Efficient Solution of Poisson-Boltzmann Equation for Electrostatics of Large Molecules

Xuejun Hao **Amitabh Varshney**
Department of Computer Science and UMIACS
University of Maryland at College Park
College Park, MD 20742
{hao, varshney}@cs.umd.edu

Keywords: Poisson-Boltzmann equation (PBE), finite difference method, tetrahedral decomposition, iso-surface generation, level-of-detail hierarchy

Abstract

Development of fast computational methods to solve the Poisson-Boltzmann equation (PBE) for molecular electrostatics is important because of the central role played by electrostatic interactions in many biological processes. The accuracy and stability of the solution to the PBE is quite sensitive to the boundary layer between the solvent and the solute which defines the solvent-accessible surface. In this paper, we propose a new interface-layer-focused PBE solver for efficiently computing the electrostatic potential for large molecules. Our method analytically constructs the solvent-accessible surface of molecules and then builds nested iso-surface layers outwards and inwards from the surface using the distance field around the surface. We then develop a volume simplification algorithm to adaptively adjust the density of the irregular grid based on the importance to the PBE solution. We also generalize finite difference methods on our irregular grids using Taylor series expansions. Our algorithm achieves about three times speedup in the iterative solution process of PBE, with more accurate results on an analytical solvable testing case, compared with the popular optimized DelPhi program. Our approach can also be applied directly to solve partial differential equations arising in other application domains.

1 INTRODUCTION

Electrostatic interactions are of central importance for many biological processes [Leach 2001]. Electrostatics influences various aspects of nearly all biochemical reactions [Baker et al. 2001], such as macromolecular folding and conformational stability. It also determines the structural and functional properties of biological samples, such as shapes, binding energies, and association rates [Sharp and Honig 1990]. So the modeling of electrostatics in molecular modeling packages [Humphrey et al. 1996] has great practical, as well as, theoretical importance, for structure-based drug design and protein folding.

Electrostatic properties of biological samples can be modeled by quantum mechanical methods or classical electrostatics. Quantum mechanical methods are more accurate, but due to their immense computational demands, can only be applied to small molecules. Classical electrostatics interactions are modeled as interactions between partial atomic charges. They depend not only on 3D structure of molecules and charge distributions, but also on the environment. Biological processes occur in aqueous solution, so solvent plays an important role in determining the electrostatics of solute molecules. Due to the huge computational cost, solvent properties are normally described in terms of average values. That means, instead of treating each solvent atom explicitly, we treat them as a continuum with average properties. The more important solute molecules are described in atomic details [Honig and Nicholls 1995]. These result in a Poisson-Boltzmann equation (PBE) for describing the electrostatic interactions in solution.

In this paper, we propose a new algorithm for efficiently solving PBE. Our method is based on the observation that the accuracy and stability of the solution to PBE is quite sensitive to the boundary layer between the solvent and the solute. So an accurate construction of this boundary with adaptively controlled grid density should improve the solution. In our algorithm, we first analytically construct the solvent-accessible surface of the molecules, then build a tetrahedral decomposition of the 3D space around the surface and also build a distance field from the surface. Next we build iso-surfaces by the marching-tetrahedra method on the distance field with progressively greater distances. This results in nested isosurfaces at varying distances from

the solvent-accessible surface. After that, we apply a volume simplification algorithm to simplify the tetrahedral grid and adaptively adjust grid density based on the influence it exerts on the solution. We maintain a higher resolution for the solution-sensitive region in the vicinity of the solvent-accessible surface. We have found that this improves the accuracy and stability of the solution while speeding up the computation. In our work we have generalized the traditional finite difference computations on regular grids to irregular grids by Taylor series expansions. We validate our algorithm on an analytical solvable case and compare the results with the popular DelPhi program.

The main contributions of this paper are:

1. We design an efficient algorithm to solve PBE by taking advantage of the exact geometry around the solution-sensitive region.

