

Minimal Molecular Surfaces and Their Applications

P. W. BATES,¹ G. W. WEI,^{1,2} SHAN ZHAO³

¹Department of Mathematics, Michigan State University, Michigan 48824

²Department of Electrical and Computer Engineering, Michigan State University, Michigan 48824

³Department of Mathematics, University of Alabama, Alabama 35487

Received 15 November 2006; Revised 13 May 2007; Accepted 26 May 2007

DOI 10.1002/jcc.20796

Published online 25 June 2007 in Wiley InterScience (www.interscience.wiley.com).

Abstract: This article presents a novel concept, the minimal molecular surface (MMS), for the theoretical modeling of biomolecules. The MMS can be viewed as a result of the surface free energy minimization when an apolar molecule, such as protein, DNA or RNA is immersed in a polar solvent. Based on the theory of differential geometry, the MMS is created via the mean curvature minimization of molecular hypersurface functions. A detailed numerical algorithm is presented for the practical generation of MMSs. Extensive numerical experiments, including those with internal and open cavities, are carried out to demonstrate the proposed concept and algorithms. The proposed MMS is typically free of geometric singularities. Application of the MMS to the electrostatic analysis is considered for a set of twenty six proteins.

© 2007 Wiley Periodicals, Inc. J Comput Chem 29: 380–391, 2008

Key words: biomolecular surface; minimal surface; mean curvature flow; evolution equation; molecular surface

Introduction

Molecular models have widespread applications in modern science and technology. The atom and bond model of molecules was proposed by Corey and Pauling in 1953,¹ and continues to be a cornerstone in physical science. The regular polyhedral and periodic lattice model plays an important role in crystallography and solid state physics. The molecular and atomic orbital models provide a visual basis for a quantum mechanical description of molecules and their dynamics. The difficulty of modeling and visualization of large complex biomolecules has motivated the development of a variety of physical and graphical models. Among them, the molecular surface (MS)² is one of the most important models in molecular biology. The stability and solubility of macromolecules, such as proteins, DNAs and RNAs, are determined by how their surfaces interact with solvent and other surrounding molecules. Therefore, the structure and function of macromolecules depend on the features of their molecule-solvent interfaces.³ The MS is defined by rolling a probe sphere with a given radius around the set of atomic van der Waals spheres.^{2,4,5} It has been applied to protein folding,⁶ protein-protein interfaces,⁷ protein surface topography,³ oral drug absorption classification,⁸ DNA binding and bending,⁹ macromolecular docking,¹⁰ enzyme catalysis,¹¹ calculation of solvation energies,¹² and molecular dynamics.¹³ The concept of molecule and solvent interfaces is of paramount importance to the implicit solvent models^{14,15} and polarizable continuum methods.¹⁶ MSs are generated by a variety of methods, including use of the Gauss-Bonnet theorem,^{17,18} overlapped multiple spheres,¹⁹ space

transformation,²⁰ alpha shape theory,²¹ the contour-buildup algorithm,²² variable probe radius²³ and parallel methods.²⁴ While most methods represent the resulting MS by triangulation,^{25–27} a Cartesian grid based method was proposed by Rocchia et al.²⁸ A partial differential equation approach of MS was proposed by Wei et al.²⁹ However, the existing biomolecular surface models encounter theoretical and computational difficulties, due to the possible presence of self-intersecting surfaces, cusps, and other singularities.^{25,26,30,31} Moreover, such models are inconsistent with the surface free energy minimization, which likely leads to a minimal surface separating the apolar biomolecule from a polar solvent.

Because of the energy minimization principle, minimal surfaces are omnipresent in nature. Their study has been a fascinating topic for centuries.^{32–34} French geometer, Meusnier, constructed the first non-trivial example, the catenoid, a minimal surface that connects two parallel circles, in the 18th century. In the 1760s, Lagrange discovered the relation between minimal surfaces and a variational principle, which is still a cornerstone of modern mechanics. Plateau studied minimal surfaces in soap films in the mid-nineteenth century. In liquid phase, materials of largely different polarizabilities, such as water and oil, do not mix, and the material in smaller quantity forms ellipsoidal drops, whose surfaces are minimal subject to the gravitational constraint. The self-assembly of minimal cell membrane surfaces in water has

Correspondence to: G. W. Wei; e-mail: wei@math.msu.edu

Contract/grant sponsor: NSF grant; contract/grant number: DMS-0616704

been discussed.³⁵ Curvature effects in static cell membrane deformations have been considered by Du et al.³⁶ The Schwarz P minimal surface is known to play a role in periodic crystal structures.³⁷ The formation of β -sheet structures in proteins is regarded as the result of surface minimization on a catenoid.³⁸ A minimal surface metric has been proposed for the structural comparison of proteins.³⁹ However, to the best of our knowledge, a natural minimal surface that separates a less polar macromolecule from its polar environment such as the water solvent has not been considered yet.

The generation of minimal surfaces with given boundary constraints can be pursued using Matlab or Mathematica.⁴⁰ Evolution equation approaches were also proposed to generate minimal surfaces with predetermined boundaries.^{41,42} However, there is no algorithm available that generates minimal surfaces constrained by obstacles, such as arbitrarily distributed atoms in biomolecules, to the best of our knowledge.

The objectives of the present article are twofold, i.e., to propose a novel concept, the minimal molecular surface (MMS), for the modeling of biomolecules, and to develop a new algorithm for the practical generation of MMSs under biomolecular constraints. Since the surface free energy is proportional to the surface area, an MMS contributes to the molecular stability in solvent. Therefore, there must be an MMS associated with each stable macromolecule in its polar environment, and our differential geometric approach appears to produce the desired results. A brief report of the proposed concept and algorithm was presented elsewhere.^{43,44}

This paper is organized as follows. In Theoretical Modeling, we provide the theoretical modeling of the MMS. A hypersurface representation of biomolecular system is defined. The normal and mean curvature of the hypersurface is evaluated and used to evolve the hypersurface. Methods and algorithms are described in Methods and Algorithms. The hypersurface function is initialized based on atomic constraints, and evolved via the mean curvature minimization. A level surface is extracted from the steady state hypersurface function to obtain the MMS. Results and Discussion is devoted to numerical results and discussions. We validate the proposed method via several numerical experiments. The capability of representing open and internal cavities is demonstrated. We show that the proposed method is free of typical geometric singularities. Electrostatic analysis is carried out with the proposed MMS.

Theoretical Modeling

Hypersurface and its Mean Curvature

Consider a C^2 immersion $f: U \rightarrow \mathbb{R}^4$, where $U \subset \mathbb{R}^3$ is an open set. Here $f(u) = (f_1(u), f_2(u), f_3(u), f_4(u))$ is a hypersurface element (or a position vector), and $u = (u_1, u_2, u_3) \in U$.

Tangent vectors (or directional vectors) of f are $X_i = \frac{\partial f}{\partial u_i}$. The 4×3 Jacobian matrix of the mapping f is given by $Df = (X_1, X_2, X_3)$.

The first fundamental form is a symmetric, positive semi-definite metric tensor of f , given by $I: = (g_{ij}) = (Df)^T \cdot (Df)$. Its matrix elements can also be expressed as $g_{ij} = \langle X_i, X_j \rangle$, where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product in \mathbb{R}^4 , $i, j = 1, 2, 3$.

Let $v(u)$ be the unit normal vector given by the Gauss map $v: U \rightarrow S^3$,

$$v(u_1, u_2, u_3) := X_1 \times X_2 \times X_3 / \|X_1 \times X_2 \times X_3\| \in \perp_u f, \quad (1)$$

where the cross product in \mathbb{R}^4 is a generalization of that in \mathbb{R}^3 . Here, $\perp_u f$ is the normal space of f at point $p = f(u)$. The vector v is perpendicular to the tangent hyperplane $T_u f$ to the surface at p . Note that $T_u f \oplus \perp_u f = T_{f(u)} \mathbb{R}^4$, the tangent space at p . By means of the normal vector v and tangent vector X_i , the second fundamental form is given by

$$II(X_i, X_j) = (h_{ij}) = \left\langle -\frac{\partial v}{\partial u_i}, X_j \right\rangle. \quad (2)$$

The mean curvature can be calculated from

$$H = \frac{1}{3} h_{ij} g^{ij}, \quad (3)$$

where we use the Einstein summation convention, and $g^{ij} = g_{ij}^{-1}$.

Let $U \subset \mathbb{R}^3$ be an open set and suppose \bar{U} is compact with boundary ∂U . Let $f_\varepsilon: \bar{U} \rightarrow \mathbb{R}^4$ be a family of hypersurfaces indexed by $\varepsilon > 0$, obtained by deforming f in the normal direction according to the mean curvature. Explicitly, we set

$$f_\varepsilon(x, y, z) := f(x, y, z) + \varepsilon H v(x, y, z). \quad (4)$$

We wish to iterate this leading to a minimal hypersurface, that is $H = 0$ in all of U , except possibly where barriers (atomic constraints) are encountered.

