# Combining an Elastic Network With a Coarse-Grained Molecular Force Field

Periole, Xavier; Cavalli, Marco; Marrink, Siewert-Jan; Ceruso, Marco A.

*Published in:*
Journal of Chemical Theory and Computation

*DOI:*
10.1021/ct9002114



*Publication date:*
2009

Supporting Information for:

# Combining an Elastic Network With a Coarse-Grained Molecular Force Field: structure, dynamics and intermolecular recognition

Xavier Periole, Marco Cavalli, Siewert-Jan Marrink and Marco A. Ceruso*

Email: mceruso@sci.ccny.cuny.edu

## 1. Parameterization of bonded terms in ELNEDIN models.

Modifications with respect to the MARTINI-2.1[1,2] force field were necessary since the backbone bead of a residue in ELNEDIN is placed at the location of the Cα atom rather than at the location of the center of mass (COM) of the backbone atoms (N, Cα, C, O).

The motivation behind this change was the removal of the dependence of the sequential backbone-backbone distance on the local secondary structure. This is illustrated in Figure S1, which shows that when placing the backbone bead at the location of the Cα the distribution of sequential backbone-backbone distances has a unique peak centered at 3.80 Å independently of the secondary structure of the residues involved. The distributions were computed using a culled set of PDB structures[3] (cullpdb_pc20_res1.6_R0.25_d070317_chains1190) for which the resolution was better than 1.6 Å and the sequence identity between sequences was less than 20%. The set was taken directly from the Dunbrack's web site (http://dunbrack.fccc.edu).

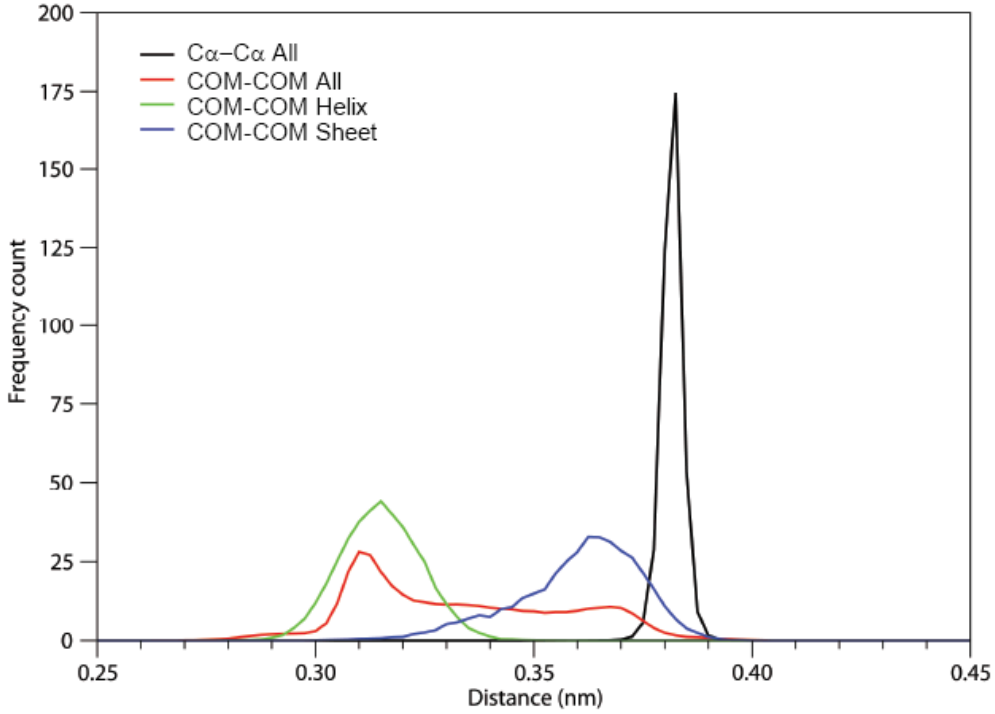


**Figure S1**: Distribution of sequential backbone-backbone distances using either a COM or a Cα placement of the backbone bead. (black) Distribution Cα-Cα distances including all categories of secondary structure. (red) Distribution of COM-COM distance including all categories of secondary structure. Distribution of sequential COM-COM distances for helical residues only (green) and for residues in β-sheets (blue).