2. We show how to achieve better results by using an application-driven hierarchy of detail for 3D space decomposition.

2 PREVIOUS WORK

Modern electrostatic models are based on Poisson-Boltzmann equation (PBE). PBE is non-linear but can be approximated by a linear equation with a possible analytic solution [Lee and Richards 1971; Tanford and Kirkwood 1957], if there are no highly charged molecules or high ionic strengths. Among numeric methods to solve PBE, the finite difference method (FDM) [Warwicker and Watson 1982] is the most widely used one, in which the molecule is mapped onto a three-dimensional grid. Ionizable atoms are assigned to grid points and the electrostatic potential at each grid point is calculated using the finite difference approximation of the PBE. The accuracy of the results is highly dependent on grid spacing, while the computational cost increases steeply with the number of grid points. One approach to reduce the cost is called focusing [Gilson et al. 1988], in which the mesh of the grid is reduced only in the vicinity of ionizable groups of particular interest with potentials from coarser grids used as initial guesses.

Adaptive space-subdivision approach [Baker et al. 2000; Baker et al. 2001; Holst et al. 2000] has been used to address the high cost of using a regularly-spaced grid. This approach increases the accuracy of the solution by explicitly giving a higher spatial resolution to the solvent-solute boundary region, to which the solution and the rate of convergence is highly sensitive. The drawbacks are the price for slow convergence resulting from the over-subdivision around the boundary region and the difficulty in exactly placing the points on the solute-solvent boundary.

In applications such as computational fluid dynamics, adaptive mesh refinement (AMR) techniques [Berger and Colella 1989; McCormick 1989] have been developed to refine the grid only in the regions where it is needed. They do not incur the over-subdivision problem. They achieve this by making their mesh refinements quite flexible so that the coarser and finer mesh do not have to coincide at their boundaries. In our approach we use similar ideas for efficiently computing molecular electrostatics by using the solvent-accessible surface boundary.

3 MOLECULAR ELECTROSTATICS BACKGROUND

The main trends of molecular electrostatics theory follow the principles of classical electrostatics, treating the solvent as a continuum in terms of average value. In this section, we lay out the physical foundations for molecular electrostatics.

The electrical potential satisfies the Poisson equation in a uniform dielectric medium [Jackson 1975]:

$$\nabla^2\phi(\overrightarrow{r}) + \frac{4\pi\rho(\overrightarrow{r})}{\epsilon} = 0$$

where $\phi(\overrightarrow{r})$ is the electrostatic potential, and $\rho(\overrightarrow{r})$ is the charge density. ϵ is the dielectric constant in uniform media.

If the charges are continuously distributed, the electrical potential is given by the following integral:

$$\phi(\overrightarrow{r}) = \iiint \frac{\rho(\overrightarrow{r'})}{\epsilon\left|\left|\overrightarrow{r} - \overrightarrow{r'}\right|\right|} d\overrightarrow{r'}$$

where $\rho(\overrightarrow{r'})$ is the charge density. The integral is over the space.

If the dielectric ϵ varies through space, then we arrive at a general form of the Poisson equation:

$$\nabla[\epsilon(\overrightarrow{r})\nabla(\phi(\overrightarrow{r})] + 4\pi\rho(\overrightarrow{r}) = 0$$

where $\epsilon(\overrightarrow{r})$ is a function of position. Normally the solute is treated as a uniform medium with a low dielectric constant of about $2 \sim 4$, relative to the dielectric constant in vacuum. The solvent is also treated as a uniform medium with a relative dielectric of about 80 [G. Allen (editor) 1999].

Incorporating the environmental response, we get a general equation for the molecular electrostatics the Poisson-Boltzmann Equation (PBE) [Sharp and Honig 1990]:

$$\nabla[\epsilon(\overrightarrow{r})\nabla(\phi(\overrightarrow{r})] - \kappa'^2(\overrightarrow{r})\sinh[\phi(\overrightarrow{r})] + 4\pi\rho(\overrightarrow{r}) = 0$$

where κ' is the modified Debye–Huckel parameter and defined as:

$$\kappa'^2 = \frac{8\pi N_a e^2 I}{1000kT}$$

where N_a is Avogadro's number, e is the electron charge, k is Boltzmann constant, T is the absolute temperature, and I is the ionic strength of the bulk solution. The variables ϕ, ϵ, κ', and ρ are all functions of the position vector $\overrightarrow{r}$. The general PBE above incorporates electronic and dipole polarization through ϵ and ion-screening through κ'.

