Design of reduced point charge models for proteins

Leherte, Laurence; Vercauteren, Daniel

Publication date:
2015

Document Version
Peer reviewed version

Design of reduced point charge models for proteins

Laurence Leherte, Daniel P. Vercauteren
University of Namur – Belgium
TheoBioChem 2015 – Cagliari

UNIVERSITÉ DE NAMUR
nomedic NAmur MEdicine & Drug Innovation Center
Laboratoire de Physico-Chimie Informatique

Introduction

Reduced point charge models (RPCMs) for proteins are obtained from topological analyses of smoothed charge density (CD) distribution functions. For each amino acid, the RPCMs involve two backbone charges and up to six charges on the side chain. RPCM-based molecular dynamics (MD) trajectories are compared to all-atom ones for Ubiquitin-based systems (1UBQ, 1Q0W).

1. Method

1. Smoothing of the Coulomb potential

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

Unsmoothed molecular electrostatic potential (MEP). Amber99 atomic charges [1] are assigned to atoms using PDB2PQR [2].

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

Smoothed MEP [3]

The Poisson equation is applied to generate the corresponding smoothed atomic charge density (CD) distribution function, $\rho_{a,s}$:

$$-\nabla^2 V_{a,s}(\mathbf{r}) = \frac{\rho_{a,s}}{\varepsilon_0} \longrightarrow \frac{\rho_{a,s}}{\varepsilon_0} = \frac{q_a}{(4\pi s)^{3/2}} e^{-r^2/4s}$$

2. Example

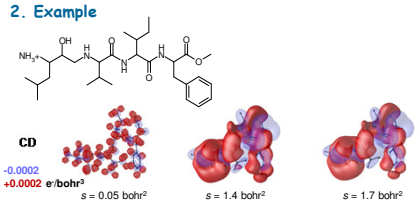


Fig. 1 Smoothed CD of a peptide-like molecule.

3. Location of critical points (CP) in $\rho_{A,s}$

A hierarchical merging algorithm, based on the idea of Leung *et al.* [4], is used to locate local extrema in $\rho_{A,s}$.

- At scale $s \sim 0$, each atom of a molecular structure is considered as a starting point of the merging procedure.
- As s increases, each point moves along a gradient path to reach a location in the 3D space where:

$$|\nabla \rho_{CP}(s)| = 0$$

These trajectories are defined by:

$$\mathbf{r}_{CP}(s) = \mathbf{r}_{CP}(s - \Delta s) + \frac{\Delta}{\rho_{CP}(s - \Delta s)} \nabla \rho_{CP}(s)$$

Δ = displacement step

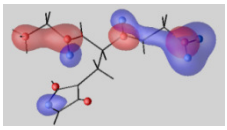


Fig. 2 Isocontours of the CD of Gly-Hisδ-Gly smoothed at s = 1.7 bohr².
-0.005
+0.005 e⁻/bohr³

4. Charge fitting

Charges are fitted either to unsmoothed Amber99 MEPs or MEFs [5],
- considering various amino acid rotamers [6],
- with constraints: total electric charge & total dipole moment.

Side chain charges are first assigned [7,8], then backbone charges are fitted using the side chain charge values as constraints.

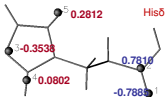


Fig. 3 Reduced point charge model based on the local extrema or critical points (CP) of the Hisδ charge density.

Templates are obtained for each amino acid residue. A second model, named 'AT', is similar to the original 'CP' but most of the point charges are now forced to be located on atoms.

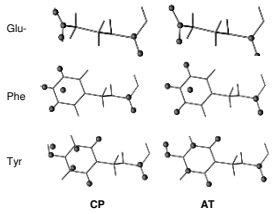


Fig. 4 Reduced point charge models of three amino acid residues.

5. Effect of fitting conditions on charges and forces

	Charges fitted to ...	Range of charge values (|e⁻|)	Absolute charge values of the main chain (|e⁻|)	RMSD vs. 1.4-10.0 Å all-atom forces (kcal/mol.Å)	RMSD vs. 1.0-1.4 Å all-atom forces (kcal/mol.Å)
CP_V	MEP	-0.85 - 1.35	0.77 ± 0.09	1.28	6.90
CP_F	MEF	-0.80 - 1.03	0.69 ± 0.08	1.32	6.36
AT_V	MEP	-0.81 - 1.03	0.73 ± 0.09	1.38	7.05
AT_F	MEF	-0.76 - 1.03	0.64 ± 0.07	1.41	6.37

Charges fitted to forces allows to better approximate short-range forces [5].

6. Molecular electrostatic potential

Ubiquitin (1UBQ.pdb) - 76 amino acids - 1231 atoms

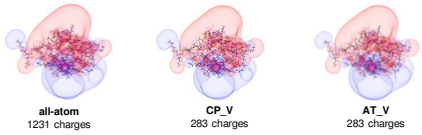
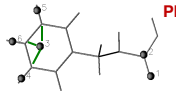


Fig. 5 MEP isocontours of Ubiquitin: -0.05, +0.05 e⁻/bohr.