For our purposes, let us choose $f(u) = (x, y, z, S)$, where $S(x, y, z)$ is a function of interest. We have the first fundamental form:

$$(g_{ij}) = \begin{pmatrix} 1 + S_x^2 & S_x S_y & S_x S_z \\ S_x S_y & 1 + S_y^2 & S_y S_z \\ S_x S_z & S_y S_z & 1 + S_z^2 \end{pmatrix}. \quad (5)$$

The inverse matrix of (g_{ij}) is given by

$$(g^{ij}) = \frac{1}{g} \begin{pmatrix} 1 + S_y^2 + S_z^2 & -S_x S_y & -S_x S_z \\ -S_x S_y & 1 + S_x^2 + S_z^2 & -S_y S_z \\ -S_x S_z & -S_y S_z & 1 + S_x^2 + S_y^2 \end{pmatrix}, \quad (6)$$

where $g = \text{Det}(g_{ij}) = 1 + S_x^2 + S_y^2 + S_z^2$ is the Gram determinant. The normal vector can be computed from eq. (1)

$$v = (-S_x, -S_y, -S_z, 1)/\sqrt{g}, \quad (7)$$

The second fundamental form is given by

$$(h_{ij}) = \left(\frac{1}{\sqrt{g}} S_{x_i x_j} \right), \quad (8)$$

i.e., the Hessian matrix of S .

We consider a family $f_\varepsilon = (x, y, z, S_\varepsilon)$, where

$$S_\varepsilon(x, y, z) = S(x, y, z) + \varepsilon H \frac{1}{\sqrt{g}}. \quad (9)$$

The explicit form for the mean curvature can be written as

$$H = \frac{1}{3} \nabla \cdot \left(\frac{\nabla S}{\sqrt{g}} \right). \quad (10)$$

Thus, we arrive at the following evolution scheme

$$S_\varepsilon(x, y, z) = S(x, y, z) + \frac{\varepsilon}{3\sqrt{g}} \nabla \cdot \left(\frac{\nabla S}{\sqrt{g}} \right). \quad (11)$$

To balance the growth rate of the mean curvature operator, we replace H by $3gH$ in eq. (11), which is permissible since g is nonsingular. This leads to the final scheme

$$S_\varepsilon(x, y, z) = S(x, y, z) + \varepsilon \sqrt{g} \nabla \cdot \left(\frac{\nabla S}{\sqrt{g}} \right). \quad (12)$$

Surface Free Energy Minimization

Let us denote the surface free energy of a molecule as $E = \int_U \sigma(x, y, z) d\Omega$, where U encloses the molecule, σ the energy density and $d\Omega = \sqrt{g} dx dy dz$. The energy minimization via the first variation leads to the Euler Lagrange equation,

$$\frac{\partial e}{\partial S} - \frac{\partial}{\partial x} \frac{\partial e}{\partial S_x} - \frac{\partial}{\partial y} \frac{\partial e}{\partial S_y} - \frac{\partial}{\partial z} \frac{\partial e}{\partial S_z} = 0, \quad (13)$$

where $e = \sigma \sqrt{g}$. The explicit form of $\sigma(x, y, z)$ is required in practical applications. For a homogeneous surface, $\sigma = \sigma_0$, a constant, eq. (13) leads to the vanishing of the mean curvature $\sigma_0 \nabla \cdot \left(\frac{\nabla S}{\sqrt{g}} \right) = 3\sigma_0 H = 0$ everywhere except for a set of protected points. This result is consistent with the evolution eq. (12).

Methods and Algorithms

The procedure of the present algorithm is the follows. First we minimize the mean curvature H of a hypersurface function S , while protecting the molecular van der Waals surfaces. Then we extract the desirable MMS from the hypersurface function by choosing a level surface of S .

Minimization of the Mean Curvature

We directly iterate eq. (12) so that $S_\varepsilon(x, y, z) \rightarrow S(x, y, z)$ and $H \rightarrow 0$, except for the constraint surface. For a given set of atomic coordinates, we prescribe a step function initial value for $S(x, y, z)$, i.e., a nonzero constant S_0 inside a sphere of radius \tilde{r} about each atom and zero elsewhere. Alternatively, a Gaussian

or other smooth initial value can be placed around each atomic center. The value of $S(x, y, z)$ is updated in the iteration except for obstacles, i.e., a set of boundary points given by the collection of all of the van der Waals sphere surfaces or any other desired atomic sphere surfaces. The mean curvature, H , can be approximated by any standard numerical method. For simplicity, we use the standard second order central finite difference. Because of the stability concerns, we choose $\varepsilon < \frac{h^2}{2}$, where h is the smallest grid spacing. The iteration converges $S_\varepsilon(x, y, z) \rightarrow S(x, y, z)$ whenever $H \rightarrow 0$ everywhere except for certain protected boundary points where the mean curvature takes constant values. The MMS is differentiable and consistent with surface free energy minimization.

The hypersurface minimization process can be formulated as a mean curvature (geometric) flow^{45–52}

$$\frac{\partial S}{\partial t} = 3\sqrt{g}H = \sqrt{g} \nabla \cdot \left(\frac{\nabla S}{\sqrt{g}} \right). \quad (14)$$

Many variants of eq. (14) can be found in the literature. Since at the hypersurface boundary, the Gram determinant g is dominant by $\|\nabla S\|^2$, it does not make much difference in practice to modify eq. (14) as

$$\frac{\partial S}{\partial t} = \|\nabla S\| \nabla \cdot \left(\frac{\nabla S}{\|\nabla S\|} \right). \quad (15)$$

Other time evolution equations that lead to minimize the mean curvature or some approximation of the mean curvature will work too.

An alternative approach can be pursued via the minimization of the mean curvature of the hypersurface function in the framework of the first variation as discussed in surface free energy minimization. This can be done by coupling with appropriate constraints given by the set of extrema from the molecular van der Waals surfaces. The corresponding Euler–Lagrange equation provides the condition of extremality, and the equation for the minimal molecular hypersurface. The (nonlinear) mean curvature expression can be linearized and discretized as an elliptic equation, and a minimizing sequence can be generated via iterating solvers. This approach is somewhat related to previous theories, such as the Mumford–Shah variational functional,⁵³ and the Euler–Lagrange formulation of surface variation.^{54–61}

Isosurface Extraction

The hypersurface $S(x, y, z)$ obtained via the mean curvature minimization is not the MMS that we seek. Instead, it gives rise to a family of level surfaces, which include the desired MMS. It turns out that $S(x, y, z)$ is very flat away from the MMS, while it sharply varies at the MMS. In other word, $S(x, y, z)$ is virtually a step function at the desirable MMS. Therefore, it is easy to extract the MMS as an isosurface, $S(x, y, z) = C$. It is convenient to choose $C = (1 - \delta)S_I$, where S_I is the initial amplitude, and $\delta > 0$ is a very small number and can be calibrated by standard tests. Computationally, by taking $S_I = 1000$, satisfactory results can be attained by using δ values ranging from

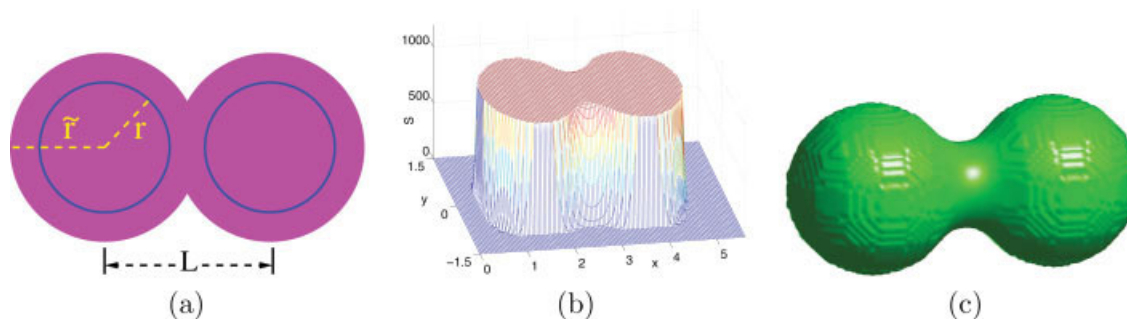


Figure 1. MMS generation. (a) Illustration of r , \tilde{r} , and L at a cross section $z = 0$; (b) $S(x, y, z = 0)$ shows a family of level surfaces. (c) The isosurface extracted from $S = 990$. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

0.004 to 0.01. Mathematically, this process is closely related to the level set algorithm devised by Osher and Sethian.^{59,62,63}

Numerically, isosurface extraction can be done with existing software, such as Matlab and VMD. In fact, finding efficient algorithms for isosurface extraction is an active research topic for volume visualization or scientific visualization. The marching cubes algorithm⁶⁴ and its improvements^{65–70} can be adopted for the purpose of constructing a stand-alone software for the MMS generation.