4 INTERFACE-FOCUSED FDM

We find from previous discussion that the region around the solvent-accessible surface is critical to the accuracy of the FDM solution of PBE. This is due to two facts. First, all the atomic charges are within the molecule. Second, there are huge differences in the dielectric constants and ionic strengths between the two regions separated by the boundary layer. As several biological processes occur at or near the molecular surface, the accuracy of the solution to PBE close to this boundary layer is critical. Not coincidentally, the stability and accuracy of numerical methods also depend largely on the discretization of the grid in this region. To the best of our knowledge, no previous algorithm for solving PBE for molecules exists that builds the tetrahedral solution grid based on the solvent-accessible surface at the solvent-solute boundary. In this paper we present a new algorithm to solve PBE more efficiently by building an adaptive tetrahedral space-decomposition about the solvent-accessible molecular surface. The main idea is to give higher priority and resolution to the boundary region, and lower priority and resolution to other non-critical regions. We also adjust the grid density to be close to uniform in each region.

4.1 Analytical Solvent-accessible Surface

Previous methods to solve PBE approximate the solvent-accessible surface after building a 3D grid around the molecule. For each grid point, a binary marker indicates whether it is inside the molecule or inside the solvent. The solvent-accessible surface is then defined as passing between those grid points which have different markers. This way, the accuracy of the surface is limited to the grid mesh resolution on which it is generated, and the real surface points do not in general, coincide with the preset grid points.

Several analytical solvent-accessible surface generation algorithms have been published [Akkiraju and Edelsbrunner 1996; Bajaj et al. 2003; Klein et al. 1990; Sanner and Olson 1997; Varshney et al. 1994]. After we analytically generate the solvent-accessible surface using the approach in [Varshney et al. 1994], we incorporate it in the 3D grid used for the solution of the PBE. Guaranteeing that the grid points at the boundary layer are actually on the exact surface, improves the accuracy of the algorithm.

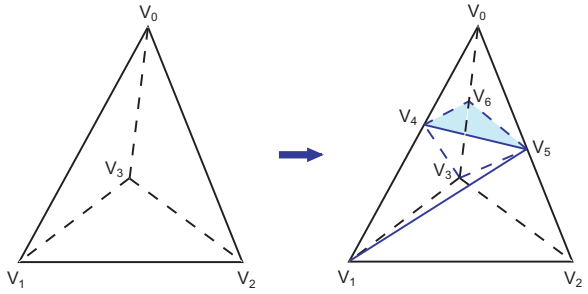


Figure 1: Marching tetrahedron

4.2 Tetrahedral Decomposition of Space and Iso-surface Layers Construction

The accuracy of the FDM solution to PBE depends on the ionic strength assignment, which is zero within $2\mathring{A}$ from the molecular surface, and constant beyond. So we also need to generate an accurate ionic screening surface, which lies in the solvent and is $2\mathring{A}$ away from the molecular surface.

After we construct a 3D grid over the molecule using the exact solvent-accessible surface, we tetrahedralize the volume. We use an odd/even scheme for splitting rectilinear and curvilinear grids into tetrahedra as done in [Max et al. 1990]. Then we generate a signed-distance map of the space, which measures the distance of each grid point to the solvent-accessible surface (points inside the molecule are assigned negative distances). We use a method similar to the one described in [Gibson 1998] to build the distance map.

