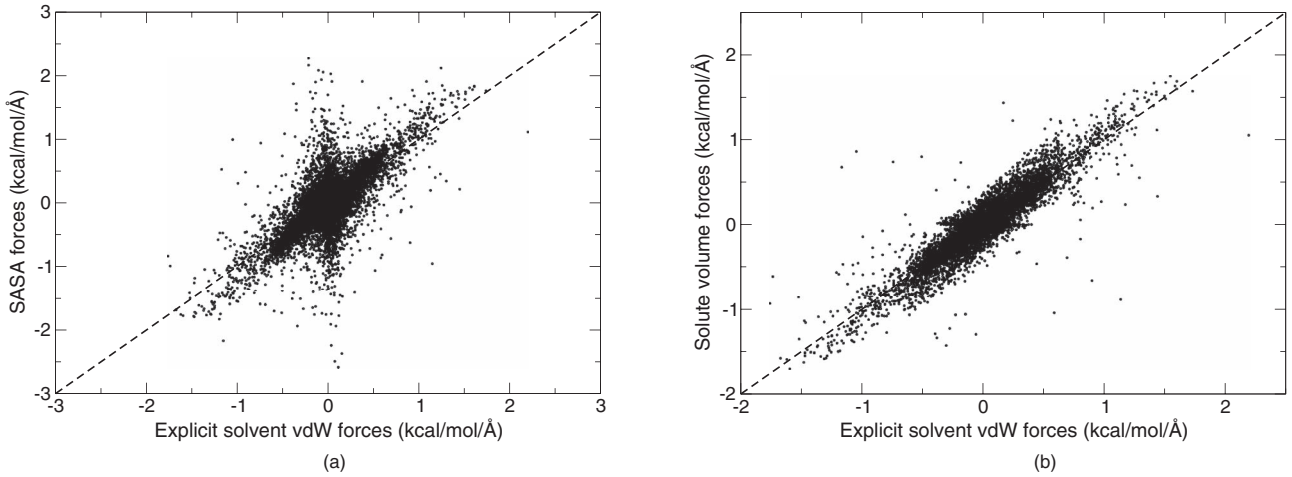

FIG. 1. Comparison between mean-field cavitation forces from explicit solvent and nonpolar terms, (a) SASA, $r = 0.72$ and (b) solute volume, $r = 0.91$.

Our first protein folding test case is the 10-residue protein, chignolin.[14] We and others[23–27] have been able to fold this protein with explicit solvent replica exchange simulations as see in Table II and Fig. 3(d). Successful folding with implicit solvent has been limited.[28] Comparing various surface-area and volume terms combine with GBMV polar solvation (Table II), we find that only the volume-based nonpolar solvation correctly folds up the protein, albeit just barely (Figure 3(c)). The larger coefficient ($p = 60$) from the unfiltered mean fit seems to "crush" the folded structure leading to an exceptionally large predicted melting temperature. For reference, the experimental melting temperature of chignolin is 312 K.[29]

The SASA-based nonpolar term failed to fold chignolin, although the $\gamma = 40$ cal/mol/Å^2 generated more near-native structures compared to $\gamma = 5.42$ cal/mol/Å^2 as seen in Figs. 3(a) and 3(b). In Figure 4(a), rescoring some of the decoys generated by the explicit solvent simulation with SASA ($\gamma = 40$ cal/mol/ Å^2) plus GBMV energy shows that this energy function can easily detect the native fold. Another lab has shown that the more approximate GBSW method[28] folds chignolin.[30] While the GBSW model of polar solvation is faster than GBMV, it is a more approximate alternative to Poisson solvation[1] and not always suitable for protein structure prediction for other proteins.[31,32]

