Supplementary Information

Self-Assembly of the Full-length Amyloid Aβ42 Protein in Dimers

Yuliang Zhang, Mohtadin Hashemi, Zhengjian Lv and Yuri L. Lyubchenko*

Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, 68198, USA.

*Corresponding author:

Yuri L. Lyubchenko

Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, 68198,

USA

ylyubchenko@unmc.edu

SIMULATION METHODS

Monomer simulation procedure. To mimic the experimental design, a Cys residue was added to the N-terminus. Since the N-terminus is flexible and less important than the central or C-terminus regions of Aβ peptide.[1] Thus, the N-terminus usually was selected as tethering target in experiments. The addition Cys residue does not significantly affect the peptide structure and aggregation behavior.[2, 3] The index number of this Cys residue was set to 0 to keep the context of the other residues as the actual Aβ42 protein. The structure was then solvated in a truncated octahedron box with 10620 TIP3P water molecules. The minimum distance between the protein surface and the edges of the water box was 1.5 nm, so that any interactions between periodic copies, due to periodic boundary condition (PBC), are avoided. The protonation states of Lys and Arg residues were set to mimic neutral pH conditions, at which both contain 1 positive charge. The nitrogen atoms at the δ position of the His residues were also protonated. 32 Na^+ and 29 Cl^- ions were added to neutralize the system charges and keep a constant salt concentration of 150 mM. Other details of the simulations setup were adopted from our previous work.[2]

Dimer simulation on the specialized supercomputer Anton. The dimer was solvated into a tetragonal box, with the edges of 8.2 nm and 8.4 nm, along with 18031 TIP3P water molecules. The minimum distance between the proteins and the box edge was 1.5 nm. In order to allow relaxation and free tumbling before intermolecular contact, the COM distance of two monomers was set to 4 nm. 56 Na^+ and 50 Cl^- ions were placed in the box to neutralize the protein charges and to maintain an ionic concentration of 150 mM. The protonation of charged residues was processed the same way as the monomer simulations. The viparr.py script from Maestro-Desmond package was employed to load the Amber ff99SB-ILDN force field and the TIP3P

water model and to constrain the mobility of the hydrogen atoms using the M-SHAKE algorithm.[4] Then, the system was equilibrated using 20 ns NPT cMD simulations on our local HCC cluster. The resulting system from the last frame of the 20 ns simulations was used as the initial input for 4 μs cMD simulation run on Anton. The input parameters were optimized by the guess_chem command available on the Anton machine. The multigrator scheme from Anton was used to achieve elevated flexibility of the setup during the integration steps. All simulations utilized the Martyna-Tobias-Klein (MTK)[5] and the Nosé-Hoover algorithms[6] for constant pressure of 1 bar and constant temperature of 300 K. The short-range unbound interactions beyond 0.9 nm were ignored and the long-range electrostatics were calculated by the particle-mesh Ewald (PME) algorithm[7]. The integration time step was 2 fs and the output frequency was 240 ps.

Accelerated MD (aMD) simulation. Briefly, the dimer structure was extracted from the last frames in the cMD simulation on Anton and all of the hydrogen atoms were removed to avoid conflicts within the conversion between different MD packages. Then, the tleap command from Ambertool(Amber 12; University of California:San Francisco, CA, 2012) was used to solvate, neutralize, and make a 150 mM ion concentration within the dimer system with the same force field and the same solvent model conditions as the cMD simulations described above. Other than proteins, the final system contained 11480 water molecules, 38 Na^+ ions, and 32 Cl^- ions. The charged residues, Lys, Arg, and His, were processed in the same way as described in the monomer section. Then, the output products were taken as input files to run 6 step cMD simulations following the online tutorial prescriptions (URL: http://ambermd.org/tutorials/advanced/tutorial22/) for energy minimization and system

relaxation. Finally, 500 ns aMD simulation was performed via the Amber12 PMEMD (Amber 12; University of California: San Francisco, CA, 2012).