Next, we generate a sequence of iso-surfaces from the distance map using a tetrahedral variant of the Marching Cubes algorithm [Lorensen and Cline 1987]. We use tetrahedra instead of cubes for simplicity and stability. We insert new grid points into the 3D grid such that they form surfaces at a fixed distance away from the real solvent-accessible surface. One case of marching tetrahedra is shown in Figure 1. Here the processing of tetrahedron $V_1V_2V_3V_0$ generates triangle $V_4V_5V_6$, which splits the original tetrahedron into four new tetrahedra: $V_4V_5V_6V_0$, $V_4V_3V_5V_6$, $V_1V_5V_3V_4$, and $V_1V_2V_3V_5$.

4.3 Volume Simplification

For accuracy, we need higher grid resolution near the PBE solution-sensitive region around the molecular surface, while for efficiency we prefer to have lower grid resolution in regions that are away from the surface. In regions of the grid that approximately have the same resolution, we would like the grid density to be close to

uniform to avoid slow convergence that sometimes arises due to very fine grid spacing. We use an adaptive volume simplification algorithm to adjust the grid resolution progressively and seamlessly based on the distance from the molecular surface.

Our volume simplification is based on an edge-collapse scheme. Specifically, we use a *half-edge* collapse method [Pajarola 2001], in which we pick one of the end vertices of the edge to collapse to. This guarantees that if we collapse an on-surface edge, the new vertex will also be on the surface. Each half-edge collapse decreases the vertex count by one, and decreases the triangle and tetrahedron count based on its local connectivity. As an example, Figure 2 shows the collapse of edge E resulting in a decimation of vertex V_2 (merged with vertex V_1) and tetrahedron $V_0V_1V_2V_5$, while the tetrahedron $V_0V_2V_3V_5$ has been adjusted to $V_0V_1V_3V_5$.

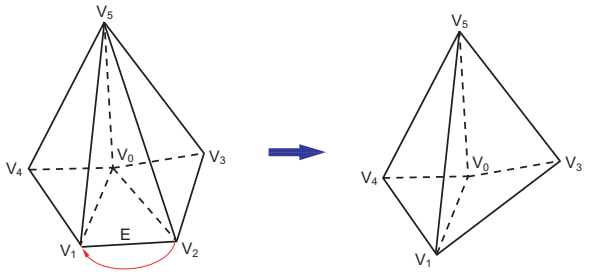


Figure 2: Half-Edge collapse for Tetrahedral Grid Decimation

The purpose of the tetrahedral volume simplification is to enable more efficient computation of the molecular electrostatics. During the simplification, we have to be careful not to simplify the volume into a state in which some grid points lose part of their necessary neighbors which are required for FDM computation. We also avoid generating negative volume tetrahedra (i.e., tetrahedra with a wrong orientation) by checking the sign of the tetrahedral volumes during the simplification process. We verify that the above conditions do not arise as a result of a candidate half-edge collapse, before we allow that half-edge collapse to proceed.

4.4 Derivatives for Irregular Grids

FDM solver of PBE has to compute first and second derivatives of the 3D potential field at each grid point. The derivatives can be computed for regular grids by taking the finite differences between the potential value at each grid point with values at their axis-aligned neighboring grid points. The regular structure of the regular grids makes this procedure straight forward. For irregular grids, the situation is complicated by the fact that not only the distances between grid points vary, but also the neighboring points are rarely axis-aligned. So instead we compute the derivatives of the potential at each grid point using the values at the points that define the axis-inclusion tetrahedra of the grid point. We interpolate to get the value at a new point which is the axis-aligned neighbor of the grid point. Then the derivatives at the grid point can be computed in the same way as for regular grids. This approach is similar in spirit to the approach suggested by Moore [Moore 1999], though he has focused on irregular rectangular grids. Moore has also done a careful error estimate for this situation.

5 RESULTS AND DISCUSSION

In this section we discuss the results obtained using our algorithm. We have used a 2GHz Pentium 4 PC running Windows 2000 with a nVIDIA GeForce3 graphics card. We present results on an analytical solvable case and compare them with the results by the well-known DelPhi implementation. We also show our results on real molecular datasets and display the results on smooth solvent-accessible molecular surfaces.