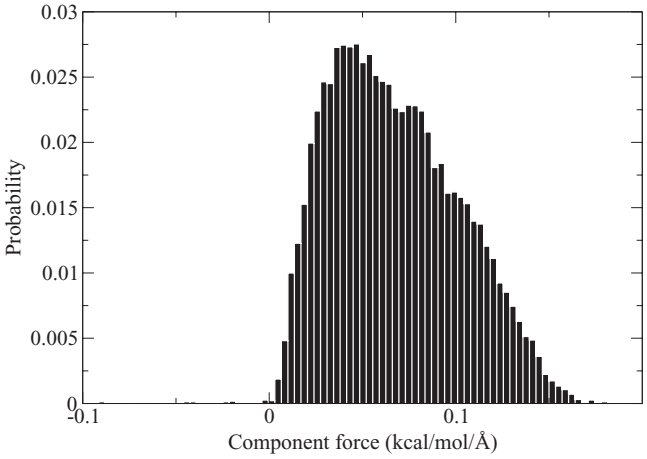

FIG. 2. Nonpolar solvent force distribution along a single axis for an atom on the surface of a Trp-cage protein decoy.

For the other two test systems, Trp-cage and Trp-zip, the results in Table III are mixed. Trp-cage folding with the volume term somewhat outperformed the surface area term. In a previous study, we folded the Trp-cage protein with a weak surface tension parameter, $\gamma = 5.42$ cal/mol/Å. In this work, with a stronger parameter, $\gamma = 40$ cal/mol/Å, we also folded the protein but the accuracy of the native basin was diminished by about 1 Å C_α RMSD compared to that study. The volume-based nonpolar term produced moderately better results. Intriguingly, SASA and volume energy terms yield diametrically opposite melting temperatures despite being fit to the same filtered explicit solvent data.

Compared to the other two proteins, the Trp-zip results were surprising (Table III). The potential energy with the SASA nonpolar term folded the peptide into the native basin after only 3 ns (results not shown). The volume term simulation with $p = 30$ cal/mol/Å^2, on the other hand, failed to find the native basin after 100 ns. Rescoring the SASA simulations in Fig. 4(b) reveals that the volume term correctly detects the native basin.

If the volume term does in fact favor the native basin, then why is the native basin not sampled? The same question can be asked to why the SASA term detects the native chignolin basin but does not reach it. We think that the cavitation term nonspecifically raises the energy of the transition state between compact basins thereby introducing kinetic traps within these compact basins. The principle is that with a mean-field cavitation approximation, instantaneous expansions of the solute are disfavored thereby reducing the probability of structural rearrangements into new compact forms. This effect could explain why it is important to keep the p and γ coefficients small regardless of the implied fit to explicit solvent force data. In addition, we see that the misfolded basins generated by current molecular dynamics force fields and implicit solvent models conflict with the experimental observation that many small proteins exhibit two-state folding. If there were no artificial misfolded decoy states, then kinetic

TABLE II. Chignolin folding simulations starting from the *trans* conformation: model accuracies (RMSD, Å) and melting temperatures, T_{m} (K). Experimental T_m is 312 K.[29]

Nonpolar term	Parameters	Lowest energy structure	Avg. top 10	Best RMSD @ 270 K	T_m(K) (max C_p)	T_m(K) ($\Delta G_{fold} = 0$)
None	n/a	4.7	6.3	2.5	<270	n/a
SASA	$\gamma = 5.42$, $r_{\mathrm{probe}} = 1.4$ Å	6.0	6.1	2.3	n/a	n/a
SASA	$\gamma = 40$, $r_{\mathrm{probe}} = 0.7$ Å	2.8	3.3	0.6	369	379
SASA	$\gamma = 70$, $r_{\mathrm{probe}} = 0.3$ Å	3.9	3.4	0.8	464	>500
Volume	$p = 30$	3.8	1.9	0.6	383	398
Volume	$p = 60$	1.0	1.8	0.6	479	480
TIP3P		1.3	1.3	0.4	304	312

TABLE III. Metrics for folding the Trp-cage (T_{m} [exp] = 315 K[34]) and Trp-zip (T_{m} [exp] = 323 K[16]) peptides. SASA parameters: $\gamma = 40$ cal/mol/Å^2, $r_{\mathrm{probe}} = 0.7$ Å; Volume parameter: $p = 30$ cal/mol/Å^3.