The bonded parameters (bond and angle equilibrium values and the corresponding force constants) between backbone beads as well as between the backbone and side chain beads, and within the side chains (for consistency) were re-parameterized as per the new position of the backbone bead. To this end distributions of distances and angles were extracted from atomistic molecular dynamics simulations (see below) from which bond and angle reference values, and force constants were derived assuming that the distributions were Gaussian and could be described by harmonic potentials. Reference values for bonds and angles were set to the centre (average value) of the distributions. The force constants $K_{BOND}$ for bond stretching potentials were obtained as:

$$K_{BOND} = \frac{k_B T}{\sigma^2}$$

with $T$ the temperature (302 K, see below), $k_B$ the Boltzmann constant, and $\sigma^2$ the variance of the distribution. In MARTINI angle potentials are cosine-based, therefore the force constant $K_{ANGLE}$ for angle bending terms were obtained as:

$$K_{ANGLE} = \frac{k_B T}{\sigma_\theta^2 \sin^2(\theta_0)}$$

where $\sigma_\theta^2$ is the variance of the angular distribution and $\theta_0$ is the average of the distribution.

The atomistic (AA) molecular dynamics simulations from which the above parameters were derived (with the same force field and simulation settings as described in the main manuscript) used the temperature-replica exchange molecular dynamics protocol[4]. The parameters for each amino acid were obtained by simulating solvated tri-peptide systems with sequence Ala-X-Ala where X was one of the 20 natural amino acids. The simulations used 8 replicas at temperatures of 296, 302, 308, 314, 320, 327, 333 and 340 K, respectively. Following 100 ps of equilibration during which no exchange was attempted, each tri-peptide was simulated for 5 ns with exchanges attempted between consecutive replicas (temperatures) every picosecond. The distributions were built from the conformational ensemble explored at 302 K.

Note that the distributions obtained from the simulations were preferred to the ones obtained from PDB structures (data not shown). The main reason being that while the reference values for distances and angles are similar in both approaches, the width of the distributions were significantly different. The set of PDB structures yielded very narrow distributions when compared to the ones produced by the atomistic simulations. This resulted in larger force constants and therefore a stiffer model when the PBD distributions were used. It is important to note that this was not the case when the COM was used[2].

## 2. Structural mapping and bond connectivity of aromatic residues

The structural mapping from AA to coarse grained (CG) of the side chains containing a ring was modified with respect to MARTINI-2.1. The structural mapping and bond connectivity of Phe, Tyr, His and Trp used in the manuscript are illustrated in Figure S2. The main changes were the use of two bonds to maintain the ring structure of both the Phe and the Tyr, and the consideration of asymmetry in rings of His and Trp.

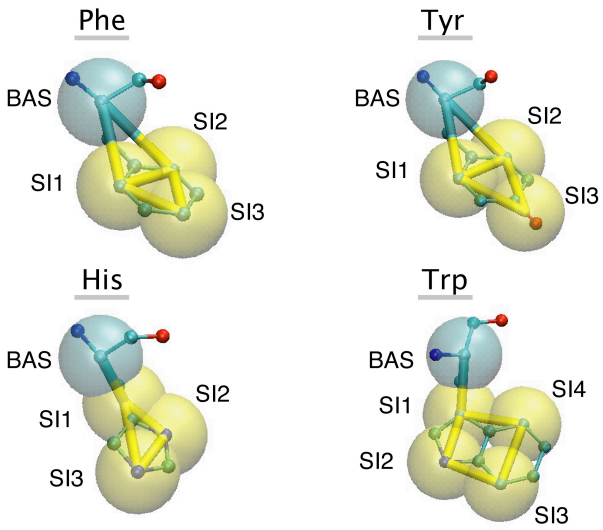


**Figure S2**: Structural mapping and bond connectivity of residues Phe, Tyr, His and Trp. The atomistic model (excluding hydrogen atoms) are shown in ball and stick representation. The backbone nitrogen atom is colored blue, the carbonyl carbon is cyan and the carbonyl oxygen red. The thicker stick represents bonds present in the CG model and the large transparent balls show the CG beads.

## 3. Bonded parameters for amino acid residues.

Note that the bead type used for each CG side chain and backbone CG atom is the same as in the original MARTINI-2.1, and is thus still a function of the secondary structure of the residue[2]. In the following two tables, BAS designates the backbone bead and SIX the $X^{th}$ side chain bead.

In Table 2, bonds whose force constant was found to be higher than 100000 kJ/mol/nm$^2$ during the parameterization stage are marked as "constrained" to indicate that they were constrained during the simulations using the LINCS algorithm[5].

**Table 1**: Backbone parameters[a]

| Type | Ref. value | Force constant |
|---|---|---|
| BAS-BAS distance | 0.38[b] | 150000 |
| BAS-BAS-BAS angle | 120[c] | 40 |

[a] Distances are given in nm, angles in degrees and force constants in kJ/mol/nm$^2$ or kJ/mol for bond and angle potentials, respectively.

[b] This value is the average value of $C^{\alpha}$-$C^{\alpha}$ distances observed in PDB protein structures and the atomistic simulations of tri-peptides. In a simulation of a protein using ELNEDIN the BAS-BAS distance is taken from the model structure of the protein.