Results and Discussion

Validation

To validate the proposed theory and algorithm, we first consider the generation of the MMS of a diatomic molecule. We set the atomic radius as r and the distance between the atomic centers as L . Let the atomic centers be $c_1 = (-\frac{L}{2}, 0, 0)$ and $c_2 = (\frac{L}{2}, 0, 0)$. We take an enlarged domain to be $D = \{(x, y, z) : |x, y, z - c_i| < \tilde{r}, \text{ for } i = 1, 2\}$, where $\tilde{r} > L/2 > r$ [see Fig. 1(a) for an illustration]. First we consider a step function initial value $S = S_l = 1000$ in the domain D and $S = 0$ elsewhere. After iterations, the steady state solution of $S(x, y, z)$ is a function on a 3D domain. A cross section of the graph of S is depicted in Figure 1b. As discussed earlier, $S(x, y, z)$ is virtually a step function at the desired MMS, which suggests the uniqueness of the solution. We choose the level set $S(x, y, z) = 0.99S_l$ to obtain the MMS shown in Figure 1c. It consists of parts of the surfaces of the two atoms, i.e., contact surfaces, and a catenoid, i.e., a reentrant surface that connects the two atoms. Although the shape of the present MMS looks similar to the MS of a diatom, we note that the generation procedure of the MMS is different from that for the MS, in which a reentrant surface is generated by rolling a probe.

Although the formation of the connected MMS is automatic, an initial connected domain is an important constraint. We found that if we choose $r < \tilde{r} < L/2$, two isolated spheres will not join to form the MMS. Therefore initial connectivity ($\tilde{r} > L/2$) is crucial for the formation of MMSs.

It is important to know how an initially connected level set of $S(x, y, z)$ will eventually separate into two regions when L is sufficiently large. For two atoms with the same radius $r_1 = r_2 = r$, lower and upper bounds of the MMS area can be taken to be $4\pi r^2$ and $8\pi r^2$, respectively for the diatomic system. When L is

small, the MMS consists of a catenoid and parts of two spheres, and the MMS area is smaller than the upper bound, see Figure 2(a1). When the separation length L is gradually increased, the MMS area grows continuously, while the neck of the MMS surface becomes thinner and thinner, see Figure 2(a2) and 2(a3). At a critical distance $L_c > 2r$, the MMS breaks into two disjoint pieces. The present study predicts $L_c \simeq 2.426r$. In fact, this result is robust with respect to initial domain D as long as $\tilde{r} > \frac{L_c}{2} \simeq 1.213r$. We actually take \tilde{r} to be slightly larger than $1.213r$, say $\tilde{r} = 1.3r$, in our computations. We found that a Gaussian initial value gives the same prediction. Note that although the topological change of the MMS seems to be abrupt at the critical distance because of the break of the catenoid, the MMS surface area is actually a continuous function of the separating distance L . The continuity of the MMS surface area will be discussed rigorously elsewhere.

Diatomic systems with different radii are also considered in Figure 2. Similar transition patterns can be seen in these systems. By comparing the critical states among different cases, a linear dependence of the critical value L_c with respect to r_1 and r_2 can be identified as $L_c \simeq 1.213(r_1 + r_2)$. As the MS area is proportional to the surface Gibbs free energy, the critical value (L_c) might provide an indication of the molecular disassociation critic and could be used in molecular modeling.

We next consider the MMS of the benzene molecule which consists of six carbon atoms and six hydrogen atoms. The carbon atoms are in sp^2 hybrid states with delocalized π stabilization. The MMSs of the benzene molecule with van der Waals radii (r_{vdw}) and other atomic radii are depicted in Figure 3. By using the van der Waals radii, a bulky MMS is obtained. A topologically similar while smaller MMS is formed using the set of standard atomic radii. No ring structure is seen until the atomic radii are reduced by a factor of 0.9, see Figure 3c. Clearly, all atoms are connected via catenoids. Eventually, the MMS decomposes into 12 pieces when radii are further reduced to slightly below their critical values, see Figure 3d. This again confirms to our prediction of critical separation distance $L_c \simeq 1.213(r_1 + r_2)$.

Cavities

Inaccessible internal cavities and open cavities (pockets) are commonly encountered in macromolecules. It is thus interesting and important to explore the representation of these cavities in

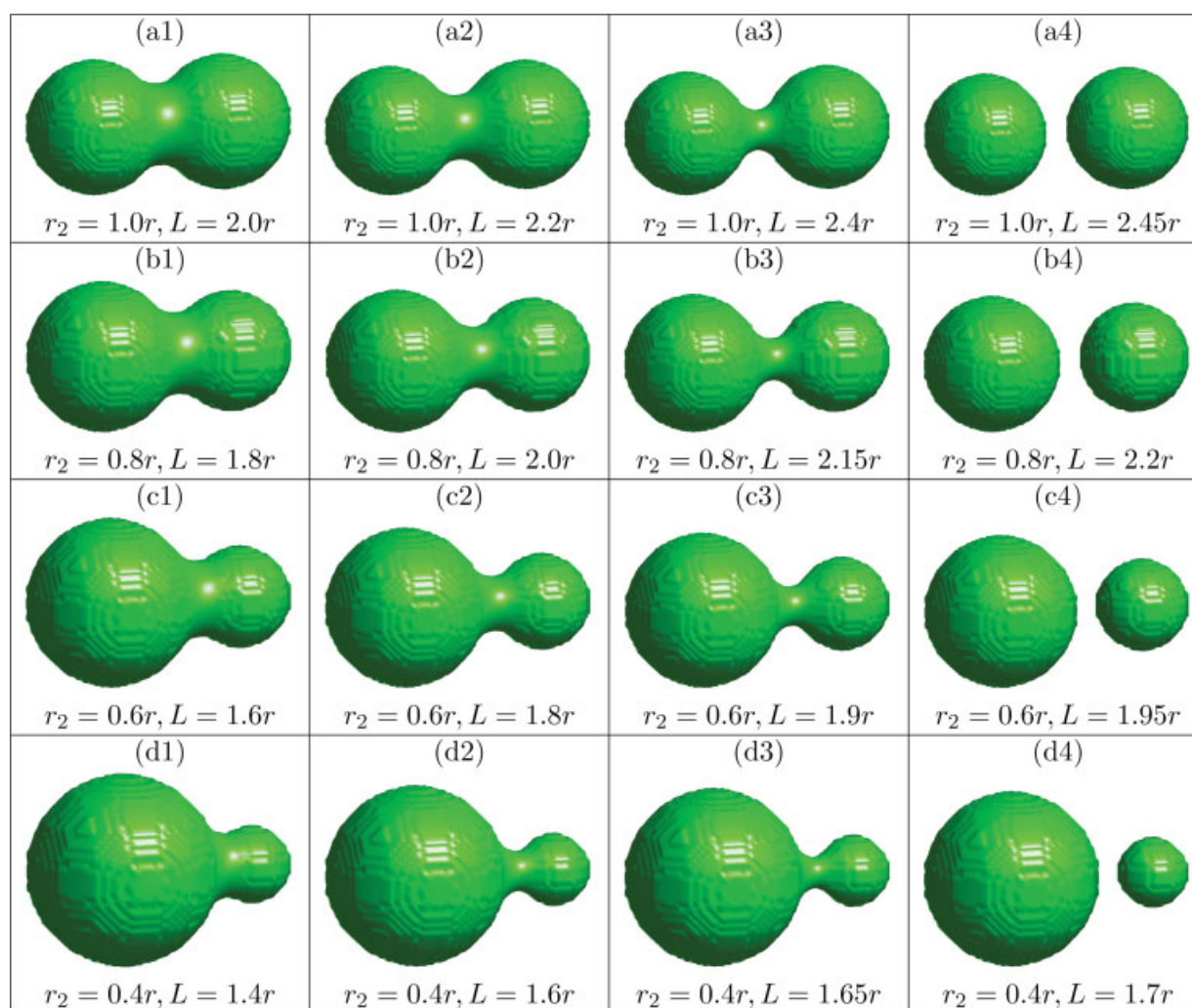


Figure 2. The MMS of a diatomic molecule with $r_1 = r$ and different r_2 and at different separation lengths L . [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

the present MMS algorithm. As an example, we consider a buckyball molecule (buckminsterfullerene C_{60}). The buckyball has 12 pentagons which do not share any edge. To study the performance of our algorithm for open cavities, we create a C_{54} “molecule” by removing 6 carbon atoms, i.e., one hexagon being excluded. However, the coordinates of remaining 54 atoms are unchanged.