Peptide	Nonpolar term	Lowest energy (Å)	Average top 10 (Å)	Best RMSD (Å)	T_m(K) (max C_p)	T_m(K) ($\Delta G_{fold} = 0$)
Trp-cage	SASA	1.8	1.9	1.3	<270	304
Trp-cage	Volume	1.5	1.8	0.6	>500	533
Trp-zip	SASA	0.8	1.8	0.3	<270	316
Trp-zip	Volume	6.1	5.6	3.0	n/a	n/a

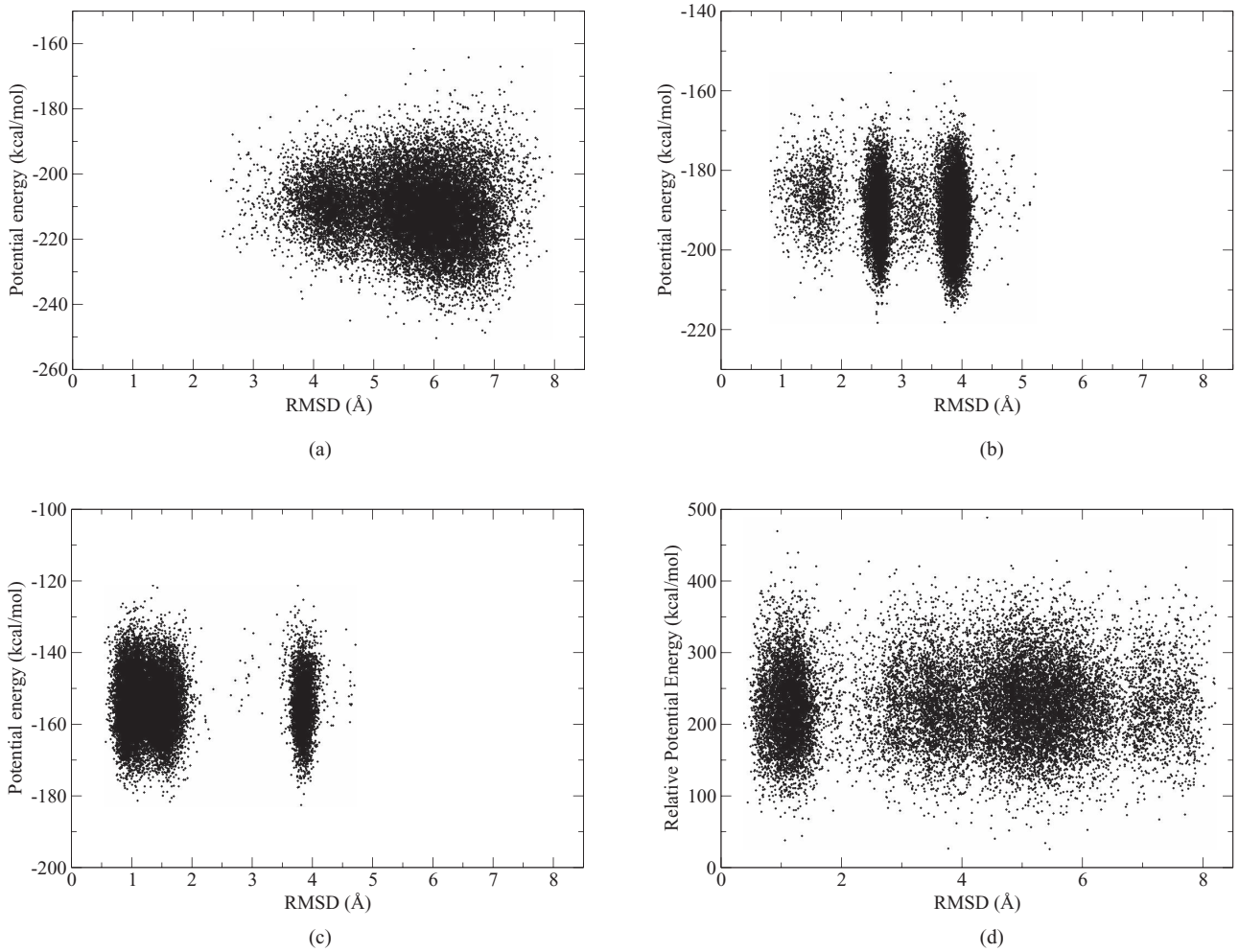

FIG. 3. Sampling of chignolin structures using different simulation protocols: (a) GBMV + SASA, $\gamma = 5.42$ cal/mol/Å^2, $r_{probe} = 1.4$ Å; (b) GBMV + SASA, $\gamma = 40$ cal/mol/Å^2, $r_{probe} = 0.7$ Å; (c) GBMV + volume, $p = 30$ cal/mol/Å^3; (d) TIP3P explicit solvent (−13 400 kcal/mol subtracted from the energies).