[c] This value is the average value of the $C^{\alpha}_1$-$C^{\alpha}_2$-$C^{\alpha}_3$ in the tri-peptide simulations. In the simulations the angle references were set to this value although it might be different from that in the starting PDB structure.

(Table 2 on following page)

**Table 2**: Side chain parameters[a]

| Residue | # of beads | Parameter | Reference value | Force constant[b] |
|---|---|---|---|---|
| Gly | 1 | - | - | - |
| Ala | 1 | - | - | - |
| Cys | 2 | d(BAS-SI1) | 0.24 | 94000 |
| Val | 2 | d(BAS-SI1) | 0.20 | constrained |
| Leu | 2 | d(BAS-SI1) | 0.265 | 81500 |
| Ile | 2 | d(BAS-SI1) | 0.225 | 13500 |
| Met | 2 | d(BAS-SI1) | 0.31 | 2800 |
| Pro | 2 | d(BAS-SI1) | 0.19 | constrained |
| Asn | 2 | d(BAS-SI1) | 0.25 | 61000 |
| Gln | 2 | d(BAS-SI1) | 0.30 | 2400 |
| Asp | 2 | d(BAS-SI1) | 0.255 | 65000 |
| Glu | 2 | d(BAS-SI1) | 0.31 | 2500 |
| Thr | 2 | d(BAS-SI1) | 0.195 | constrained |
| Ser | 2 | d(BAS-SI1) | 0.195 | constrained |
| Lys | 3 | d(BAS-SI1) | 0.25 | 12500 |
| | | d(SI1-SI2) | 0.30 | 9700 |
| | | θ(BAS-SI1-SI2) | 150.0 | 20.0 |
| Arg | 3 | d(BAS-SI1) | 0.25 | 12500 |
| | | d(SI1-SI2) | 0.35 | 6200 |
| | | θ(BAS-SI1-SI2) | 150.0 | 15.0 |
| His | 4 | d(BAS-SI1) | 0.195 | constrained |
| | | d(SI1-SI2) | 0.193 | constrained |
| | | d(SI2-SI3) | 0.216 | constrained |
| | | d(SI1-SI3) | 0.295 | constrained |
| | | θ(BAS-SI1-SI2) | 135.0 | 100.0 |
| | | θ(BAS-SI1-SI3) | 115.0 | 50.0 |
| Phe | 4 | d(BAS-SI1) | 0.34 | 7500 |
| | | d(BAS-SI2) | 0.34 | 7500 |
| | | d(SI1-SI2) | 0.24 | constrained |
| | | d(SI1-SI3) | 0.24 | constrained |
| | | d(SI2-SI3) | 0.24 | constrained |
| | | θ(BAS-SI1-SI2) | 70.0 | 100.0 |
| | | θ(BAS-SI1-SI3) | 125.0 | 100.0 |
| Tyr | 4 | d(BAS-SI1) | 0.335 | 6000 |
| | | d(BAS-SI2) | 0.335 | 6000 |
| | | d(SI1-SI2) | 0.24 | constrained |
| | | d(SI1-SI3) | 0.31 | constrained |
| | | d(SI2-SI3) | 0.31 | constrained |
| | | θ(BAS-SI1-SI2) | 70.0 | 100.0 |
| | | θ(BAS-SI1-SI3) | 130.0 | 50.0 |
| Trp | 5 | d(BAS-SI1) | 0.255 | 73000 |
| | | d(SI1-SI2) | 0.22 | constrained |
| | | d(SI2-SI3) | 0.25 | constrained |
| | | d(SI3-SI4) | 0.28 | constrained |
| | | d(SI4-SI1) | 0.255 | constrained |
| | | θ(BAS-SI1-SI2) | 142.0 | 30.0 |
| | | θ(BAS-SI1-SI3) | 143.0 | 20.0 |
| | | θ(BAS-SI1-SI4) | 104.0 | 50.0 |
| | | θ(SI1-SI2-SI4-SI3) | 180.0 | 200.0 |

[a] Distances are given in nm, angles in degrees and force constants in kJ/mol/nm$^2$ and kJ/mol for bond and angle potentials respectively. The symbol d(X-Y) designates the distance between beads X and Y, and θ(X-Y-Z) designates the angle between beads X, Y and Z.

[b] The term "constrained" indicates that the bond was constrained during the simulations.

**REFERENCES**

(1) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The Martini Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007,** *111*, 7812-7824.

(2) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The Martini Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008,** *4*, 819-834.

(3) Wang, G.; Dunbrack, R. L., Jr. Pisces: A Protein Sequence Culling Server. *Bioinformatics* **2003,** *19*, 1589-1591.

(4) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999,** *314*, 141-151.

(5) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. Lincs: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997,** *18*, 1463-1472.