2. Molecular dynamics applications

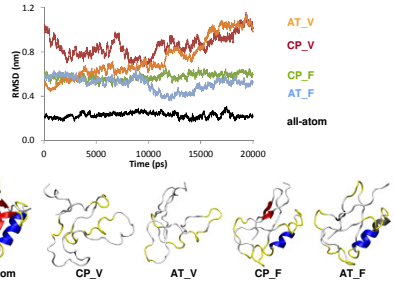
1. Simulation conditions

Gromacs 4.5.5 [9]
Amber99SB and TIP4P-Ew force fields, PME
All force field terms are preserved except the # of protein charges → Cb_{-14} energy values and forces are strongly modified
Non-atomic point charges = virtual sites defined *vs.* selected atoms



Equilibration : 40 ns
Production : 20 ns
NPT (1 bar, 300 K)

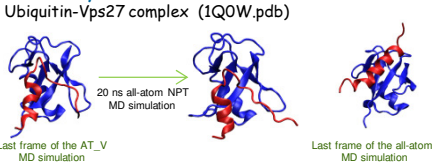
2. RMSD and final snapshots at 300 K (Ubiquitin)



The increased RMSD values reflect a deconstruction of the protein structure, especially with CP_V and AT_V sets of charges.

3. Stability of deconstructed conformations

Ubiquitin-Vps27 complex (1Q0W.pdb)



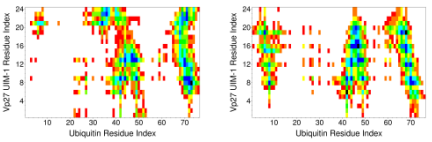


Fig. 6 Shortest contact map Vps27 UIM-1 – Ubiquitin (1Q0W). (left) AT_V MD simulation, (right) all-atom MD simulation.

RPCMs allow to generate deformed but stable protein conformations.

4. Intra-molecular H bonds

Distributions are strongly affected with a RPCM.

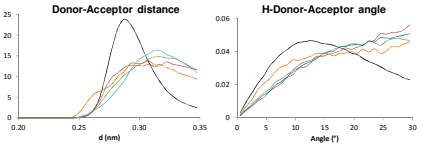


Fig. 7 Distance and angle distributions of intra-molecular H-bonds calculated for solvated Ubiquitin (1UBQ).

5. Protein-water interface

Distance and angle distributions present trends similar to the all-atom case.

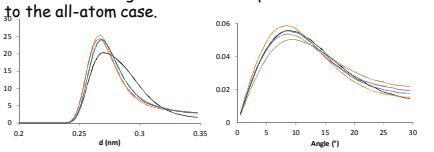


Fig. 8 Distance and angle distributions of Ubiquitin-water H-bonds (1UBQ).

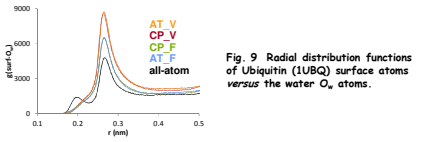


Fig. 9 Radial distribution functions of Ubiquitin (1UBQ) surface atoms *versus* the water O_w atoms.

Table 1. Mean numbers N of water molecules located in a layer of thickness 0.35 nm from the Ubiquitin surface and their self-diffusion coefficient D

	all-atom	CP_V	CP_F	AT_V	AT_F
N	360	564	576	445	435
# H-bonds	198	254	242	205	204
D (10^{-5} cm²/s)	2.31	2.06	2.04	2.15	2.22

The first shell of H_2O molecules is unstructured and appears to be more compact. The dynamics is slower.

Conclusions

- RPCMs allow the approximation of the MEP of rigid proteins. They also allow simulations of flexible structures by MD provided they involve a good description of the short range Coulomb energy terms.
- Charges fitted to electrostatic forces allow a better approximation of the short-range forces.
- Charges located on atoms allow a better approximation of the Cb_{14} energy terms.
- Secondary structure elements can be destructed due, notably, to a loss in the number of H-bonds. It allows the sampling of new conformations that can be stable under all-atom MD conditions.
- RPCMs involve modifications of the interfacial water structure and dynamics.

[1] Duan *et al.* J. Comput. Chem. 24 (2003) 1999; [2] pdb2pqr.sourceforge.net; [3] Hart *et al.* J. Comput. Chem. 21 (2000) 531; [4] Leung *et al.* IEEE T. Pattern Anal. 22 (2000) 1396; [5] Leherte, Mol. Simul. (in press); [6] Simms *et al.* Prot. Eng. Des. Select. 21 (2008) 369, www.dynameomics.org; [7] Leherte *et al.* J. Chem. Theory Comput. 5 (2009) 3279; [8] Leherte *et al.* J. Comput.-Aided Mol. Des. 25 (2011) 913; [9] Pronk *et al.* Bioinformatics 29 (2013) 845

Acknowledgments F. Wautelet and L. Demelenne for program installation and maintenance; The Plateforme Technologique de Calcul Intensif (PTCI) located at the University of Namur, Belgium, supported by the F.R.S.-FNRS