Design of a reduced point charge model for proteins: Applications to molecular electrostatic potential and solvation energy calculations

Leherte, Laurence; Vercauteren, Daniel

Publication date:
2010

Document Version
Peer reviewed version

Design of a reduced point charge model for proteins: Applications to molecular electrostatic potential and solvation energy calculations

Laurence Leherte, Daniel P. Vercauteren - University of Namur - Belgium (FUNDP)

Abstract

To model and simulate large molecular systems like proteins, the development of reduced representations is currently an active field of research.
We present an approach to design and evaluate reduced point charge models obtained from smoothed electron density distribution functions.
The location of charges is determined through a merging/clustering procedure of atom trajectories generated in EDs built from smoothed Amber99 electrostatic potential functions.
A fitting algorithm is applied to evaluate the charge values. Radii are also evaluated.
Applications to proteins are presented.

Contact

Laurence Leherte
Laboratoire de Physico-Chimie Informatique,
Unité de Chimie Physique Théorique et Structurale,
University of Namur (FUNDP)
Belgium
laurence.leherte@fundp.ac.be

Oral Presentation

Area: (7) Biological Chemistry
Session Title: Protein, Peptide, and Peptidomimetics Design (#149) [6E]
December 18, 7:00-9:00 PM
Location: 304B (Convention Center)

Introduction

In a previous work [1], we developed an approach to locate extrema in a smoothed molecular electrostatic potential (MEP) $\Phi_{A,t}$, and applied it to the design of amino acid reduced point charge models. Such models were then used for the evaluation of the MEP of KcsA, a potassium ion channel structure. While efficient in approximating the corresponding all-atom MEP, the models however present two drawbacks: the points are located away from the skeleton of the structure, and their location and charge value may be dependent on the amino acid conformation.
In this poster, the approach is modified so as to determine the point charge centers from electron density (ED) distribution functions $\rho_{A,t}$ obtained using the Poisson equation.

Method

1. Molecular Electrostatic Potential and Electron Density

$$V_A(\vec{r}) = \sum_{a \in A} \frac{q_a}{\left|\vec{r} - \vec{R}_a\right|}$$

Unsmoothed MEP generated by molecule A. Amber99 atomic charges [2] are assigned to atoms using PDB2PQR [3].

$$V_{A,t}(\vec{r}) = \sum_{a \in A} \frac{q_a}{\left|\vec{r} - \vec{R}_a\right|} erf\left(\frac{\left|\vec{r} - \vec{R}_a\right|}{2\sqrt{t}}\right)$$

Smoothed MEP [4]

$$-\nabla^2 V_{A,t} = \frac{\rho_{A,t}}{\varepsilon_0}$$

Poisson Equation → Smoothed ED $\rho_{A,t}$
(t = smoothing degree)

2. Location of Coarse Grains in $\rho_{A,t}$

A hierarchical merging algorithm, based on the idea of Leung *et al.* [5], is used to locate local extrema in $\rho_{A,t}$.
• At scale t = 0, each atom of a molecular structure is considered as a starting point of the merging procedure.
• As t increases, each point moves along a gradient path to reach a location in the 3D space where: $\vec{\nabla}\rho_{A,t} = 0$

These trajectories are defined by:

$$\vec{r}_{\rho_{A,t}} = \vec{r}_{\rho_{A,t-\Delta t}} + \frac{\Delta}{\rho_{A,t}} \vec{\nabla}\rho_{A,t}$$

Δ = displacement step

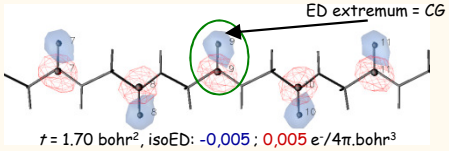
3. Charge and Radius Fitting

• **Charges** - achieved through the program QFIT [6] to get *CG* point charges fitted from an unsmoothed MEP grid, considering various rotamers with the following constraints: molecular charge and dipole.
• **Radii** - achieved through the program APBS [7] to get *CG* radii adjusted to approach the mean all-atom solvation energy obtained for various rotamers.

4. Backbone *CGs*

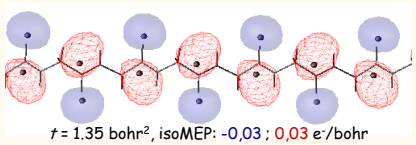
a. An extended strand ß-Gly_{15} is built considering (Ω = 180°, Φ = -139°, Ψ = 135°) using the program SMMP05 [8].

b. The merging/clustering program is applied separately to negative and positive charges of ß-Gly_{15}. The central motif is isolated and further used as template for the backbone of any amino acid residue.



t = 1.70 bohr², isoED: -0,005 ; 0,005 e⁻/4π.bohr³

Advantages: *CGs* are located close to the molecular structure, and are not significantly dependent on the conformation, as opposed to our previous work [1]:



t = 1.35 bohr², isoMEP: -0,03 ; 0,03 e⁻/bohr

5. Side Chain *CGs*

a. Each isolated amino acid is considered with various conformations [9] built using SMMP05 [8]. For instance:

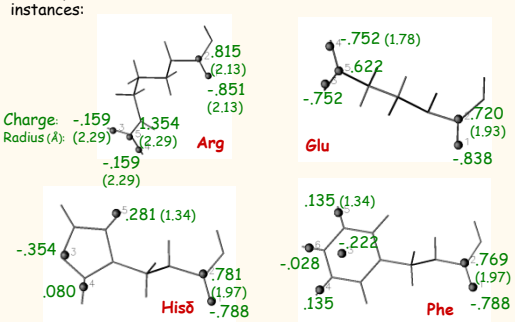
	Conformation	χ1 (°)	χ2 (°)	χ3 (°)	χ4 (°)	Occurrence (%)
Arg	*g-, t, g-, g-*	300	180	300	300	9.5
	g-, t, g-, t	300	180	300	180	11.9
	g-, t, g+, t	300	180	60	180	12.2
	g-, t, t, t	300	180	180	180	12.2

b. Atomic charges are assigned using PDB2PQR [3].

c. *CGs* are obtained using the merging/clustering program.

d. Charges and radii are assigned in two steps. First, side chain *CGs* are treated. Second, backbone charges and radii are adjusted while keeping the side chain charges/radii to their pre-determined values.

e. Templates are thus obtained for each amino acid residue. For instances:

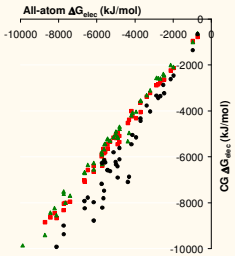


Applications

1. Set of 53 Protein Structures [10]

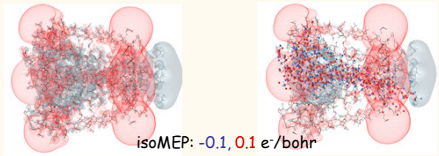
• Positioning of *CG* points is achieved through QUATFIT [11], a superposition algorithm, using the amino acid templates and the PDB structures of the proteins.
• End positive and negative charges (± 0.929 e⁻) are added on terminal N and OXT atoms.

	Means *vs.* all-atom values
rmsd V (kJ/mol)	25.2 ± 3.7
rmsd μ (D) (%)	18.8 ± 11.3 6.2 ± 6.6
rmsd $\Delta G_{aams\ (1)}$ (kJ/mol) (%)	31.1 ± 25.7 3.9 ± 3.6
rmsd $\Delta G_{cgms\ (2)}$ (kJ/mol) (%)	27.2 ± 20.2 0.4 ± 5.7



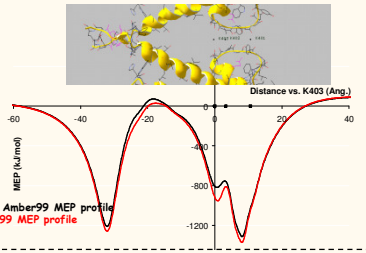
1. all-atom molecular surface
2. corrected (+ 10%) coarse-grain molecular surface (non corrected values are in black on the Figure)

2. KcsA Channel (1BL8.pdb)



All-atom model
5888 atoms, q_{tot} = +4 e⁻
μ = 1348.4 , 512.3 , 273.7 D
ΔG_{elec}= -11.2 10^3 kJ/mol

***CG* model**
1328 *CGs*
μ = 1336.3 , 512.4 , 244.0 D (rmsd μ = 2.2 %)
rmsd V = 2.7 kJ/mol
ΔG_{elec}= -11.8 10^3 kJ/mol (rmsd ΔG_{aams} = 4.9 %)



Conclusions

A reduced point charge model for proteins, built from smoothed EDs, allows to approach all-atom electrostatic properties (MEP, dipole moment, ΔG_{elec}). Transferability of amino acid templates is based on: templates built for isolated residues (no secondary structure dependency), separate treatment of backbone and side chain charges, reduced conformational dependency of *CG* locations in ED distributions.

(1) Leherte & Vercauteren, J. Chem. Theory Comput. 5 (2009) 3279 ; **(2)** Wang *et al.*, J. Comput. Chem. 21 (2000) 1049 ; **(3)** pdb2pqr.sourceforge.net ; **(4)** Amara & Straub, Phys. Rev. B 53 (1996) 13857 ; **(5)** Leung *et al.* IEEE T. Pattern Anal. 22 (2000) 1396 ; **(6)** Borodin *et al.* Force Field Fitting Toolkit, www.eng.utah.edu/~gdsmith/fff.html ; **(7)** Baker *et al.*, Proc. Natl. Acad. Sci. USA 98 (2001) 10037, apbs.sourceforge.net ; **(8)** Eisenmenger *et al.* Comp. Phys. Comm. 174 (2006) 422, www.smmp05.net ; **(9)** Simms *et al.*, Prot. Eng. Des. Select. 21 (2008) 369, www.dynameomics.org ; **(10)** Tjong & Zhou, J. Chem. Theory Comput. 4 (2008) 507 ; **(11)** Heisterberg, Ohio Supercomputer Center, translation from FORTRAN to *C* and Input/Output by Labanowski, Ohio Supercomputer Center, 1990.