According to papers[8, 9], a boost potential $\Delta V(r)$ is introduced to the original potential $V(r)$ to raise the energy surface near a minimum. Using this method, the proteins are able to escape from potential wells, thereby enhancing the sampling of the conformational space. Boost potential is applied conditionally, when the $V(r)$ is smaller than the selected threshold energy E, the simulation will be run with the modified potential $V^*(r) = V(r) + \Delta V(r)$; if $V(r)$ is larger than E, the simulation will use the true potential $V^*(r) = V(r)$. With $\Delta V(r)$ defined as:

$$\Delta V(r) = \begin{cases} 0, & V(r) \geq E \\ \dfrac{(E - V(r))^2}{\alpha + (E - V(r))}, & V(r) < E \end{cases} \quad (1)$$

where, α is the tuning parameter that administers the depth and roughness of the modified potential. The smaller α is the less rough the modified potential will be.

The dual boost approach, in which both torsional and total energies are taken into account,[10] was utilized to explore the Aβ42 dimerization process. Parameters for aMD simulation were calculated based on the last step of the cMD relaxation simulation. The appropriate total boost parameters (E_{tot} and α_{tot}) and dihedral boost parameters (E_{dih} and α_{dih}) were calculated according to the procedure from Pierce *et al.*[8] as follows:

$$EthreshP: E_{tot} = -118009\frac{kcal}{mol} + \left(0.16\frac{kcal}{mol\ atom}\ 35786\ atoms\right) = -112283\frac{kcal}{mol} \quad (2)$$

$$\alpha P: \alpha_{tot} = \left(0.16\frac{kcal}{mol\ atom}\ 35786\ atoms\right) = 5726\frac{kcal}{mol} \quad (3)$$

$$EthreshD: E_{dih} = 796\frac{kcal}{mol} + \left(4\frac{kcal}{mol\ residue}\ 86\ residues\right) = 1140\frac{kcal}{mol} \quad (4)$$

$$\alpha D: \alpha_{dih} = \frac{1}{5}\left(4\frac{kcal}{mol\ residue}\ 86\ residues\right) = 68.8\frac{kcal}{mol} \quad (5)$$

In order to keep the temperature at 300K, the Langevin thermostat was used with a collision frequency of 5 ps^{-1}. The cutoff for short-range non-bonded interactions was set to 12 Å.

Graphic Software. The figures of the contact map and the free energy landscape in the dPCA analysis were generated via Python2.7 (Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org).[11, 12] The force curves were analyzed by Matlab 2012 (MathWorks Inc., Natick, MA, USA). All of the line plots, scatter plots, and distributions were produced by Igor Pro v6.3 (WaveMetrics, Lake Oswego, OR, USA). The statistical analysis of the force distributions was performed using the Kolmogorov-Smirnov nonparametric test (SPSS 20.0; IBM Corp, Armonk, NY, USA). The angle calculations and movies were made with the VMD software package,[13] and the protein snapshots were generated by YASARA (www.yasara.org).

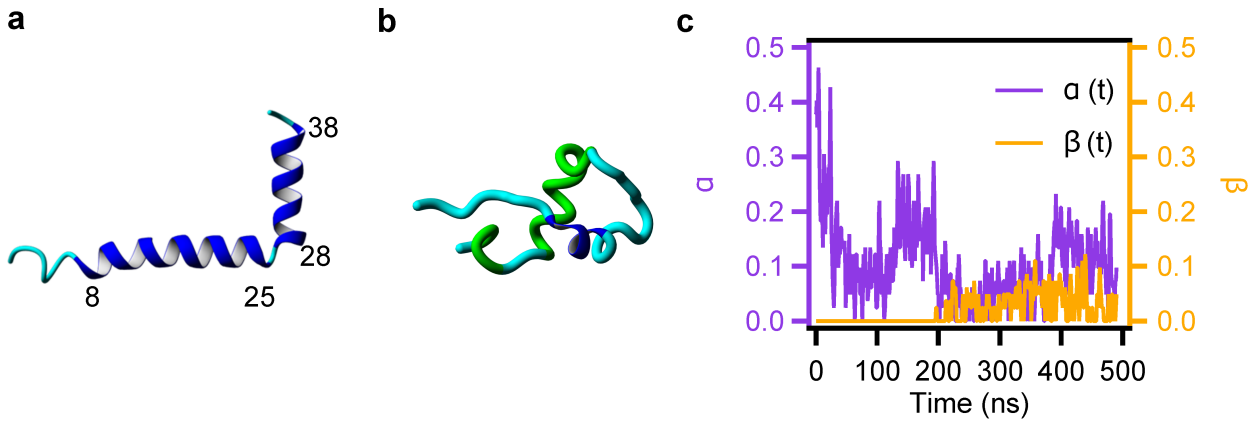


Figure S1. The dynamics of Aβ42 monomer**. (a)** Initial structure for the Aβ42 monomer simulation**.** Coordinates of Aβ42 monomers were taken from the Protein Database Bank (PDB ID: 1IYT). Two helical regions are represented by blue ribbons, encompassing residues 8–25 and 28–38; remaining residues are shown as a cyan tube. (b) The largest cluster comprising 53.73% of the entire population. (**c**) The time-dependent dynamics of the secondary structure. Purple curves represent the fractions of α-helix content and orange curves indicate the fluctuation of the fractions of β-strand content over time.