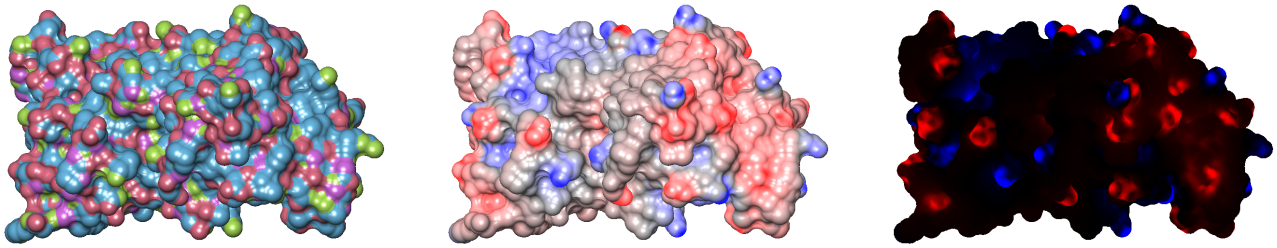
Normally it is difficult (or impossible) to obtain analytical solutions to PBE. In some special cases, we may have analytical solutions to the linearized PBE. One example is that of a spherical molecule with total charge q uniformly distributed on the surface immersed in a solvent containing mobile ions. The analytical solution to this special case is [Holst 1993]:

$$\begin{cases} \phi(r) = \frac{q}{\epsilon_2 R}\left(1 - \frac{R\kappa}{1+\kappa a}\right) & \text{inside molecule} \\ \phi(r) = \frac{q}{\epsilon_2 r}\left(1 - \frac{r\kappa}{1+\kappa a}\right) & \text{ion-exclusion layer} \\ \phi(r) = \frac{qe^{\kappa a}}{\epsilon_2(1+\kappa a)} \cdot \frac{e^{-\kappa r}}{r} & \text{inside solvent} \end{cases}$$

where q is the total sphere surface charge and R is the radius of the sphere. We have tested this analytical solvable case using both our algorithm and the DelPhi (V.4) program. The results are summarized in Table 1. For the testing case, we use a spherical surface charge with a diameter of 27$\mathring{A}$ and a positive charge of 20e (where e is the charge of an electron). The sphere is immersed in a cubic solvent volume with each side 66$\mathring{A}$ long. The average error in Table 1 is defined as the average of the relative error over all the grid points. Peak-signal-to-noise-ratio (PSNR) is defined as $20\log_{10}(\frac{signal\ energy}{noise\ energy})$. The *signal energy* is defined as the sum of the squares

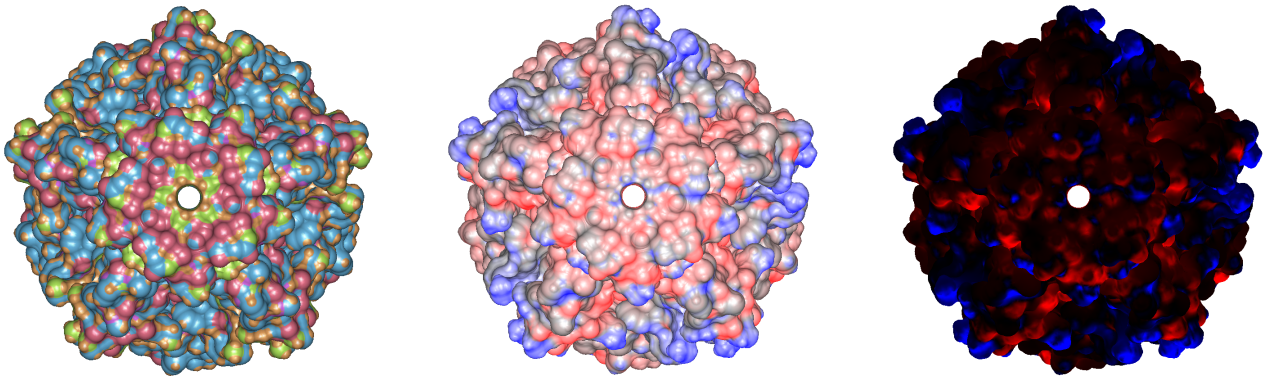
	DelPhi			Our Method
Grid size	67^3	133^3	199^3	N/A
# of pts	300,763	2,352,637	7,880,599	26,987
PSNR	8.17	19.1	25.1	27.7
Avg. error	30.88%	17.91%	13.27%	15.98%
PBE Time	0.31 sec	4.50 sec	20.09 sec	0.25 sec

