Rotamers in CHARMM19

Semen A. Trygubenko[*] and David J. Wales[†]

*University Chemical Laboratories, Lensfield Road,
Cambridge CB2 1EW, UK*

December 16, 2004

$V\Big(\mathbf{r}_1(j), \mathbf{r}_2(j), \mathbf{r}_3(j), ...\mathbf{r}_N(j)\Big)$ is the energy functional that describes a system with N atoms, where $\mathbf{r}_i(j)$ is the three-dimensional vector that contains the Cartesian coordinates of atom i for structure j, i.e. $\mathbf{r} = \Big(X_i(j), Y_i(j), Z_i(j)\Big)$, where $X_i(j)$ is the X coordinate of atom i in structure j, etc. Structure A is fully specified by a set of coordinates $\Big(\mathbf{r}_1(A), \mathbf{r}_2(A), \mathbf{r}_3(A), ...\mathbf{r}_N(A)\Big)$. Ideally, the energy functional is invariant with respect to any permutations within a subset of atoms of the same element. For example, if atoms 1 and 2 are both hydrogens, swapping their coordinates around should not make any difference for the evaluation of energy and its derivatives. This effectively means that 'numbering' or 'labelling' of the atoms within such subset is completely arbitrary. Swapping the coordinates and relabelling are therefore equivalent operations.

The CHARMM energy functional is capable of describing some of the permutational isomers. While it is not possible to swap two atoms of the same type that belong to different residues (a rearrangement of this type would involve bond breaking and thus impossible to model with CHARMM), rotations of the side chains can map an initial structure onto a permutational isomer. Such rotational isomers (rotamers) can be thought of as structures where symmetry-related atoms are relabelled. For example, in the phenylalanine side chain a rotation around a bond CB—CG by 180 degrees generates a structure that could otherwise be obtained via permutation CD1 $\Leftrightarrow$ CD2 followed by CE1 $\Leftrightarrow$ CE2. CB, CG, CD1 etc. are CHARMM atom types as specified in CHARMM19 topology and parameter files, toph19.inp and param19.inp.

Landscape searching using a discrete path sampling (DPS) method[1,2] involves precise identification of the stationary points, which is best done by comparing their exact energies and the principal components of the inertia tensor. Recently Evans and Wales applied DPS to met-enkephalin and the B1 domain of protein G. They noted that permutational isomers of stationary points using the

---

[*]email: sat39@cam.ac.uk
[†]email: dw34@cam.ac.uk

CHARMM19 force field can have slightly different energies and geometries even if tightly optimised.[3,4]

There are eight amino acids that have side chain rotamers in CHARMM19. The phenyl rings in tyrosine and phenylalanine have two pairs of carbon-hydrogen united atoms in the ring that are equivalent, namely, CD1, CD2 and CE1, CE2. Aspartic and glutamic acids each have a pair of equivalent oxygens - OD1, OD2 and OE1, OE2, respectively. Asparagine, glutamine and arginine have equivalent hydrogens in amino groups. Lysine has three equivalent hydrogens in NH_3^+-group. Standard N- and C-terminal capping groups NTER and CTER also have rotational isomers. Finally, valine and leucine side chains support a symmetry plane.

Rotamers of tyrosine, asparagine, glutamine and the C-terminal residue CTER can have different energies in the CHARMM19 force field. The energy difference can be as large as 10^{-3} kcal/mol for tightly optimised structures and is due to the asymmetry in the definition of the dihedrals in the case of tyrosine, asparagine and glutamine, and in the definition of the improper dihedral in the case of CTER.

It is easy to demonstrate this effect for tyrosine with a structure for which the axis of rotation of the phenyl ring (which passes through atoms CG and CZ) does not coincide with the torsion axis (which passes through atoms CB and CG). In this case the angles of the torsion CE1-CZ-OH-HH in the two rotamers are not complementary, which results in slightly different energy. Similar situation is found in asparagine and glutamine where rotation of aminogroups is described by asymmetric CB-CG-ND2-HD21 and CG-CD-NE2-HE21 dihedrals, respectively.

The same asymmetry is present in another tyrosine torsion, CA-CB-CG-CD1, as well as the CA-CB-CG-CD1 torsion in phenylalanine, the CB-CG-CD-OE1 torsion in glutamic acid, and the CA-CB-CG-CD1 torsion in aspartic acid, but they are not noticeable when a standard parameter file is used because the corresponding force constants are zero.

The improper torsion C-CA-OT2-OT1 is the origin of the energy difference for the CTER residue. In the general case it is difficult to symmetrise impropers in CHARMM without changing of the form of the energy functional and/or reparametrisation. Fortunately, when the equilibrium angle is zero the improper torsion term is a symmetric function of the angle and, provided symmetry-related atoms are given in 'correct' order, the improper torsion energy and its derivatives will be the same in both rotamers. Specifically, for improper torsion I-J-K-L, where I is the central atom and the torsion angle is the angle between the plane defined by atoms I-J-K and atom L, if the symmetry related atoms are J and K ϕ(I-J-K-L)$=-\phi$(I-K-J-L). While tyrosine impropers CG-CD1-CD2-CB and CZ-CE1-CE2-OH, phenylalanine CG-CD1-CD2-CB improper, CD-OE1-OE2-CG improper for glutamic acid and CG-OD1-OD2-CB improper for aspartic acid are all symmetrically specified, the CTER's C-CA-OT2-OT1 improper is not.

The valine and leucine situation is special because if one of the two structures with side chains that are mirror images of one another is a stationary point the other one is not. In other words, because the rearrangement where the $C-C(CH_3)_2$

group undergoes an inversion is not describable by the force field, only one of the two isomers can exist on the CHARMM19 potential energy surface as a stationary point.

We have fixed dihedral asymmetries in ASN, GLN, TYR, PHE, GLU, ASP, VAL and LEU by adding multiple dihedral terms between bonds with similar symmetry-related positions and halving the corresponding force constants, as was advised by Charles L. Brooks III. We also 'symmetrised' CTER improper torsion by changing the topology for that residue. Patches for toph19.inp, param19.inp, toph19_eef1.inp and param19_eef1.inp are available online.[5]

References

[1] D. J. Wales, Mol. Phys. **100**, 3285 (2002).

[2] D. J. Wales, Mol. Phys. **102**, 883 (2004).

[3] D. A. Evans and D. J. Wales, The Journal of Chemical Physics **119**, 9947 (2003).

[4] D. A. Evans and D. J. Wales, The Journal of Chemical Physics **121**, 1080 (2004).

[5] *http://www.trygub.com/charmm/rotamers/* (2004).