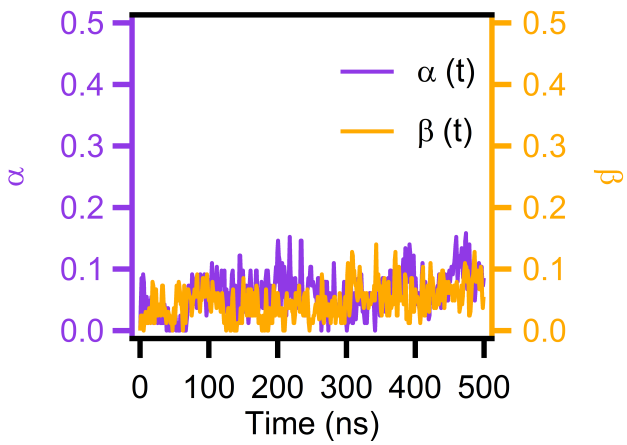


Figure S2. The time-dependent changes of the content of α-helix (purple) and β-strand (orange) obtained from aMD simulations.

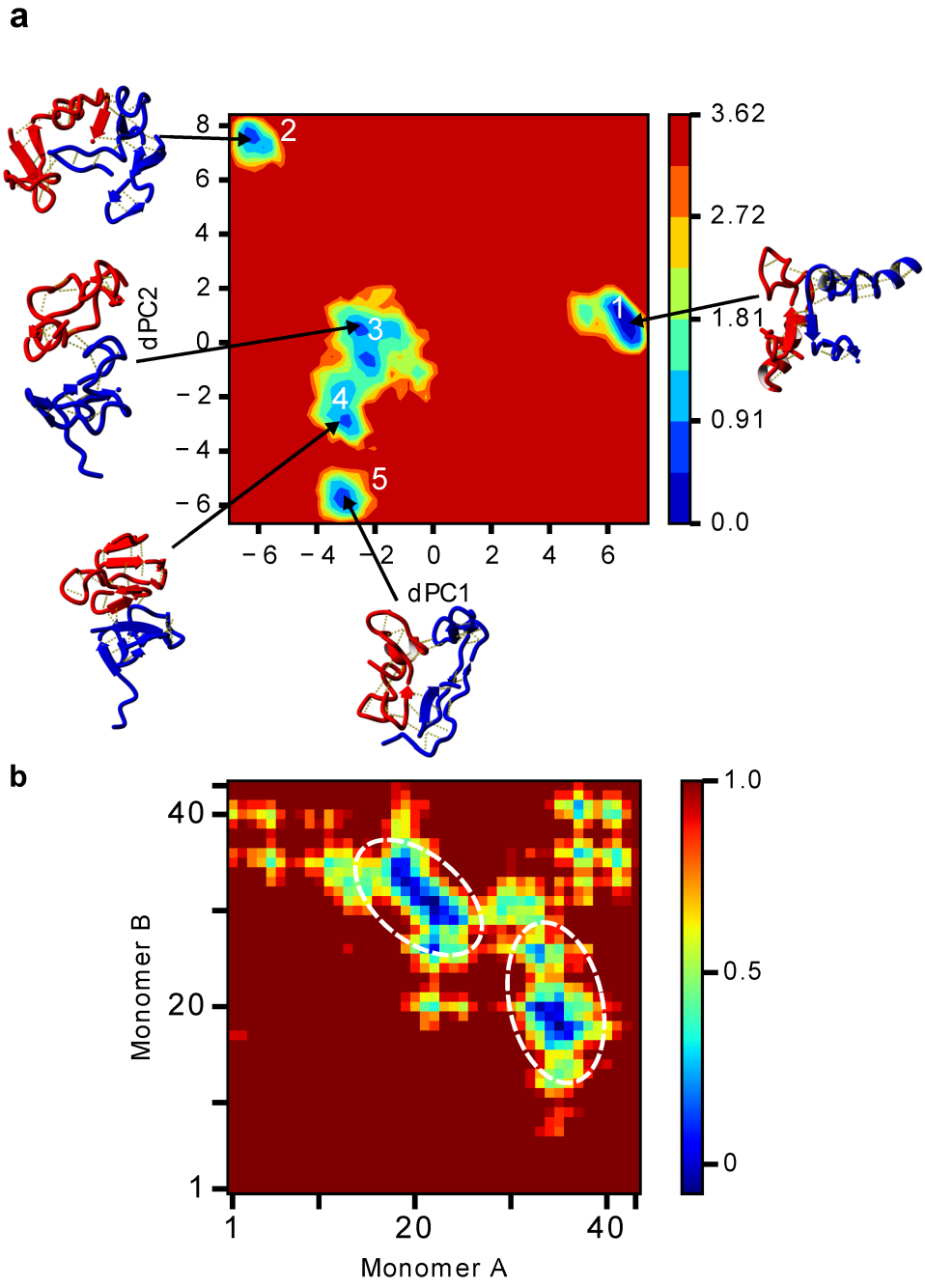


Figure S3. The analysis of REMD simulation of Aβ42 dimer. (a). The free energy landscape constructed after REMD simulations**.** Snapshots of representative structures corresponding to states with local energy minima are indicated with arrows. In the snapshots red and blue lines represent monomers A and B, respectively. The dashed lines indicate hydrogen bonds. (b). The contact map of Aβ42 dimer (structure 1). The colors in the contact map indicate the distance in nm between the pairwise residues. The regions of interest are highlighted with dashed lines.

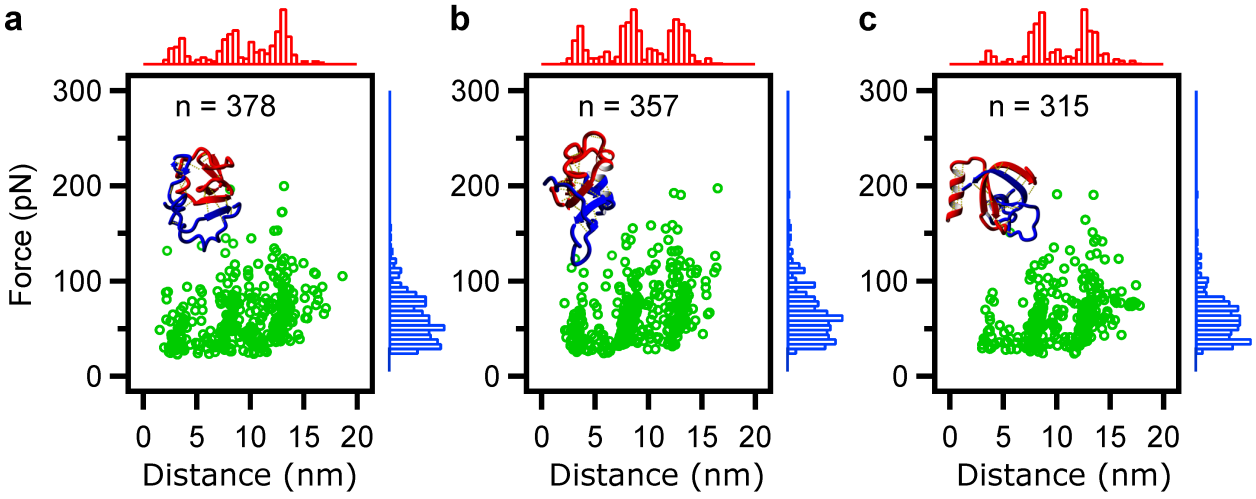


Figure S4. Rupture patters of three major structures for Aβ42 dimer corresponding to structures 2–4 in Fig. 2b**.** Each rupture event is shown with green circles. The force distributions are shown as blue histograms and placed on the right side of the scatter plot. The rupture lengths distributions (red histogram) are placed at the top of the plot. The insets are initial structures prior to the pulling.