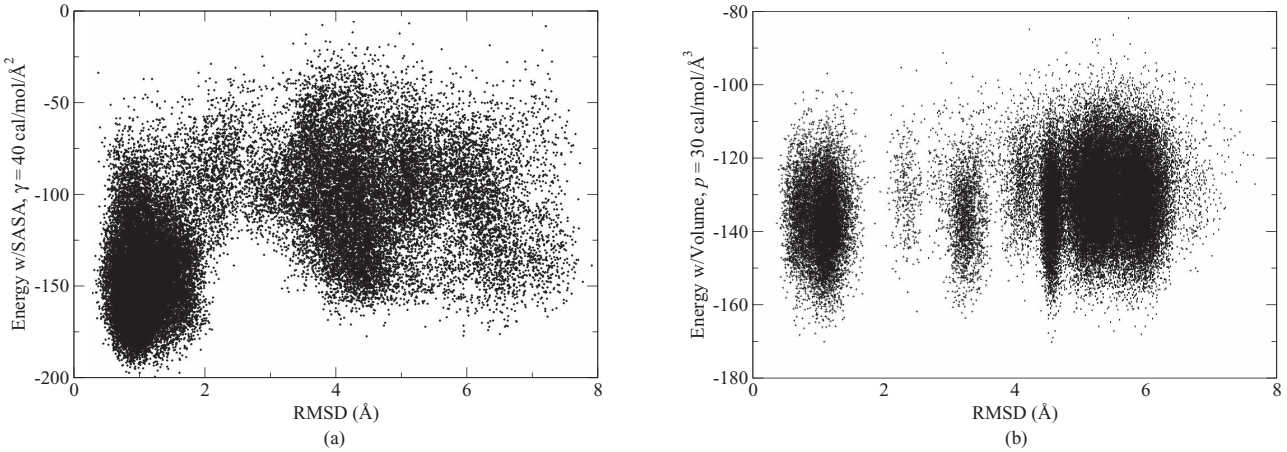

FIG. 4. Rescoring with other cavitation term. (a) SASA ($\gamma = 40$ cal/mol/Å^2) rescoring of two native basin-visiting windows of the explicit solvent chignolin run. (b) Volume term ($p = 30$ cal/mol/Å^3) rescoring of ∼15 000 random structures of the Trp-zip ($\gamma = 40$ cal/mol/Å^2) run ($T = 270$ K window).

traps would probably not be an issue. It is also interesting to point out that polar implicit solvent models (e.g., GB) tend to accelerate sampling because they generally dampen the electrostatic interactions within the solute while nonpolar cavitation terms, as seen in this work, decelerate sampling.

CONCLUSIONS

The solvent-accessible surface area is often used to model nonpolar solvation for implicit solvent calculations. While researchers have long understood its limitations, optimization of the surface tension and probe radius has only rarely been considered.[33] Volume-based nonpolar solvation has not been previously applied to simulations despite its relatively simple implementation. We show that neither the SASA nor the volume term provide a universal "fix" to implicit solvent modeling of protein folding for native structure generation. A delicate balance needs to be achieved between favoring the native basin and not significantly decelerating the sampling of large-scale structural rearrangements.

ACKNOWLEDGMENTS

M.A.O. acknowledges funding from Defense Threat Reduction Agency Grant (Grant Nos. TMTI0004_09_BH_T and CBCALL12–LS6–2–0036). Computational time was provided by the U.S. Army Research Laboratory DoD Supercomputing Resource Center.

