John von Neumann Institute for Computing

Comparing Semi-Empirical versus Classic Charge Assignments in BioMolecules and their Effect on Electrostatic Potentials

P. Kar, M. Seel, U. H. E. Hansmann, S. Höfinger

published in

From Computational Biophysics to Systems Biology (CBSB07),
Proceedings of the NIC Workshop 2007,
Ulrich H. E. Hansmann, Jan Meinke, Sandipan Mohanty,
Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **36**, ISBN 978-3-9810843-2-0, pp. 155-158, 2007.

http://www.fz-juelich.de/nic-series/volume36

Comparing Semi-Empirical versus Classic Charge Assignments in BioMolecules and their Effect on Electrostatic Potentials

Parimal Kar[1], Max Seel[1], Ulrich H. E. Hansmann[1,2], and Siegfried Höfinger[1]

[1] Michigan Technological University, Department of Physics,
1400 Townsend Drive, 49931 Houghton, MI, USA
E-mail: [pkar, seel, hansmann, shoefing]@mtu.edu

[2] John von Neumann Institute for Computing, FZ Jülich
52425 Jülich, Germany
E-mail: u.hansmann@fz-juelich.de

The program LocalSCF is used to consider an antifungal protein at the semi-empirical QM level of theory. Model Hamiltonians include AM1, MNDO, PM3 and PM5. The biomolecule is also studied with classic charge assignment using the AMBER force field. An important aspect of classic biomolecular simulation can thus be addressed, namely to what extent the usual concept of a single set of static atomic partial charges per type of amino acid will hold in general for the entire global protein structure. Semi-empirical charges will vary with different chemical neighborhood inside the protein and the question remains how severely these alterations will affect global electrostatic properties of the protein. In order to probe this effect we use grid maps of electrostatic potentials obtained from solutions to the Poisson-Boltzmann equation. Source charges are either the classic AMBER ones, or some set of semi-empirical charges from the list of models mentioned above. In comparing different potential maps we aim to recognize systematic trends as well as to identify a recommended way of proper charge assignment in proteins.

1 Introduction

Electrostatics plays an integral part in the study of structure and function of proteins at physiological conditions[1]. Theoretical considerations of the electrostatics in proteins are usually based on solutions to the Poisson-Boltzmann (PB) equation[2,3]. All these theoretical descriptions will involve a certain type of charge assignment to the atoms of the protein. Since the result of the PB calculation will inevitably depend on the particular choice made for the charges, it might be of interest to study the influence and variation resulting from different charge assignments. Of particular interest will be the comparison between a set of classic charges, ie from force fields commonly employed in the simulation of biomolecules, and charges derived from ab-inito calculations performed at a certain level of Quantum Mechanical (QM) theory.

A convenient method to compare different charge assignments to each other is to study the shape and appearance of electrostatic potential (ESP) maps. These ESP maps describe the way the protein will represent itself to its environment in electrostatic terms. Since the solution to the PB equation is included, ESP maps render a reasonably complete picture of the protein in its native environment, ie at physiological conditions. Moreover, ESP maps are a useful tool with many direct applications in structural biology. For example, from ESP maps we can learn whether a protein, (i) is likely to migrate to the membrane[4],

(ii) will potentially bind RNA or DNA[5,6], (iii) belongs to a certain family[7–9], (iv) offers a chemically attractive binding site to ligands and other proteins.

In this present study we therefore compare ESP maps based on classic charge assignments using AMBER parameters[10] with ESP maps resulting from semi-empirical charges computed with program LocalSCF[11] at several levels of semi-empirical theory, ie AM1, MNDO, PM3 and PM5. The PB program POLCH[12] is used throughout.

2 Methods

After download of the protein with pdb code EAFP2 from the pdb data bank, a PB calculation is performed using program POLCH[12] and classic AMBER partial charges[10]. Inner/outer dielectric constants are set to 1 and 80 respectively. The net charge is +4 due to the four Arg residues. ESP maps are computed on the molecular surface and on a cubic grid superimposing the protein. Only ESP maps directly mapped onto the molecular surface are used for further analysis. Semi-empirical calculations are then carried out on the protein EAFP2 using LocalSCF[11] and finally computed partial charges are extracted from the output. The net charge is +2 due to different treatment of lone-pairs in the semi-empirical models. AM1, MNDO, PM3 and PM5 methods are applied. Classic AMBER partial charges are then replaced with either charge set derived from the semi-empirical calculations and PB calculations are repeated with the changed charge assignment. Resulting ESP maps are compared in the form of difference ESP maps.