Table 1: Comparison of our method with DelPhi



(a) Solvent-accessible surface (b) Lighted on-surface potential (c) On-surface potential

Figure 3: Electrostatics on SuperOxide Dismutase (SOD) dataset (red is for negative potential, and blue is for positive potential)



(a) Solvent-accessible surface (b) Lighted on-surface potential (c) On-surface potential

Figure 4: Electrostatics on Ecoli dataset (red for negative potential, and blue for positive potential)

of the potential values over all grid points. The *noise energy* is defined as the sum of the squares of the errors over all grid points. PBE time is the time for solving the Poisson-Boltzmann Equation. One can see from Table 1 the advantages of our method. With the same accuracy, our method needs only $27K$ points instead of several million as needed by DelPhi, and takes only 0.25 seconds to converge, compared with several seconds by DelPhi. For about the same amount of time, our method can be much more accurate than DelPhi, e.g., 15.98% instead of 30.88% error. Our algorithm with $27K$ points has even higher PSNR than DelPhi with about $8M$ points.

We next display our results for some real molecules. The first dataset is superoxide dismutase (SOD) enzyme, which consists of 2196 atoms. Our second dataset is a channel on the outer membrane of the Escherichia coli (Ecoli) bacterium molecule [Sukharev et al. 2001], which consists of 10585 atoms. The results are shown in Figure 3 and 4. Figures 3(a) and 4(a) display the smooth solvent-accessible surfaces of SOD and Ecoli membrane channel using the SURF algorithm [Varshney et al. 1994]. Figures 3(b),(c) and Figures 4(b)(c) display the electrostatic potential on the surfaces, with red for negative and blue for positive potential. Both Figures 3(b) and 4(b) use the potential information to modulate lighting color with grey for neutral potential. An alternative visualization is presented in Figures 3(c) and 4(c) that display the electrostatic potential information directly, with black as neutral potential.

6 CONCLUSIONS

We have presented a new algorithm for efficiently computing electrostatic potentials for large molecular datasets. Our methods give higher priority and resolution to the solution-sensitive region to improve the accuracy and accelerate convergence rates. We build a 3D tetrahedral partition of the space directly from an analytically constructed interface layer. We also provide an algorithm to control the density and uniformity of the grid by using an edge collapse scheme. Compared with the state-of-the-art method using analytically solvable testing case, our method is faster and more accurate. The advantage of our algorithm in solving partial differential equations directly from the geometrical point of view gives it a broad range of possible applications in other application domains.

7 ACKNOWLEDGEMENTS

We would like to acknowledge the reviewers for their detailed and constructive comments which have led to a much better presentation of our results. This work has been supported in part by the NSF grants: IIS-00-81847, and ACR-98-12572/02-96148.

References

Akkiraju, N., and Edelsbrunner, H. 1996. Triangulating the surface of a molecule. *Discrete Appl. Math. 71*, 5–22.

Bajaj, C., Pascucci, V., Shamir, A., Holt, R., and Netravali, A. 2003. Dynamic maintenance and visualization of molecular surfaces. *Discrete Appl. Math. 127*, 23–51.

Baker, N., Holst, M., and Wang, F. 2000. Adaptive multilevel finite element solution of the poisson-boltzmann equation ii: refinement at solvent accessible surfaces in biomolecular systems. *J. Comput. Chem. 21*, 1343–1352.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. In *Proc. Natl. Acad. Sci. USA 98*, 10037–10041.

