Design of a reduced point charge model for proteins – Molecular Dynamics applications

Leherte, Laurence; Vercauteren, Daniel

Publication date:
2012

Document Version
Peer reviewed version

Design of a reduced point charge model for proteins: Molecular Dynamics Applications

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

Framework

This work is part of a more general project on multiresolution analysis of three-dimensional *(3D)* molecular fields. More specifically, it is related to the topographical analysis of smoothed charge density *(CD)* distribution functions. From such functions, we show how to obtain a reduced point charge model (also named here coarse grain *(CG)* model) for proteins and we assess the applicability to molecular dynamics *(MD)* simulations of protein structures.

Reduced Point Charge Model

1. Smoothed Charge Density

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

Unsmoothed electrostatic Coulomb potential generated by molecule A. Amber99 atomic charges [1] are assigned to atoms using PDB2PQR [2].

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

Smoothed electrostatic potential [3]

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

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

$$\rho_{A,t} = \frac{q_a}{(4\pi)^{3/2}} e^{-r^2/4t}$$

Analytical expression of $\rho_{A,t}$

2. Location of CGs 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. Backbone CGs

a. An extended strand β-Gly_{15} is built considering (Ω = 180°, Φ = -139°, Ψ = 135°) using the program SMMP05 [5].

b. Atomic charges are assigned using PDB2PQR [2].

c. The merging/clustering program is then applied separately to negative and positive charges of β-Gly_{15}. The central motif is isolated and further used as template for the backbone of any amino acid residue (see Figure below).

4. Side Chain CGs

a. Each isolated amino acid is considered with various conformations [6] built using SMMP05 [5].

b. Atomic charges are then assigned using PDB2PQR [2].

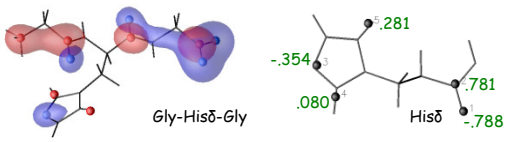
c. CGs are finally obtained using the merging/clustering program.

5. Charge Fitting

It is achieved with the program QFIT [7] to get CG point charges fitted from an unsmoothed electrostatic potential grid, considering the various rotamers with two constraints: molecular charge and dipole.

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

Templates are thus obtained for each amino acid residue [8]. For instance:



Isocontours of the CD: -0.005, +0.005 e⁻ bohr⁻³

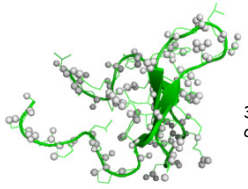
Références

(1) Wang *et al.*, J Comput Chem 21 (2000) 1049 **(2)** pdb2pqr.sourceforge.net **(3)** Amara & Straub, Phys Rev B 53 (1996) 13857 **(4)** Leung *et al.* IEEE T Pattern Anal 22 (2000) 1396 **(5)** Eisenmenger *et al.* Comp Phys Comm 174 (2006) 422, www.smmp05.net **(6)** Simms *et al.*, Prot Eng Des Select 21 (2008) 369, www.dynameomics.org **(7)** Borodin *et al.* Force Field Fitting Toolkit, www.eng.utah.edu/~gdsmith/fff.html **(8)** Leherte & Vercauteren, J Comput Aided Mol Des 25 (2011) 913 **(9)** Heisterberg, Ohio Supercomputer Center, translation from FORTRAN to C and Input/Output by Labanowski, 1990 **(10)** Leherte & Vercauteren, J Phys Chem A 115 (2011) 12531 **(11)** TINKER - Software Tools for Molecular Design, http://dasher.wustl.edu/tinker/

Molecular Dynamics Applications

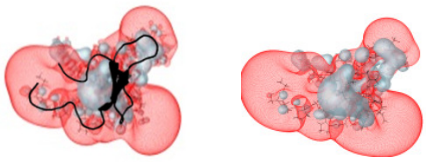
1. Ubiquitin protein-ligase Nedd4-2 (PDB access code 1WR3)

- The positioning of CGs is achieved with QUATFIT [9], 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.



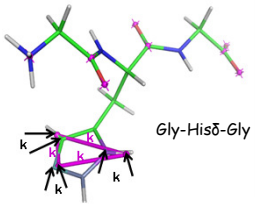
36 aa / 577 atoms / 158 point charges or CGs (white spheres) [10]

2. Isocontours of the electrostatic potential (-0.05, 0.05 e⁻/bohr)



All-atom CG

3. Implementation

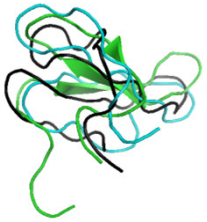


Gly-Hisδ-Gly

The new set of charges was implemented in the program TINKER [11]. The fact that the point charges are not necessarily located on atoms makes the implementation challenging and, as a first trial, the original code was not modified. Thus, the charges are treated as additional particles (m = 2) held together and to heavy atoms of the molecular structure by harmonic bonds [10].

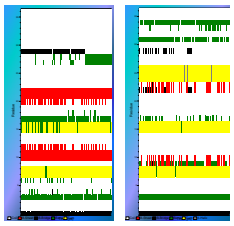
4. Results

10 ns (10^7 x 1 fs)
T = 298 K
Vacuum
CPU time AA/CG = 2



Snapshots at 10 ns
All-atom
CG

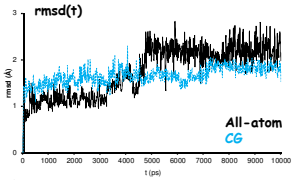
Cristal structure
PDB

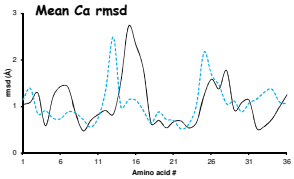


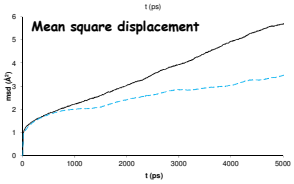
All-atom CG

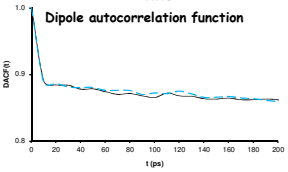
DSSP secondary structures: Only the main features of the secondary structure are preserved in the CG simulation.

Coil
β-sheet
β-bridge
Bend
Turn









Conclusions

CG locations and number are determined only from the topography of the CD distribution functions.
Charge centers differ from atom centers → challenge for MD applications.
Some secondary structure elements are preserved *vs.* the all-atom model as well as the global fold of the peptide.
Dynamical results are in agreement with the all-atom ones, but the mean square displacement is reduced.
One also observes an increased conformational stability of the CG model.
As a perspective, we are considering the instantaneous update of the CG model during the MD simulation (to avoid to assign a mass to the CGs).