[1]M. Feig, A. Onufriev, M. S. Lee, W. Im, D. A. Case, and C. L. Brooks III, J. Comput. Chem. **25**, 265 (2004).
[2]J. Chen and C. L. Brooks III, Phys. Chem. Chem. Phys. **10**, 471 (2008).
[3]E. Gallicchio, K. Paris, and R. M. Levy, J. Chem. Theory Comput. **5**, 2544 (2009).
[4]E. Gallicchio and R. M. Levy, J. Comput. Chem. **25**, 479 (2004).
[5]J. A. Wagoner and N. A. Baker, Proc. Natl. Acad. Sci. U.S.A. **103**, 8331 (2006).
[6]C. Tan, Y. H. Tan, and R. Luo, J. Phys. Chem. B **111**, 12263 (2007).
[7]B. Roux and T. Simonson, Biophys. Chem. **78**, 1 (1999).
[8]N. B. Vargaftik, B. N. Volkov, and L. D. Voljak, J. Phys. Chem. Ref. Data **12**, 817 (1983).
[9]E. Gallicchio, K. Paris, and R. M. Levy, J. Chem. Theor. Comput. **5**, 2544 (2009).
[10]B. Lee and F. M. Richards, J. Mol. Biol. **55**, 379 (1971).
[11]M. S. Lee, F. R. Salsbury, and C. L. Brooks, J. Chem. Phys. **116**, 10606 (2002).
[12]M. S. Lee and M. A. Olson, J. Chem. Theor. Comput. **6**, 2477 (2010).
[13]B. R. Brooks, C. L. Brooks III, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, J. Comput. Chem. **30**, 1545 (2009).
[14]S. Honda, K. Yamasaki, Y. Sawada, and H. Morii, Structure **12**, 1507 (2004).
[15]J. W. Neidigh, R. M. Fesinmeyer, K. S. Prickett, and N. H. Andersen, Biochemistry **40**, 13188 (2001).
[16]A. G. Cochran, N. J. Skelton, and M. A. Starovasnik, Proc. Natl. Acad. Sci. U.S.A. **98**, 5578 (2001).
[17]M. S. Lee, M. Feig, F. R. Salsbury, and C. L. Brooks, J. Comput. Chem. **24**, 1348 (2003).
[18]J. Chocholousova and M. Feig, J. Comput. Chem. **27**, 719 (2006).
[19]X. W. Wu and B. R. Brooks, Chem. Phys. Lett. **381**, 512 (2003).
[20]M. A. Olson, S. Chaudhury, and M. S. Lee, J. Comput. Chem. **32**, 3014 (2011).
[21]S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, J. Comput. Chem. **16**, 1339 (1995).
[22]J. A. Wagoner and N. A. Baker, Proc. Natl. Acad. Sci. U.S.A. **103**, 8331 (2006); **104**, 1732 (2007).
[23]D. van der Spoel and M. M. Seibert, Phys. Rev. Lett. **96**, 238102 (2006).
[24]S. Kannan and M. Zacharias, Proteins **66**, 697 (2007).
[25]S. Roy, S. Goedecker, M. J. Field, and E. Penev, J. Phys. Chem. B **113**, 7315 (2009).
[26]R. Harada and A. Kitao, J. Phys. Chem. B **115**, 8806 (2011).
[27]S. Chaudhury, M. A. Olson, G. Tawa, A. Wallqvist, and M. S. Lee, J. Chem. Theor. Comput. **8**, 677 (2012).
[28]W. Im, M. S. Lee, and C. L. Brooks III, J. Comput. Chem. **24**, 1691 (2003).
[29]C. M. Adams and R. A. Zubarev, Anal. Chem. **77**, 4571 (2005).
[30]X. Li, R. A. Latour, and S. J. Stuart, J. Chem. Phys. **130**, 174106 (2009).
[31]I. C. Yeh, M. S. Lee, and M. A. Olson, J. Phys. Chem. B **112**, 15064 (2008).
[32]B. Strodel and D. J. Wales, J. Chem. Theor. Comput. **4**, 657 (2008).
[33]M. Feig and C. L. Brooks III, Proteins **49**, 232 (2002).
[34]J. W. Neidigh, R. M. Fesinmeyer, and N. H. Andersen, Nat. Struct. Biol. **9**, 425 (2002).