3 Results and Conclusions

A structural sketch of the antifungal protein EAFP2 is shown in Table 1 (a) with corresponding representation of the molecular surface (b). Here the N-terminal end is colored in red while the C-terminus is given in blue. The ESP map based on classic AMBER charge assignment after PB calculation is represented in Table 1 (c). ESP levels are color-coded as +5 kT/e (blue), 0 kT/e (green) and -5 kT/e (red). It becomes clear that the major appearance of EAFP2 in aqueous solution is that of a macroscopic particle of largely positive ESP, hence the tendency to migrate to the membrane can be explained straightforwardly[4] (which also implies the antifungal mode of action). An initial test regarding the sensitivity to counter ions is shown in Table 1 (d). Here explicit Cl^- counter ions have been included in the PB calculation and corresponding ESP maps produced. The change in major ESP patterns introduced by counter ions is only marginal, thus the rest of the analysis is performed without consideration of counter ions. A differential ESP map representing ESP(AM1) - ESP(AMBER) is shown in Table 1 (e). Identical color-coding is used as mentioned above. It becomes clear that the AM1-based ESP map is comparable in sign, but significantly different in magnitude (individual ESP values have become less positive). Extended red patches mark off regions of most severe difference. Contrary to the change seen in the AM1-AMBER differential map, when comparing AM1 with MNDO we obtain essentially only green patches (see Table 1 (f)). Thus AM1 and MNDO deliver essentially the same ESP properties. Comparison of PM3 with AMBER is represented in the differential ESP map shown in Table 1 (g).

The trend is similar to the one seen with AM1, but the difference is less severely pronounced (ie certain extended red regions turn yellow or green). Switching further to PM5

