Design and applications of reduced point charge models of 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

EuCO-CC 2015 – Fulda

UNIVERSITÉ DE NAMUR

nomedic — NAmur MEdicine & Drug Innovation Center

Laboratoire de Physico-Chimie Informatique

Introduction

Reduced point charge models (RPCMs) of 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. Using MD simulations, RPCM-based representations of Ubiquitin systems (1UBQ, 1Q0W, 2MBB) allow to generate destructured protein conformations. In particular, deconstructed protein-ligand conformations appear to be stable under all-atom MD simulation conditions.

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

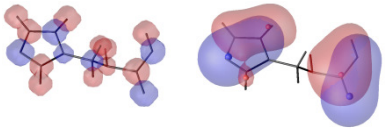


Fig. 1 Isocontours (-0.002 ; +0.005 e⁻/bohr³) and extrema (CP) of the CD of Hisδ smoothed at (left) s = 0.05 and (right) s = 1.7 bohr²

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

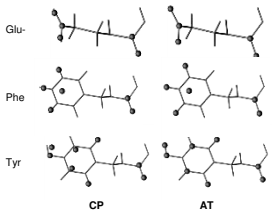


Fig. 2 Reduced point charge models of three amino acid residues.
CP: charges located at CD extrema
AT: charges located on atoms

3. Charge fitting

Charges are fitted to unsmoothed Amber99 molecular electrostatic potential (MEP) or force (MEF) grids [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.

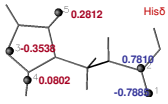
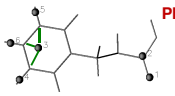


Fig. 3 Reduced point charge model based on the local extrema of the Hisδ charge density with charge values fitted to the all-atom MEP (abbreviated CP_MEP model)

2. Molecular dynamics applications to Ubiquitin

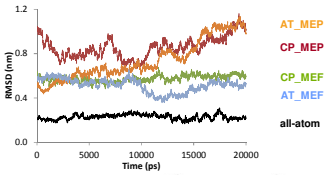
1. Simulation conditions

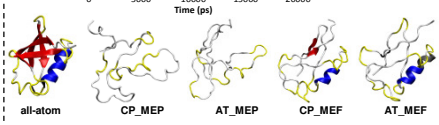
Program: Gromacs 4.5.5 [9]
Force fields: Amber99SB and TIP4P-Ew, PME
All force field terms are preserved except for the number of protein charges → Cb_{14} and Cb short range (SR) energy values and forces are strongly modified.
Non-atomic point charges = virtual sites defined *vs.* selected atoms



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

2. RMSD and final snapshots at 300 K





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

3. Intra-molecular H bonds

Distributions are strongly affected when using a RPCM.

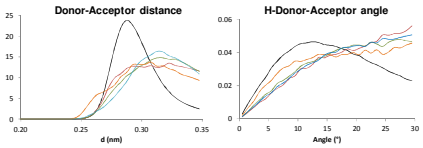


Fig. 4 Distance and angle distributions of intra-molecular H-bonds calculated for solvated Ubiquitin at 300 K

4. Protein-water interface

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

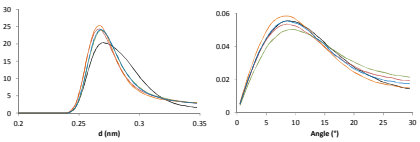


Fig. 5 Distance and angle distributions of Ubiquitin-water H-bonds

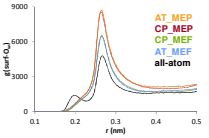


Fig. 6 Radial distribution functions of Ubiquitin surface atoms *versus* the water O_w atoms

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 calculated from 20 ns MD simulations

Charge model	all-atom	CP_MEP	CP_MEF	AT_MEP	AT_MEF
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.

3. Molecular dynamics applications to Ubiquitin-ligand complexes

5. Stability of deconstructed conformations

RPCMs allow to generate deformed but stable conformations

- Ubiquitin-Vps27 complex (1Q0W.pdb)

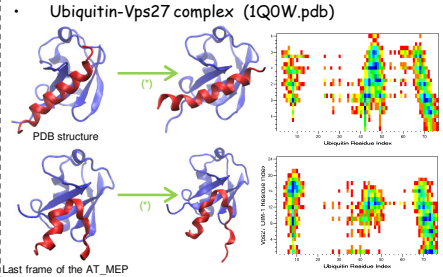


Fig. 7 Shortest contact maps of Vps27 UIM-1 – Ubiquitin calculated from 120 ns all-atom MD simulations at 300 K (top) native, (right) destructured conformation

- Ubiquitin-iota UBM1 complex (2MBB.pdb)

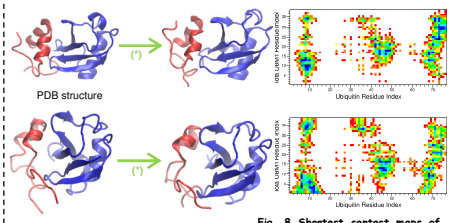


Fig. 8 Shortest contact maps of iota UBM1 – Ubiquitin calculated from 120 ns all-atom MD simulations at 300 K (top) native, (right) destructured conformation

Table 2. Contributions to the potential energy (kJ.mol⁻¹) of systems 1Q0W and 2MBB averaged over 100 ns all-atom MD trajectories

	1Q0W		2MBB	
	Native	Deconstructed	Native	Deconstructed
Total (excl. water, ions, Cb LR)	17,088 ± 189	17,144 ± 201	21,322 ± 202	21,352 ± 239
Protein-ligand (excl. Cb LR)	-474 ± 73	-301 ± 100	-617 ± 98	-725 ± 99
Protein-solvent (excl. Cb LR)	-4,238 ± 200	-4,555 ± 220	-5,425 ± 205	-5,634 ± 221
Ligand-solvent (excl. Cb LR)	-8,913 ± 258	-8,967 ± 266	-8,634 ± 260	-8,580 ± 356

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.
- RPCMs involve modifications of the interfacial water structure and dynamics.
- Secondary structure elements can be destructured due, notably, to a loss in the number of H-bonds. It allows the sampling of new conformations that are stable under all-atom MD conditions. It can be due to stabilized protein-solvent and/or protein-ligand interactions. No unique trend is observed so far.

[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. (2015); **[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