Without further constraints, the directly generated MMSs for both original and open buckyballs are found to be free of cavities, because creating a cavity will lead to a larger total surface area. To capture inaccessible and open cavities, certain geometric constraints have to be introduced. From the biological point of view, the surface of a cavity is an interface that encloses foreign objects, such as solvent molecules or ligands. Therefore, the cavity surface is subject to the minimization. A simple way to create a domain for the solvent is to make use of the solvent accessible surface⁴ which is mapped out by rolling a probe of radius r_p on the van der Waals surface of the molecule. Therefore, the domain outside the solvent accessible surface⁴

can be regarded as the solvent domain, where a zero initial value $S(x, y, z) = S_1$ is prescribed and the initial value at the solvent accessible surface is protected during the iteration. In this approach, minimal solvent surfaces (MSSs) are found for internal cavities. For open cavities, a combination of minimal molecular outer surfaces and minimal solvent inner surfaces is obtained. The total surface area is minimal under the designed constraint. In this approach, we have to choose $r_p > \tilde{r} - r$. Obviously, other choices of the cavity constraints are possible. However, a detailed discussion of these possibilities is beyond the scope of the present work.

In the present study, we set $r_p = 1.5 \text{ \AA}$. The MMSs of the buckyball are shown in Figure 4. The first row shows the exterior MMSs. While the second row depicts a cross section of the MMSs. Different van der Waals radii r are used in our plot. At a small radius $r = 0.5 \text{ \AA}$, the MMS breaks into 60 small atomic spheres. Clearly, the inner cavity is very small at $r = 1.7 \text{ \AA}$, but becomes large as r is reduced, see the second row of Figure 4. In this case, the surface of the inner cavity is a minimal solvent surface.

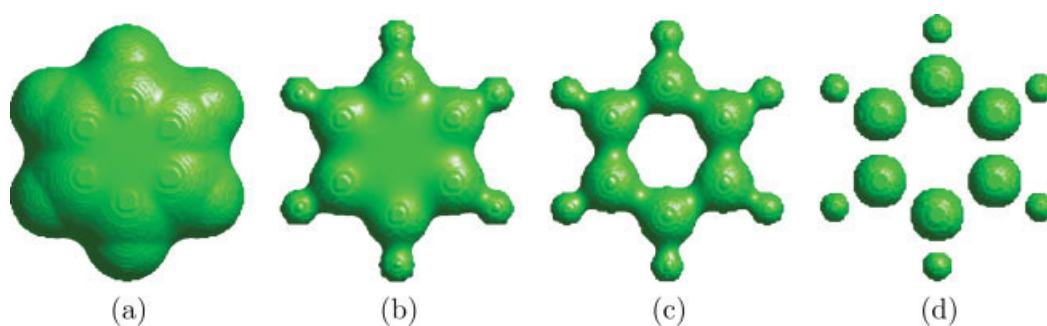


Figure 3. The MMS of benzene with van der Waals radii and scaled atomic radii. (a) Van der Waals radii, $r_C = 1.7 \text{ \AA}$ and $r_H = 1.2 \text{ \AA}$; (b) Atomic radii, $r_C = 0.7 \text{ \AA}$ and $r_H = 0.38 \text{ \AA}$; (c) Scaled atomic radii, $r_C = 0.63 \text{ \AA}$ and $r_H = 0.34 \text{ \AA}$; (d) Scaled atomic radii, $r_C = 0.56 \text{ \AA}$ and $r_H = 0.30 \text{ \AA}$. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

We next study behavior of an open cavity using the fictitious molecule C_{54} with $r_p = 1.5$. Figure 5 presents our results. Similar to the last case, the first row shows exterior views of the MMSs, while the second row provides cross sectional views. The exterior views clearly indicate the removal of a hexagon. As can be seen from the second row, there is an internal cavity when r is large. However, an open cavity is gradually formed at the site of the hexagon as r decreases. Both internal and open cavities enlarge and finally merge into one as r is sufficiently small, see Figure 5(b3). It is interesting to note that there is a smooth transition from the minimal molecular outer surface to the minimal solvent inner surface in Figure 5 (a3) and 5(b3).

Singularities

Earlier biomolecular surfaces, such as the van der Waals surface and the solvent accessible surface, are nonsmooth. The MS was introduced to create smooth surfaces by smoothly joining atomic surfaces with the probe surface. However, the MS is not smooth everywhere (it is not C^1). There are self-intersecting surfaces, cusps, and other singularities in the MS definition. These non-smooth features cause numerical instability in the calculations of electrostatic potentials and forces using implicit solvent models. Moreover, singularities are obstacles to MS generators.^{25,26,30,31} In the present work, we shall carefully examine whether similar singularities occur in our MMS definition. To this end, we carry

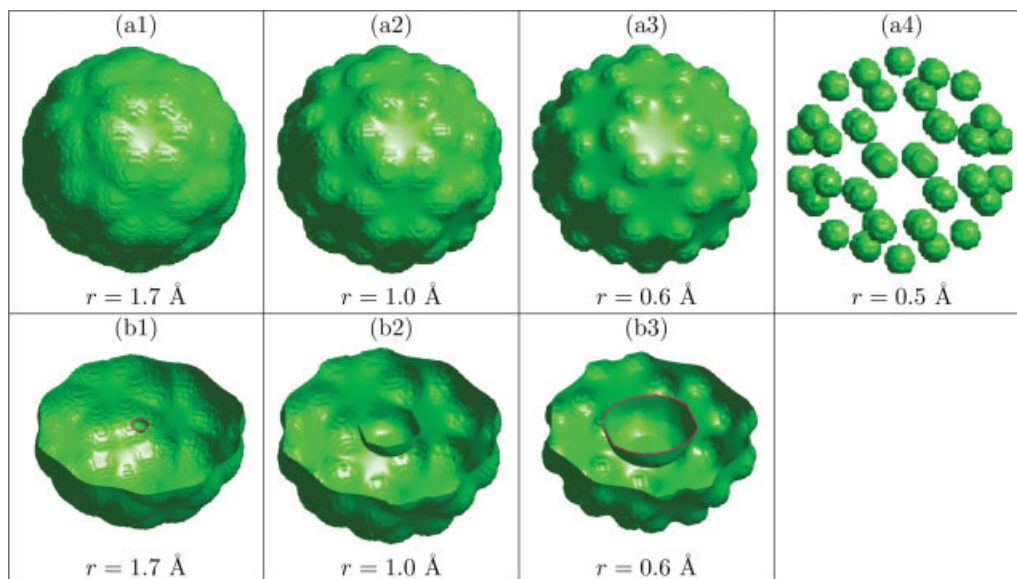


Figure 4. The MMS of the buckyball with van der Waals radius and scaled atomic radius. (a1) Van der Waals radius, $r = 1.7 \text{ \AA}$; (a2) Atomic radii, $r = 1.0 \text{ \AA}$; (a3) Scaled atomic radii, $r = 0.6 \text{ \AA}$; (a4) Scaled atomic radii, $r = 0.5 \text{ \AA}$. (b1), (b2), and (b3) show, respectively, the MMS of the same buckyball as in (a1), (a2), and (a3), with half of the data of $S(x, y, z)$ removed. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

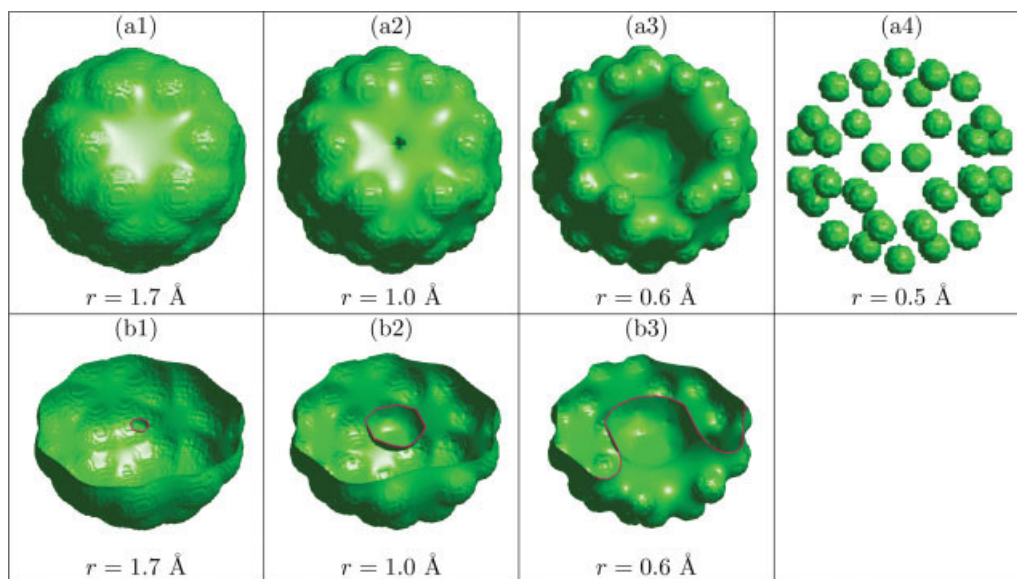


Figure 5. The MMS of the open buckyball with van der Waals radius and scaled atomic radius. (a1) Van der Waals radius, $r = 1.7 \text{ \AA}$; (a2) Atomic radii, $r = 1.0 \text{ \AA}$; (a3) Scaled atomic radii, $r = 0.6 \text{ \AA}$; (a4) Scaled atomic radii, $r = 0.5 \text{ \AA}$. (b1), (b2), and (b3) show, respectively, the MMS of the same open buckyball as in (a1), (a2), and (a3), with half of the data of $S(x, y, z)$ removed. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

out a comparative study of MS and MMS. The MSs are generated by using the MSMS code.²⁶

