

PUBLICATION CRITIQUE

NATHAN BAKER

1. ELECTROSTATICS OF NANOSYSTEMS: APPLICATION TO MICROTUBULES AND THE RIBOSOME

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences*, **98** (18), 10037-10041, 2001.

Although old, this paper marks the beginning of over 15 years of bringing applied mathematics methods to the biological sciences community in the form of finite element and finite difference software which can be easily employed by mathematics novices to understand the structure and function of biomolecules. This manuscript introduced a new software package, APBS (<http://www.poissonboltzmann.org/>), which is currently used by 25,600+ users worldwide. APBS was developed in collaboration with Prof. Michael Holst using the adaptive multigrid finite element [10] and finite difference libraries that I contributed to as part of his research group. The APBS software package solves the Poisson-Boltzmann [6] and related nonlinear partial differential equations that describe the electrostatic properties of molecules in water. The APBS package has been the focal point of my research on new mathematical models and algorithms for analyzing biomolecular solvation. Examples of innovation in this area include new topology-based methods for comparing proteins based on their electrostatic potential [9]; new descriptions of and metrics for nonpolar solvation in molecular systems [8]; fast methods for finding optimum protein titration states in very large $\mathcal{O}(2^N)$ search spaces, where N is $\mathcal{O}(10 - 100)$ for moderate-size proteins [1]; and ensemble-based classifiers for improving titration state prediction accuracy [2]. Because of the impact of our algorithms and the APBS software, this paper has been cited over 3700 times.

2. DIFFERENTIAL GEOMETRY BASED SOLVATION MODEL I: EULERIAN FORMULATION

Chen Z, Baker NA, Wei GW. Differential geometry based solvation model I: Eulerian formulation. *Journal of Computational Physics*, **229** (22), 8231-8258, 2010.

This paper introduced a new class of solvation models with accuracy improved by a self-consistent treatment that substantially reduced the number of parameters that users must choose and optimize. Traditional continuum models for molecular solvation decouple the polar and nonpolar contributions to the energy; e.g., modeling the polar solvation energy via the Poisson-Boltzmann equation and the nonpolar energy with geometric treatments. The use of these artificially separated treatments has led to confusion and inaccurate models in the solvation community and unnecessarily increased the number of extra parameters that non-experts must determine when applying the models. This paper describes an integrated model which couples my past work in improved nonpolar solvation models [8] with Prof. Guowei Wei's research on Laplace-Beltrami equations to model protein geometry. The resulting models showed excellent accuracy in predicting small molecule solvation energies [3]–[5] and were implemented in both Eulerian (paper above) and Lagrangian solvers [7]. The model described in this and subsequent papers offer improved accuracy and substantially reduce the number of parameters that must be chosen and optimized for biological applications.

Lei H, Yang X, Zheng B, Lin G, Baker NA. Quantifying the influence of conformational uncertainty in biomolecular solvation. *SIAM Multiscale Modeling and Simulation*, in press. <http://arxiv.org/abs/1408.5629>

Whereas the previous paper addressed model uncertainty in biomolecular solvation, this manuscript describes a framework for addressing the numerous sources of parameter uncertainty associated with biomolecular electrostatics and solvation models. This initial paper chooses a set of parameters with well-defined uncertainty: the coordinates provided in X-ray crystal structures accompanied by the B-factors which represent thermally induced (and experimentally limited) uncertainties in atomic positions. Our framework combines generalized polynomial chaos (gPC) methods for uncertainty quantification with two optimization approaches for increasing the efficiency and accuracy of the surrogate model: compressed sensing ℓ_1 minimization is used to solve the under-determined linear system for the gPC basis coefficients while minimizing the number of non-zero basis contributions and the active random subspace method of Constantine and others is used to maximize the sparsity of the measurement matrix. Our approach accurately reconstructs the variation in solvation properties due to conformational uncertainty and, more importantly, lays the foundation for understanding the numerous other sources of uncertainty in solvation models that negatively impact their application in biological and chemical research applications.

REFERENCES

- [1] E. Hogan, K. Monson, and N. A. Baker, “Energy minimization of discrete protein titration state models using graph theory,” *arXiv preprint arXiv:1507.07021*, 2015.
- [2] L. J. Gosink, E. A. Hogan, T. C. Pulsipher, and N. A. Baker, “Bayesian model aggregation for ensemble-based estimates of protein pK_a values,” *Proteins: Structure, Function, and Bioinformatics*, vol. 82, no. 3, pp. 354–363, 2014.
- [3] M. D. Daily, J. Chun, A. Heredia-Langner, G. Wei, and N. A. Baker, “Origin of parameter degeneracy and molecular shape relationships in geometric-flow calculations of solvation free energies,” *The Journal of Chemical Physics*, vol. 139, no. 20, p. 204108, 2013.
- [4] D. G. Thomas, J. Chun, Z. Chen, G. Wei, and N. A. Baker, “Parameterization of a geometric flow implicit solvation model,” *Journal of Computational Chemistry*, vol. 34, no. 8, pp. 687–695, 2013.
- [5] Z. Chen, S. Zhao, J. Chun, D. G. Thomas, N. A. Baker, P. W. Bates, and G. Wei, “Variational approach for nonpolar solvation analysis,” *The Journal of Chemical Physics*, vol. 137, no. 8, p. 084101, 2012.
- [6] P. Ren, J. Chun, D. G. Thomas, M. J. Schnieders, M. Marucho, J. Zhang, and N. A. Baker, “Biomolecular electrostatics and solvation: a computational perspective,” *Quarterly Reviews of Biophysics*, vol. 45, no. 04, pp. 427–491, 2012.
- [7] Z. Chen, N. A. Baker, and G. Wei, “Differential geometry based solvation model II: Lagrangian formulation,” *Journal of Mathematical Biology*, vol. 63, no. 6, pp. 1139–1200, 2011.
- [8] J. A. Wagoner and N. A. Baker, “Assessing implicit models for nonpolar mean solvation forces: the importance of dispersion and volume terms,” *Proceedings of the National Academy of Sciences*, vol. 103, no. 22, pp. 8331–8336, 2006.
- [9] X. Zhang, C. L. Bajaj, B. Kwon, T. J. Dolinsky, J. E. Nielsen, and N. A. Baker, “Application of new multi-resolution methods for the comparison of biomolecular electrostatic properties in the absence of global structural similarity,” *Multiscale Modeling & Simulation*, vol. 5, no. 4, p. 1196, 2006.
- [10] M. Holst, N. Baker, and F. Wang, “Adaptive multilevel finite element solution of the Poisson–Boltzmann equation I. algorithms and examples,” *Journal of Computational Chemistry*, vol. 21, no. 15, pp. 1319–1342, 2000.