Supplementary Information

for: **"PRIMO/PRIMONA: A coarse-grained model for proteins and nucleic acids that preserves near-atomistic accuracy"**

Srinivasa M. Gopal, Shayantani Mukherjee, Yi-Ming Cheng, Michael Feig

Construction of local coordinate system for scheme 1

In scheme 1, an atomic site A ($\vec{r}_A$) is constructed based on the distance A-B (b), the angle A-B-C (θ), and the dihedral A-B-C-D (ϕ). Let $\vec{r}_B$, $\vec{r}_C$, $\vec{r}_D$ be the position vectors of known atoms B, C and D respectively. Using the above information, the position of Atom in a local coordinate system ($\vec{r}_A^{\,l}$) can be represented in a local coordinate system (where b is the radius, θ is the polar angle and ϕ is the azimuthal angle) by the spherical coordinate transformation as

$$\vec{r}_A^{\,l} = \begin{bmatrix} b\sin(\theta)\sin(\phi) \\ b\sin(\theta)\cos(\phi) \\ b\cos(\theta) \end{bmatrix} \quad (1)$$

The local coordinate system is constructed as follows. First the local z-axis ($\hat{w}$) is defined as

$$\hat{w} = \frac{\vec{r}_C - \vec{r}_B}{|\vec{r}_C - \vec{r}_B|} \quad (2)$$

Local X-axis ($\hat{u}$) is defined as the unit normal to plane defined by $\hat{w}$, $\vec{r}_B$ and $\vec{r}_D$ vectors as

$$\hat{u} = \frac{\vec{r}_{BD} \times \hat{w}}{|\vec{r}_{BD} \times \hat{w}|} \quad (3)$$

Since X and Z axis are defined, there is a unique Y-axis which is

$$\hat{v} = \hat{w} \times \hat{u} \quad (4)$$

The transformation[1,2] (**T**) which maps atoms B, C and D to a local coordinate system such that where B is placed at origin, C on Z-axis and D in Y-Z plane is given by

$$T = \begin{bmatrix} u^x & v^x & w^x & 0 \\ u^y & v^y & w^y & 0 \\ u^z & v^z & w^z & 0 \\ -\hat{u}.\hat{r}_B & -\hat{v}.\hat{r}_B & -\hat{w}.\hat{r}_B & 1 \end{bmatrix} \quad (5)$$

We need the inverse of above transformation (**I**) to express local position of atom A $\vec{r}_A^{\,l}$ in terms of global coordinates.

$$I = \begin{bmatrix} u^x & u^y & u^z & 0 \\ v^x & v^y & v^z & 0 \\ w^x & w^y & w^z & 0 \\ r_B^{\,x} & r_B^{\,y} & r_B^{\,z} & 1 \end{bmatrix} \quad (6)$$

$$\vec{r}_A = I.\vec{r}_A^{\,l}$$

Comparison of reconstruction methods for villin decoy set

Residue/Type	RMSD (σ_{RMSD}) in Å				
	PRIMO	**SICHO/CA**	**SICHO**	**BB**	**CA**
All	0.083 *(0.01)*	1.183 *(0.11)*	1.498 *(0.10)*	1.503 *(0.20)*	1.978 *(0.18)*
Back	0.064 *(0.02)*	0.753 *(0.07)*	1.191 *(0.13)*	0.000 *(0)*	0.647 *(0.10)*
Side	0.095 *(0.02)*	1.480 *(0.16)*	1.739 *(0.13)*	2.104 *(0.28)*	2.695 *(0.25)*
C_α	0 *(0)*	0 *(0)*	0.863 *(0.13)*	0.000 *(0)*	0.000 *(0)*
C_α+C_β	0.065 *(0.01)*	0.501*(0.05)*	1.060 *(0.11)*	0.047 *(0.01)*	0.403 *(0.09)*
Arg	0.047 *(0.02)*	1.086 *(0.39)*	1.345 *(0.41)*	2.853 *(1.82)*	3.671 *(1.82)*
Asn	0.007 *(0.00)*	1.013 *(0.37)*	1.414 *(0.32)*	2.032 *(0.74)*	2.383 *(0.52)*
Asp	0.012 *(0.00)*	0.735 *(0.23)*	1.103 *(0.34)*	1.945 *(0.64)*	2.123 *(0.43)*
Gln	0.056 *(0.02)*	0.975 *(0.29)*	1.301 *(0.28)*	2.277 *(0.77)*	2.245 *(0.78)*
Glu	0.055 *(0.02)*	0.742 *(0.21)*	1.265 *(0.280)*	2.021 *(0.70)*	2.280 *(0.69)*
Leu	0.139 *(0.06)*	0.888 *(0.26)*	1.345 *(0.23)*	1.093 *(0.35)*	1.512 *(0.44)*
Lys	0.045 *(0.01)*	0.848 *(0.20)*	1.356 *(0.22)*	2.255 *(0.53)*	2.477 *(0.53)*
Met	0.051 *(0.01)*	2.405 *(0.70)*	2.259 *(0.51)*	1.875 *(0.68)*	2.891 *(1.10)*
Phe	0.097 *(0.02)*	2.581 *(0.46)*	2.683 *(0.34)*	1.470 *(1.02)*	3.267 *(0.81)*
Pro	0.144 *(0.08)*	0.264 *(0.19)*	0.914 *(0.43)*	0.297 *(0.30)*	0.460 *(0.31)*
Ser	0.094 *(0.01)*	0.295 *(0.20)*	0.945 *(0.29)*	1.025 *(0.46)*	1.267 *(0.39)*
Thr	0.124 *(0.02)*	0.894 *(0.55)*	1.297 *(0.51)*	0.520 *(0.57)*	1.279 *(0.68)*
Trp	0.105 *(0.01)*	0.704 *(0.59)*	1.510 *(0.87)*	3.259 *(1.60)*	3.627 *(1.76)*
Val	0.251 *(0.14)*	0.557 *(0.38)*	1.007 *(0.45)*	0.313 *(0.41)*	0.812 *(0.53)*

References

1. Newman, W. M.; Sproul, R. F. Principles of Interactive Computer Graphics; McGraw-Hill: New York, 1979.
2. Bower, M. J.; Cohen, F. E.; Dunbrack, R. L. Journal of Molecular Biology 1997, 267(5), 1268-1282.