Figure 6 illustrates MS and MMS for a three-atom topology. MSs are depicted in the first row. In Figure 6(a1), we show that

cusps occur when the probe radius is small. An increase in the probe radius results in a self-intersecting surface, see Figure 6(a2). The edge of such a self-intersecting surface is singular. Its generation is due to the fact that the probe sitting above the

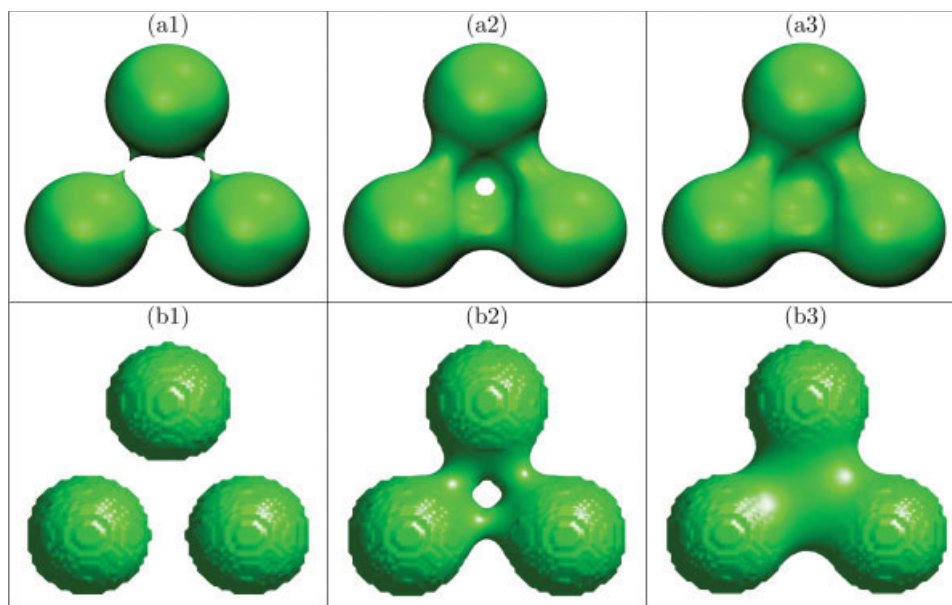


Figure 6. Singularity studies for a three-atom system with radius $r = 1.5$ and coordinates $(x, y, z) = (-2.3, 0, 0)$, $(2.3, 0, 0)$, and $(0, 3.984, 0)$. MSs and MMSs are shown in the first and second rows, respectively. (a1) $r_p = 0.5$; (a2) $r_p = 0.9$; (a3) $r_p = 1.0$; (b1) $r_p = 0.5$; (b2) $r_p = 0.6$; (b3) without probe constraint. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

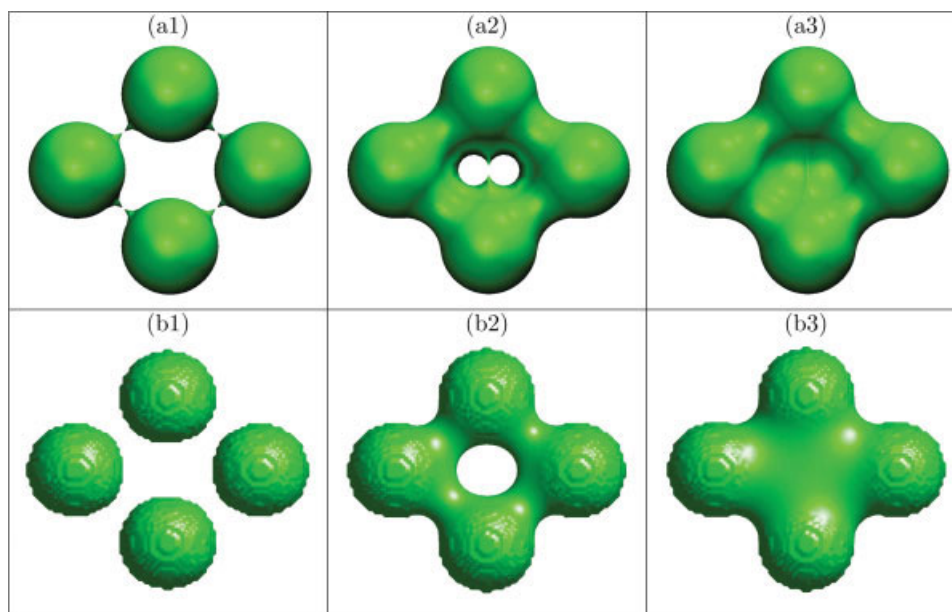


Figure 7. Singularity studies for a four-atom system with radius $r = 1.5$ and coordinates $(x, y, z) = (-2.87, 0, 0), (0, -2.36, 0), (2.87, 0, 0), \text{ and } (0, 2.36, 0)$, MSs and MMSs are shown in the first and second row, respectively. (a1) $r_p = 0.4$; (a2) $r_p = 1.1$; (a3) $r_p = 1.2$; (b1) $r_p = 0.4$; (b2) $r_p = 0.8$; (b3) without probe constraint. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

three atoms intercepts with itself when it is below the three atoms. For the MMS, the same fixed coordinate and atomic radius are used. The probe radius introduced in the last section is used to create a hole at the center of these three atoms. We note that this hole could also be generated without using the probe radius, but through varying either coordinate values or atomic radius. We have varied the probe radius over a large range and found that MMSs are free of cusps and self-intersecting surfaces (see the second row of Fig. 6). We note that in the last case, Figure 6(b3), no probe constraint is used, which is computationally equivalent to r_p being very large.

We next consider a four-atom system. The first row of Figure 7 depicts MS results. Four pairs of cusps in Figure 7(a1) can be clearly seen for a small probe radius. As the probe radius increases to $r_p = 0.5$, four atoms are connected via smooth reentrant surface patches. However, as the probe radius is further increased, a combination of cusps and self-intersecting surface singularities occurs, see Figure 7 (a2). We have explored the parameter space of the MMS, and found no singularities in practice. The MMSs change from four isolated atoms into a four-bead ring, and finally into a bulky four-atom surface as the probe radius is increased.

Intuitively, the MMS generated by minimizing the mean curvature operator generically should be free of cusp and self-intersection surface singularities. However, a mathematical proof of this is not trivial and is out of the scope of the present work.