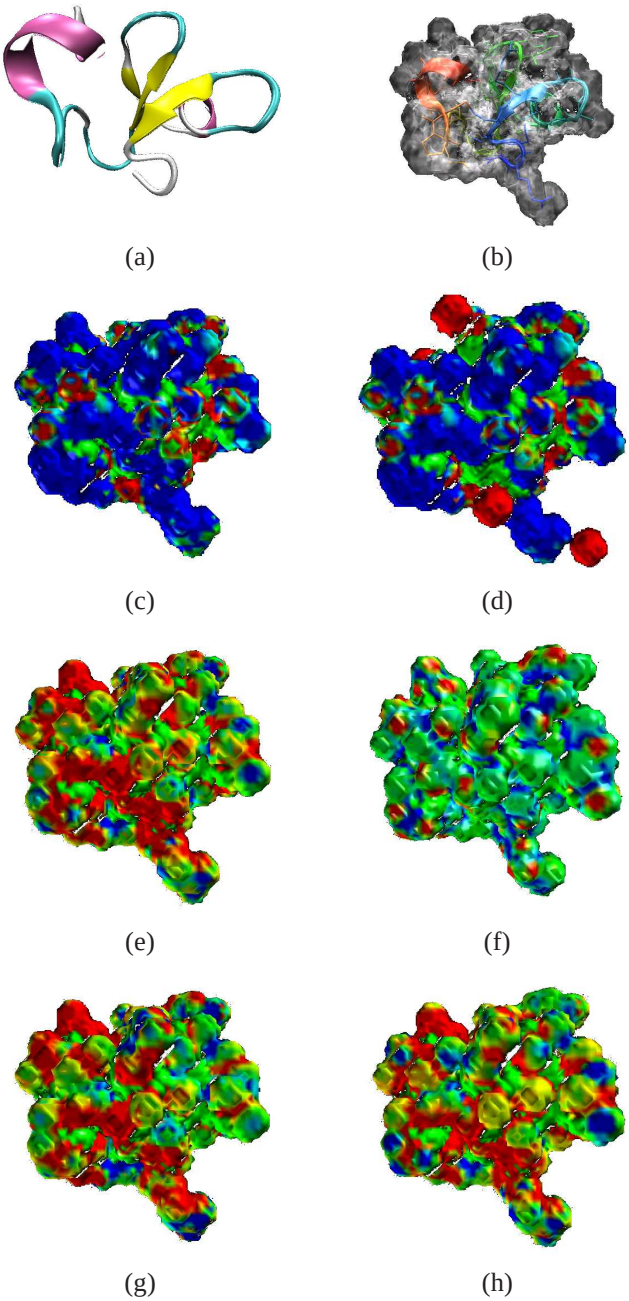

Table 1. Electrostatic Potential (ESP) maps for the antifungal protein EAFP2 (pdb code). Major structural elements are shown in (a) and a corresponding representation of the molecular surface is shown in (b), where the N-terminal helix is given in red and the C-terminus is shown in blue. The ESP mapped onto the molecular surface after solution of the PB equation based on AMBER charge assignment is shown in (c). Blue patches correspond to the +5 kT/e level, green regions represent neutral ESP and red domains indicate negative ESP of -5 kT/e. The marginal change when including 4 explicit Cl^- counter ions is shown in (d). A differential ESP map representing the difference between ESP(AM1) and ESP(AMBER) is shown in (e) with the same color-coding scheme used in (c). Further differential maps are ESP(AM1)-ESP(MNDO) (f), ESP(PM3)-ESP(AMBER) (g) and ESP(PM5)-ESP(AMBER) (h).

description is continuing the trend, ie lessening the deviation from the AMBER-based map again (see Table 1 (h)). Closer examination of the residues lying beneath the red-colored patches (indicating most severe deviation) reveals a specific role of Arg residues and the charges assigned to the N-atoms of Asn and Gln. In summary, semi-empirical charge assignments deliver a consistent picture of significant differences seen for the charged residues. However, individual semi-empirical models differ considerably amongst each other. With increasing sophistication of the semi-empirical model the deviation from the classic AMBER results becomes less severe.

Acknowledgments

This work was supported in part by the National Institutes of Health Grant GM62838.

References

1. B. Honig, A. Nicholls, *Classical electrostatics in biology and chemistry*, Science **268**, 1144–1149, 1995.
2. J. Warwicker, H. C. Watson, *Calculation of the electric potential in the active site cleft due to alpha-helix dipoles*, J. Mol. Biol. **157**, 671–679, 1982.
3. R. J. Zauhar, R. S. Morgan, *A new method for computing the macromolecular electric potential*, J. Mol. Biol. **186**, 815–820, 1985.
4. Y. Xiang, R. H. Huang, X. Z. Liu, Y. Zhang, D. C. Wang, *Crystal structure of a novel antifungal protein distinct with five disulfide bridges from Eucommia ulmoides Oliver at an atomic resolution* , J. Struct. Biol. **148**, 86–97, 2004.
5. S. S. Kim, R. g. Zhang, S. E. Braunstein, A. Joachimiak, A. Cvekl, R. S. Hegde, *Structure of the Retinal Determination Protein Dachshund Reveals a DNA Binding Motif*, Structure **10**, 787–795, 2002.
6. Q. Ye, R. M. Krug, Y. J. Tao, *The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA*, Nature **444**, 1078–1082, 2006.
7. J. W. Yu, J. M. Mendrola, A. Audhya, S. Singh, D. Keleti, D. B. deWald, D. Murray, S. D. Emr, M. A. Lemon, *Genome-Wide Analysis of Membrane Targeting by S. cerevisiae Pleckstrin Homology Domains*, Mol. Cell **13**, 677–688, 2004.
8. D. Morikis, J. D. Lambrish, *Physical methods for structure, dynamics and binding in immunological research*, Trends Immunol. **25**, 700–707, 2004.
9. K. Schleinkofer, U. Wiedemann, L. Otte, T. Wang, G. Krause, H. Oschkinat, R. C. Wade, *Comparative Structural and Energetic Analysis of WW DomainPeptide Interactions*, J. Mol. Biol. **344**, 865-881, 2004.
10. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules*, J. Am. Chem. Soc. **117**, 5179–5197, 1995.
11. N. A. Anikin, V. M. Anisimov, V. L. Bugaenko, V. V. Bobrikov, A. M. Andreyev, *LocalSCF Method for Semi-empirical Quantum-Chemical Calculation of Ultra-large Bio-molecules*, J. Chem. Phys. **121**, 1266–1270, 2004.
12. S. Höfinger, *Solving the Poisson-Boltzmann Equation with the Specialized Computer Chip MD-GRAPE-2*, J. Comp. Chem. **26**, 1148–1154, 2005.