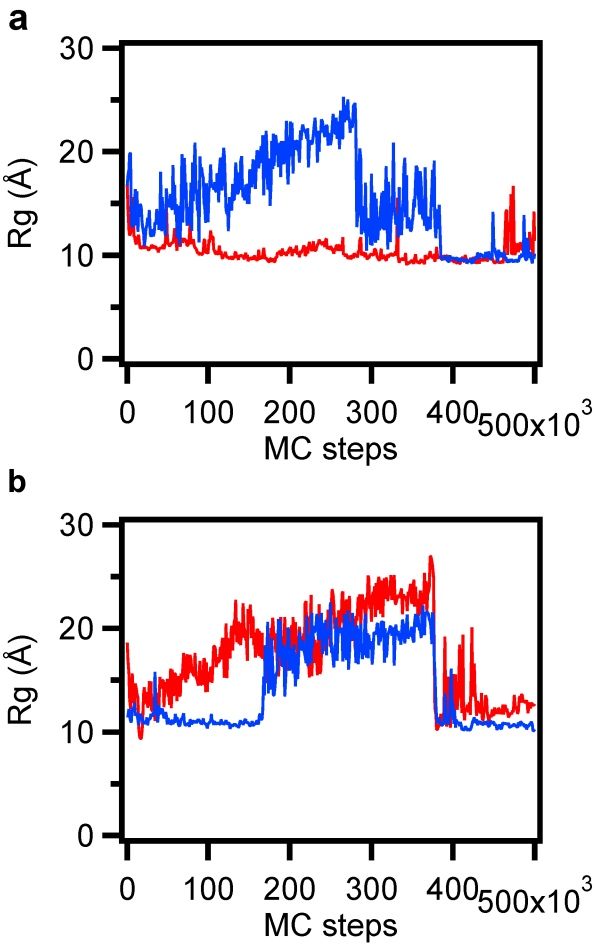
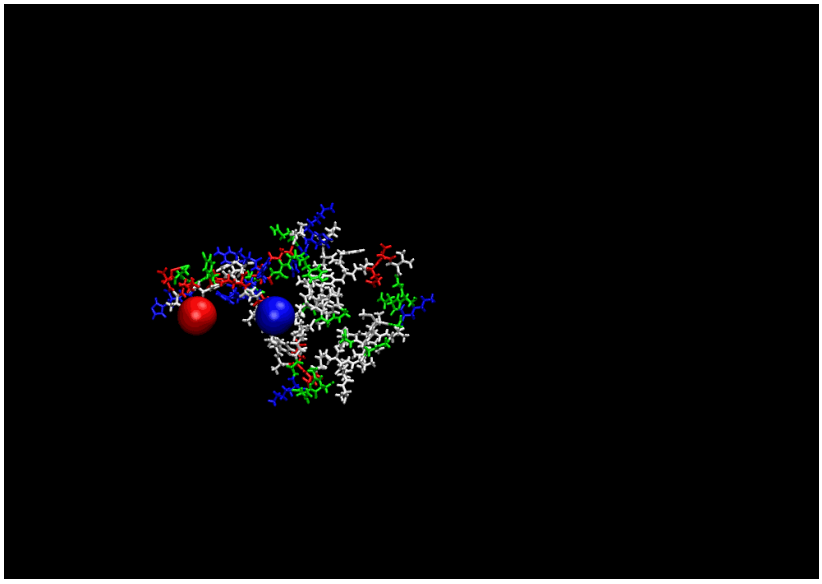
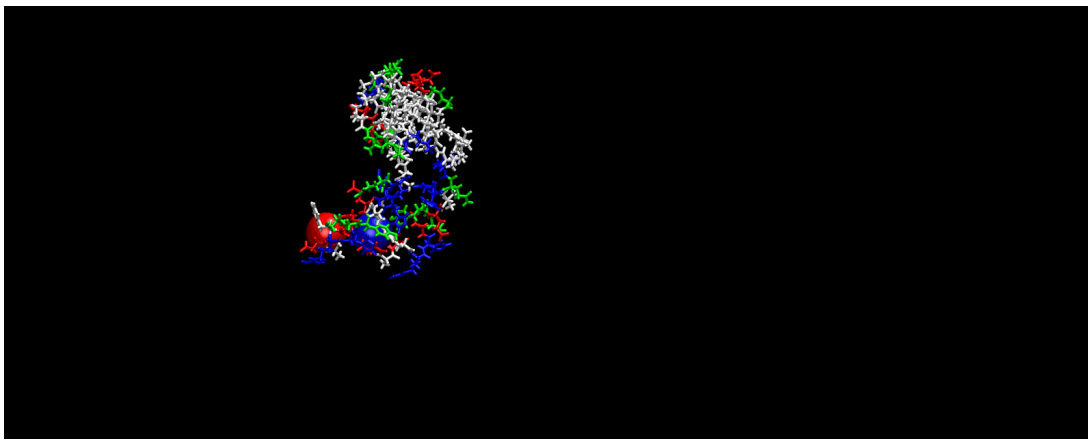