Applications

We now consider some more elaborate applications of the MMS. Our first task is to generate the MMS of a complex bio-

molecule, a B-DNA double helix segment with 494 atoms (NDB ID: BD0003; PDB ID: 425D). The MMS of the B-DNA generated by using $\bar{r} = 1.3r_{vdw}$ and mesh size $h = 0.2 \text{ \AA}$ is given in Figure 8(a). Similar MMSs can be generated for all other biomolecules in the Protein Data Bank and Nucleic Acid Database. For a comparison, the MS generated by using MSMS²⁶ is depicted in Figure 8(b), with the same set of van der Waals radii and probe radius $r_p = 1.5 \text{ \AA}$. It is interesting to note that the MMS better emphasizes the skeleton of the DNA's double helix structure. Moreover, the MMS is much smoother than the MS, indicating a natural separation boundary between the less polar biomolecule and the polar solvent. Furthermore, the enclosed volume of the MMS is larger than that of the MS at the grooves of the DNA because of the surface minimization. To quantify the difference between the MMS and the MS, the root mean square deviation (RMSD) of the distance of two surfaces is calculated as follows. Consider the Cartesian grid of the MMS generation with x , y , and z meshlines. We first partition the domain into many z -slices based on the z meshline and seek for deviations between the MMS and MS contour lines within each z -slice. This involves the determination of the MMS contour at the given isosurface value. Under the x and y coordinate lines, such a contour is actually piecewise linear, i.e., consists of only straight line segments. For each MS surface triangle edge generated via the MSMS software, we examine if it intersects with the current z -slice. If so, we compute the smallest distance between the intersection point and the line segments of the MMS of the same z -slice. By collecting all deviations, the RMSD between the MMS and MS is estimated to be about 1.048 \AA , which would imply significant differences in many

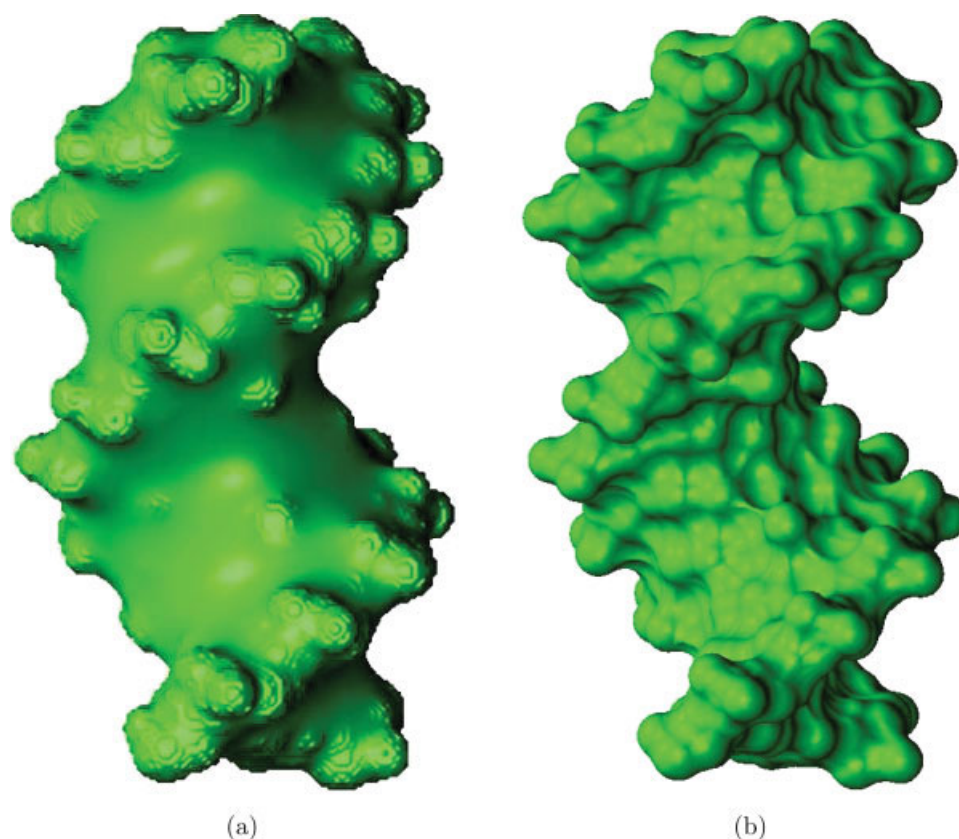


Figure 8. The MMS (a) and the MS (b) of a B-DNA double helix segment. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

physical properties. Further study is required to fully understand the impact and utility of the proposed MMS for biological modeling.

We next consider the MMS of hemoglobin, an important metalloprotein in red blood cells (PDB ID: 1hga). The system has 4649 atoms. The MMS with $\tilde{r} = 1.3r_{vdW}$ and $h = 0.3$ Å and the MS with $r_p = 1.5$ Å are depicted in Figures 9a and 9b, respectively. While the MMS in the B-DNA example remains essentially the same whether the cavity constraint is imposed or not, the present example does require the enforcement of cavity constraint. Only when the solvent accessible surface is considered, can a small pinhole be seen in the MMS near the center of four globular protein subunits. In comparing with the MS of the hemoglobin, it is obvious that the size of pinhole of the MMS is smaller than that of the MS. However, the size of the pinhole of the MMS can be adjusted via probe radius r_p . A smaller r_p will lead to a larger pinhole.

Finally, we consider the application of the MMS to the electrostatic analysis. By defining the MMS as the solvent–solute dielectric interface, the electrostatic potentials of proteins can be attained via the numerical solution of the Poisson–Boltzmann equation. Twenty six proteins, most of them are adopted from a test set used in previous studies,^{71,72} are employed. Two proteins, i.e., Cu/Zn superoxide dismutase (PDB ID: 1b4l) and acetylcholin esterase (PDB ID: 1ea5), are well-known for their

important electrostatic effects. For all structures, extra water molecules are excluded and hydrogen atoms are added to obtain full all-atom models. Partial charges at atomic sites and atomic van der Waals radii are taken from the CHARMM22 force field.⁷³ However, for the Cu/Zn superoxide dismutase, the partial charges on the metal atoms, zinc and copper, and on seven surrounding residues are assigned according to the literature.⁷⁴ By setting the MMS and MS as the dielectric boundaries, electrostatic free energies of solvation ΔG are computed by using the PBEQ,⁷⁵ a finite difference based Poisson–Boltzmann solver from CHARMM,⁷⁶ at mesh sizes $h = 0.5$ Å and $h = 0.25$ Å. The PBEQ was modified to admit the MMS interface. In all test cases, the dielectric coefficient ϵ is set to 1 and 80 respectively for the protein and solvent. The MMS is generated with the probe radius of $r_p = 0.7$ Å to enforce the cavity constraints, and at the half of the grid spacing used in solving the Poisson–Boltzmann equation to ensure the accuracy. A probe radius of 1.4 Å was used for the MS.

The numerical results of electrostatic free energies of solvation are listed in Table 1. For the first twenty four proteins (except 1b4l and 1ea5), the current PBEQ results based on the MS are in excellent agreement with those reported in the literature.^{71,72} This validates our computational procedure. It can be seen from Table 1 and Figure 10(a) that results of the MMS are in good consistent with those of the MS. Figure 10(b) confirms

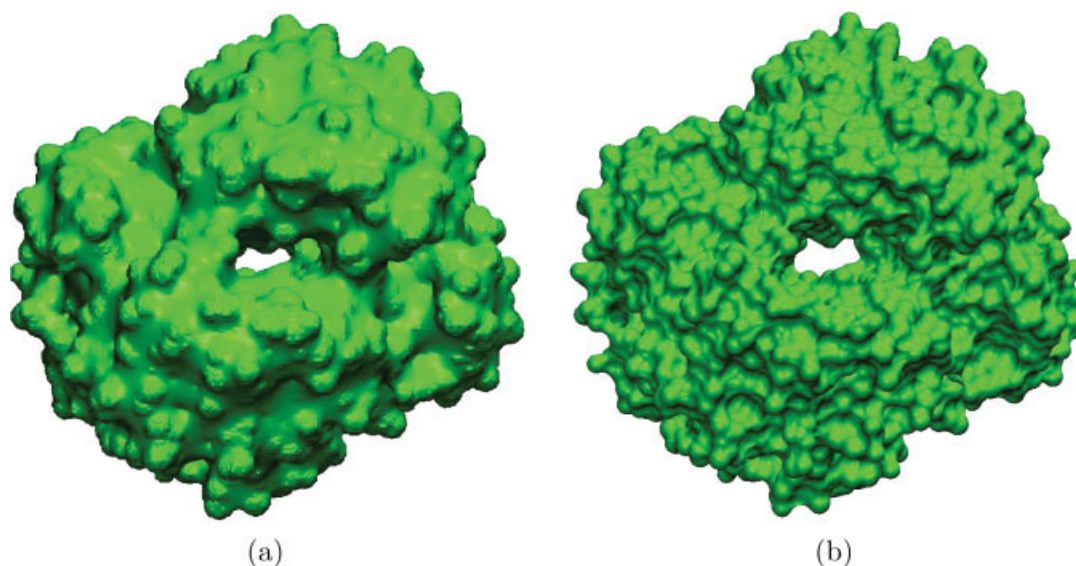


Figure 9. The MMS (a) and the MS (b) of the hemoglobin. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

that the deviations of the results computed with two surfaces are very small. It is to point out this consistence depends on the given probe radii for the MS and the MMS. Inconsistence occurs when any one of these radii is significantly changed. Nevertheless, for a given MS probe radius, we can find an MMS probe radius such that their electrostatic potentials have a good agreement for a large set of proteins. Figure 11 shows the orthographic viewing of the surface electrostatic potentials of the Cu/Zn superoxide dismutase (PDB ID: 1b4l) computed with MS and MMS at $h = 0.5 \text{ \AA}$. Clearly, there is a very good agreement between two potentials. In particular, positively charged active site can be similarly observed in the concave regions in both figures. Nevertheless, some small deviations can still be detected and their impact on the electrostatic steering and electrostatic forces is to be further analyzed in the future.