Berger, M. J., and Colella, P. 1989. Local adaptive mesh refinement for shock hydrodynamics. *Journal of Computational Physics 82(1)*, 64–84.

G. Allen (editor). 1999. *Protein : a comprehensive treatise, vol 2, 61 – 97*. JAI Press.

Gibson, S. F. 1998. Using distance maps for accurate surface representation in sampled volumes. In *IEEE Symposium on Volume Visualization*, 23–30.

Gilson, M. K., Sharp, K. A., and Honig, B. 1988. Calculating electrostatic interactions in biomolecules: method and error assessment. *J. Comp. Chem. 9*, 327 – 335.

Holst, M., Baker, N., and Wang, F. 2000. Adaptive multilevel finite element solution of the poisson-boltzmann equation i: algorithms and examples. *J. Comput. Chem. 21*, 1319–1342.

Holst, M. J. 1993. Multilevel methods for the poisson-boltzmann equation. *Ph.D. thesis, Numerical Computing Group, University of Illinois at Urbana-Champaign*.

Honig, B., and Nicholls, A. 1995. Classical electrostatics in biology and chemistry. *Science 268*, 1144 – 1149.

Humphrey, W., Dalke, A., and Schulten, K. 1996. VMD–visual molecular dynamics. *J. Mol. Graphics 14*, 33–38.

Jackson, J. D. 1975. *Classical Electrodynamics*, second ed. John Wiley and Sons.

Klein, T. E., Huang, C. C., Pettersen, E. F., Couch, G. S., Ferrin, T. E., and Langridge, R. 1990. A real-time malleable molecular surface. *J. Mol. Graphics 8(1)*, 16–24 and 26–27.

Leach, A. R. 2001. *Molecular Modelling: Principles and Applications*, second ed. Prentice Hall.

Lee, B., and Richards, F. M. 1971. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol. 55*, 379 – 400.

Lorensen, W., and Cline, H. 1987. Marching cubes: a high resolution 3D surface construction algorithm. *Computer Graphics (SIGGRAPH '87 Proceedings) 21*, 4 (July), 163–169.

Max, N., Hanrahan, P., and Crawfis, R. 1990. Area and volume coherence for efficient visualization of 3D scalar function. *San Diego Workshop on Volume Visualization, Computer Graphics 24*, 5, 27–33.

McCormick, S. F. 1989. *Multilevel Adaptive Methods for Partial Differential Equations*. Society for Industrial and Applied Mathematics.

Moore, P. K. 1999. Finite difference methods and spatial a posteriori error estimates for solving parabolic equations in three space dimensions on grids with irregular nodes. *SIAM Journal on Numerical Analysis 36:4*, 1044–1064.

Pajarola, R. 2001. Fastmesh: Efficient view-dependent meshing. In *Proceedings Pacific Graphics*, 22–30.

Sanner, M. F., and Olson, A. J. 1997. Real time surface reconstruction for moving molecular fragments. In *Pacific Symposium on Biocomputing '97*, R. B. Altman, A. K. Dunker, L. Hunter, and T. E. Klein, Eds., 385–396.

Sharp, K. A., and Honig, B. 1990. Electrostatic interactions in macromolecules: theory and applications. *Annu. Rev. Biophys. Chem. 19*, 301 – 332.

Sukharev, S., Durell, S. R., and Guy, H. R. 2001. Structural models of the MSCL gating mechanism. *J. Biophys 81(2)*, 917–936.

Tanford, C., and Kirkwood, J. G. 1957. Theory of protein titration curves. I. General equations for impenetrable spheres. *J. Am. Chem. Soc. 79*, 5333 – 5339.

Varshney, A., Brooks, Jr., F. P., and Wright, W. V. 1994. Computing smooth molecular surfaces. *IEEE Computer Graphics & Applications 15*, 5 (September), 19–25.

Warwicker, J., and Watson, H. C. 1982. Calculation of electrostatic potential in the active site cleft due to α-helix dipoles. *J. Mol. Biol. 155*, 53 – 62.