Conclusion

This article presents a novel concept, the MMS, for the theoretical modeling of biomolecules. A hypersurface function is defined with atomic constraints or obstacles from biomolecular structural information. The mean curvature of the hypersurface function is minimized through an iterative procedure. The MMS is extracted from an appropriate level surface of the hypersurface. The proposed method is systematically validated. The ability of the present method for dealing with internal and open cavities are illustrated. The MMS will not create a cavity that is smaller than a solvent molecule when an appropriate probe radius is used. We demonstrate that MMSs are typically free of singularities. Numerical experiments are carried out for a variety of systems, including simple molecules, DNAs and complex proteins. Twenty six proteins are used to illustrate the electrostatic analysis using the proposed MMS. It is believed that the proposed MMS has the potential to contribute to the development of new methods for the studies of surface chemis-

try, physics and biology, and in particular, on the analysis of stability, solubility, solvation energy, and interaction of macromolecules, such as proteins, membranes, DNAs and RNAs.

Table 1. Electrostatic Free Energies of Solvation Calculated by Using the PBEQ.

h	MMS		MS	
	0.5 Å	0.25 Å	0.5 Å	0.25 Å
1ajj	-1160.1	-1128.9	-1180.3	-1155.9
2pde	-839.9	-815.2	-856.1	-839.2
1vii	-938.1	-914.1	-947.6	-921.5
2erl	-983.4	-958.5	-987.1	-965.1
1cbn	-321.0	-299.8	-332.1	-315.8
1bor	-877.0	-853.1	-899.1	-874.8
1bbl	-1033.2	-998.5	-1039.5	-1011.1
1fca	-1245.4	-1213.1	-1241.3	-1217.6
luxc	-1196.6	-1151.5	-1201.3	-1165.3
1shl	-790.8	-760.4	-801.2	-774.6
1mbg	-1404.0	-1360.6	-1405.0	-1372.5
1ptq	-911.2	-872.9	-928.0	-897.6
1vjw	-1295.8	-1255.1	-1293.9	-1261.7
1fxd	-3352.2	-3318.4	-3356.3	-3327.3
1r69	-1137.9	-1089.1	-1149.8	-1114.7
1hpt	-871.7	-820.1	-877.7	-840.6
1bpi	-1355.9	-1309.5	-1364.7	-1330.1
451c	-1078.5	-1033.6	-1093.7	-1055.1
1a2s	-1972.7	-1929.7	-1981.1	-1944.8
1frd	-2946.1	-2883.8	-2944.2	-2891.8
1svr	-1778.9	-1725.5	-1800.3	-1756.8
1neq	-1818.8	-1759.8	-1816.7	-1768.6
1a63	-2478.2	-2403.5	-2509.5	-2438.5
1a7m	-2241.3	-2172.1	-2279.6	-2211.6
1b4l	-1772.4	-1703.3	-1813.4	-1750.8
1ea5	-6400.2	-6224.1	-6396.2	-6233.8

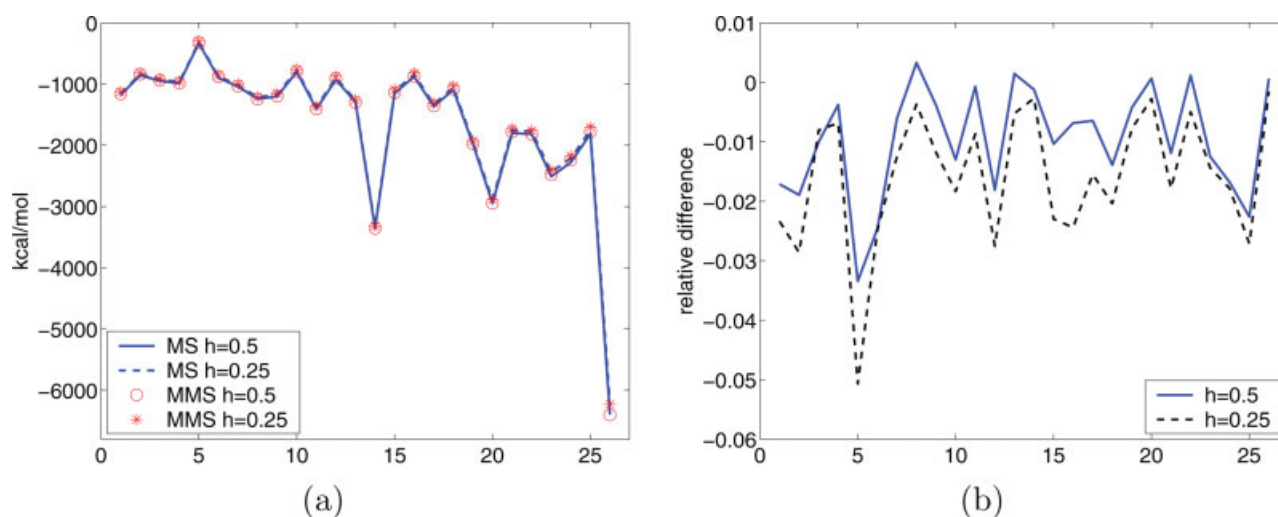


Figure 10. Comparison of electrostatic free energies of solvation ΔG of twenty six proteins listed in Table 1. (a) Electrostatic free energies of solvation ΔG ; (b) Relative differences of solvation free energies: $(\Delta G_{\text{MMS}} - \Delta G_{\text{MS}})/\Delta G_{\text{MS}}$. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

However, the MMS is not designed to replace or mimic other existing surface representations for all purposes. The proposed MMS can be computed via a stand-alone program based on the marching cubes triangulation,⁶⁴ and the computational expense of the MMS depends on the desired level of resolution. In the

current studies of proteins and DNAs, the generation of the MMS usually uses slightly more CPU time than that of the MS using the MSMS code.²⁶ Issues of efficient generation of the MMS and further application to the implicit solvent models are under our consideration.

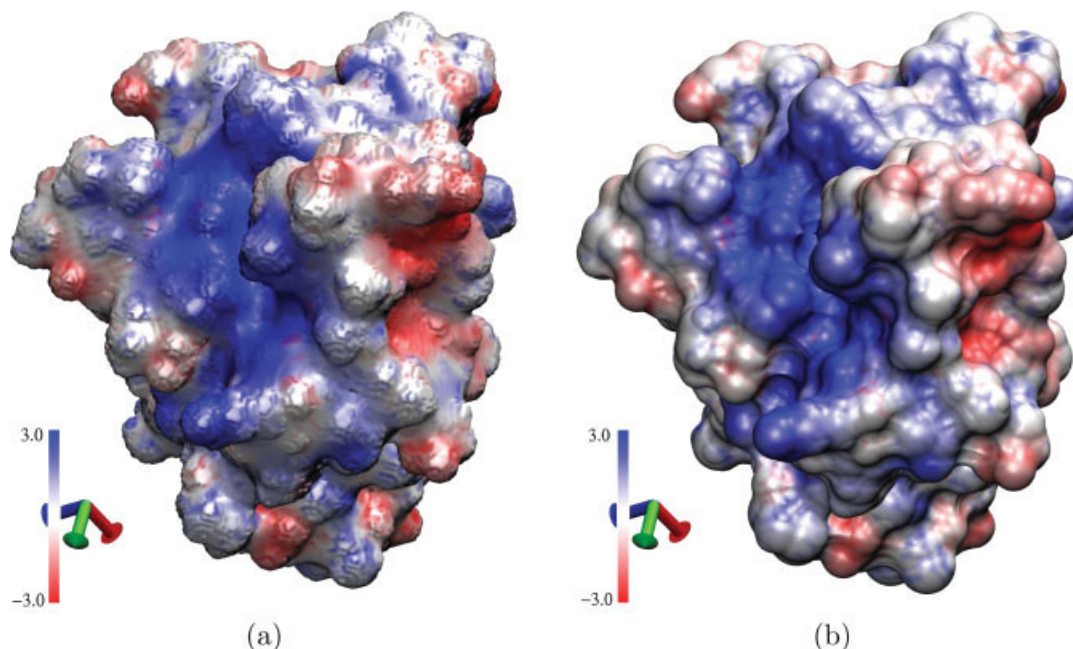


Figure 11. Surface electrostatic potentials of the Cu/Zn superoxide dismutase at $h = 0.5 \text{ \AA}$. (a) Generated with the MMS; (b) Generated with the MS.

Acknowledgments

The authors thank the anonymous referees for helpful suggestions.

References

- Corey, R. B.; Pauling, L. *Rev Sci Instr* 1953, 24, 521.
- Richards, F. M. *Annu Rev Biophys Bioengl* 1977, 6, 151.
- Kuhn, L. A.; Siani, M. A.; Pique, M. E.; Fisher, C. L.; Getzoff, E. D.; Tainer, J. A. *J Mol Biol* 1992, 228, 13.
- Lee, B.; Richards, F. M. *J Mol Biol* 1973, 55, 379.
- Connolly, M. L. *J Appl Crystallogr* 1983, 16, 548.
- Spolar, R. S.; Record, M. T., Jr. *Science* 1994, 263, 777.
- Crowley, P. B.; Golovin, A. *Proteins - Struct Func Bioinf* 2005, 59, 231.
- Bergstrom, C. A. S.; Strafford, M.; Lazorova, L.; Avdeef, A.; Luthman, K.; Artursson, P. *J Medicinal Chem* 2003, 46, 558.
- Dragan, A. I.; Read, C. M.; Makeyeva, E. N.; Milgotina, E. I.; Churchill, M. E. A.; Crane-Robinson, C.; Privalov, P. L. *J Mol Biol* 2004, 343, 371.
- Jackson, R. M.; Sternberg, M. J. *J Mol Biol* 1995, 250, 258.
- LiCata, V. J.; Allewell, N. M. *Biochemistry* 1997, 36, 10161.
- Raschke, T. M.; Tsai, J.; Levitt, M. *Proc Natl Acad Sci USA* 2001, 98, 5965.
- Das, B.; Meirovitch, H. *Proteins* 2001, 43, 303.
- Warwicker, J.; Watson, H. C. *J Mol Biol* 1982, 154, 671.
- Honig, B.; Nicholls, A. *Science* 1995, 268, 1144.
- Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J Chem Phys* 2002, 117, 43.
- Richmond, T. J. *J Mol Biol* 1984, 178, 63.
- Tsodikov, O. V.; Record, M. T.; Sergeev, Y. V. *J Comput Chem* 2002, 23, 600.
- Gibson, K. D.; Scheraga, H. A. *Mol Phys* 1987, 62, 1247.
- Fraczkiewicz, R.; Braun, W. *J Comput Chem* 1998, 19, 319.
- Liang, J.; Edelsbrunner, H.; Fu, P.; Sudhakar, P. V.; Subramaniam, S. *Proteins* 1998, 33, 1.
- Totrov, M.; Abagyan, R. *J Struct Biol* 1996, 116, 138.
- Bhat, S.; Purisima, E. O. *Prot Struct Funct Bioinformatics* 2006, 62, 244.
- Varshney, A.; Brooks, F. P., Jr.; Wright, W. V. *IEEE Comp Graph Appl* 1994, 14, 19.
- Connolly, M. L. *J Appl Crystallogr* 1985, 18, 499.
- Sanner, M. F.; Olson, A. J.; Spehner, J. C. *Biopolymers* 1996, 38, 305.
- Zauhar, R. J.; Morgan, R. S. *J Comput Chem* 1990, 11, 603.
- Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. *J Comput Chem* 2002, 23, 128.
- Wei, G. W.; Sun, Y. H.; Zhou, Y. C.; Feig, M. *arXiv:math-ph* 2005, 0511001.
- Eisenhaber, F.; Argos, P. *J Comput Chem* 1993, 14, 1272.
- Gogonea, V.; Osawa, E. *Supramol Chem* 1994, 3, 303.
- Andersson, S.; Hyde, S. T.; Larsson, K.; Lind, S. *Chem Rev* 1998, 88, 221.
- Anderson, M. W.; Egger, C. C.; Tiddy, G. J. T.; Casci, J. L.; Brakke, K. A. *Angew Chem Int Ed* 2005, 44, 3243.
- Pociecha, D.; Gorecka, E.; Vaupotic, N.; Cepic, M.; Mieczkowski, J. *Phys Rev Lett* 2005, 95, 207801.
- Seddon, J. M.; Temple, R. H. *Philos T Royal Soc London Ser A-Math Phys Eng Sci* 1993, 244, 377.
- Du, Q.; Liu, C.; Ryham, R.; Wang, K. *Commun Pure Appl Anal* 2005, 4, 537.
- Chen, B. L.; Eddaoudi, M.; Hyde, S. T.; O'Keeffe, M.; Yaghi, O. M. *Science* 2001, 291, 1021.
- Koh, E.; Kim, T. *Prot Struct Funct Bioinformatics* 2005, 61, 559.
- Falicov, A.; Cohen, F. E. *J Mol Biol* 1996, 258, 871.
- Gray, A. *Modern Differential Geometry of Curves and Surfaces with Mathematica*, 2nd ed.; CRC Press: Boca Raton, 1998.
- Chopp, D. L. *J Comput Phys* 1993, 106, 77.
- Cecil, T. *J Comput Phys* 2005, 206, 650.
- Bates, P. W.; We, G. W.; Zhao, S. The minimal molecular surface, Midwest Quantitative Biology Conference, Mission Point Resort, Mackinac Island, MI, September 29 – October 1, 2006.
- We, G. W.; Zhao, S.; Bates, P. W. A minimal surface generator, Patent disclosure with the Michigan State University Intellectual Property Office, Michigan State University, MI, 2006.
- Feng, X. B.; Prohl, A. *Math Comput* 2004, 73, 541.
- Gomes, J.; Faugeras, O. *Lect Notes Comput Sci* 2001, 2106, 1.
- Mikula, K.; Sevcovic, D. *Math Meth Appl Sci* 2004, 27, 1545.
- Osher, S.; Fedkiw, R. P. *J Comput Phys* 2001, 169, 463.
- Sarti, A.; Malladi, R.; Sethian, J. A. *Int J Comput Vis* 2002, 46, 201.
- Sbert, C.; Sole, A. F. *J Math Imag Vis* 2003, 18, 211.
- Sethian, J. A. *J Comput Phys* 2001, 169, 503.
- Sochen, N.; Kimmel, R.; Malladi, R. *IEEE T Image Proc* 1998, 7, 310.
- Mumford, D.; Shah, J. *Commun Pure Appl Math* 1989, 42, 577.
- Blomgren, P. V.; Chan, T. F. *IEEE Trans Image Process* 1990, 7, 304.
- Li, Y. Y.; Santosa, F. *IEEE T Image Proc* 1996, 5, 987.
- Carstensen, V.; Kimmel, R.; Sapiro, G. *Int J Comput Vis* 1997, 22, 61.
- Osher, S.; Rudin, L. *SIAM J Numer Anal* 1990, 27, 919.
- Osher, S.; Rudin, L. *Proc SPIE Appl Digital Image Process XIV* 1991, 1567, 414.
- Rudin, L.; Osher, S.; Fatemi, E. *Physica D* 1992, 60, 259.
- Sapiro, G.; Ringach, D. *IEEE Trans Image Process* 1995, 5, 1582.
- Sapiro, G. From active contours to anisotropic diffusion: Relation between basic PDE's in image processing. In *proceedings of ICIP Lausanne*, 1996.
- Osher, S.; Sethian, J. A. *Comput Phys* 1988, 79, 12.
- Osher, S. *SIAM J Math Anal* 1993, 24, 1145.
- Lorensen, W. E.; Cline, H. E. *Comput Graphics* 1987, 21, 163.
- Brodie, K.; Wood, J. *Comput Graphics Forum* 2001, 20, 125.
- Cignoni, P.; Marino, P.; Montani, C.; Puppo, E.; Scopigno, R. *IEEE Trans Vis Comput Graphics* 1997, 3, 158.
- Livnat, Y.; Shen, H.; Johnson, C. *IEEE Trans Vis Comput Graphics* 1996, 2, 73.
- Shroeder, W.; Martin, K.; Lorensen, B. *The Visualization Toolkit: An Object-Oriented Approach to 3D Graphics*; Prentice-Hall: Englewood Cliffs, NJ, 1996.
- Wilhelms, J.; Van Gelder, A. *ACM Trans Graphics* 1992, 11, 201.
- Shen, H. W.; Hansen, C. D.; Livnat, Y.; Johnson, C. R. In *Proceedings IEEE Visualization 96*, IEEE Press, 1996; pp. 287–294.
- Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D. A.; Brooks, C. L., III. *J Comput Chem* 2004, 25, 265.
- Zhou, Y. C.; Feig, M.; Wei, G. W. *J Comput Chem* (in press).
- MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, J. D.; Evanseck, M. J.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuczera, L.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J Phys Chem* 1998, 102, 3586.
- Shen, J.; Wong, C. F.; Subramaniam, S.; Albright, T. A.; McCammon, J. A. *J Comput Chem* 1990, 11, 346.
- Im, W.; Beglov, D.; Roux, B. *Comput Phys Commun* 1998, 111, 59.
- Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; Stats, D. J.; Swaminathan, S.; Karplus, M. *J Comput Chem* 1983, 4, 187.