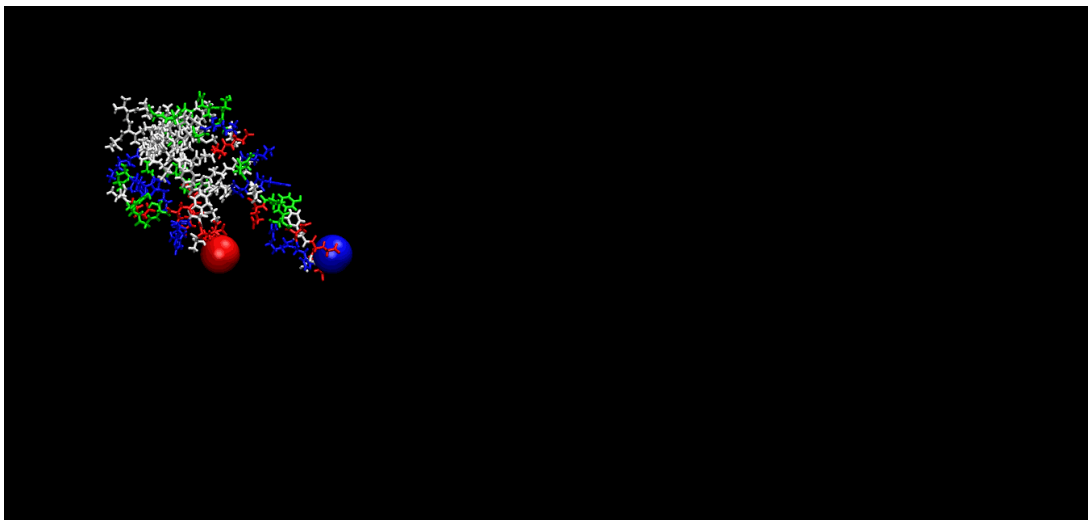
Figure S5. Fluctuation of the Radius of Gyration (Rg) for the asymmetric (a) and symmetric (b) unraveling processes. Red and blue lines correspond to the Rg changes for monomers A and B, respectively.



Movie S1. A movie illustrating the asymmetric unraveling process of dimer corresponding to the rupture pattern I in Figs. 5a. The gray color is the hydrophobic region and the green color shows the polar region. The negatively and positively charged residues are shown in red and blue, respectively. The pulling force was applied to terminal residues, indicating by the red and blue balls.



Movie S2. A movie illustrating the asymmetric unraveling process of dimer corresponding to the rupture pattern II in Fig. 5b. The gray color is the hydrophobic region and the green color shows the polar region. The negatively and positively charged residues are shown in red and blue, respectively. The pulling force was applied to terminal residues, indicating by the red and blue balls.



Movie S3 A movie illustrating the symmetric unraveling process of dimer corresponding to the rupture pattern III in Fig. 5c. The gray color is the hydrophobic region and the green color shows the polar region. The negatively and positively charged residues are shown in red and blue, respectively. The pulling force was applied to terminal residues, indicating by the red and blue balls.

References

1. T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert and R. Riek, *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 17342-17347.

2. S. Lovas, Y. Zhang, J. Yu and Y. L. Lyubchenko, *J. Phys. Chem. B*, 2013, **117**, 6175-6186.

3. Z. Lv, R. Roychaudhuri, M. M. Condron, D. B. Teplow and Y. L. Lyubchenko, *Sci. Rep.*, 2013, **3**.

4. V. Kräutler, W. F. van Gunsteren and P. H. Hünenberger, *J. Comput. Chem.*, 2001, **22**, 501-508.

5. G. J. Martyna, D. J. Tobias and M. L. Klein, *J. Chem. Phys.*, 1994, **101**, 4177-4189.

6. G. J. Martyna, M. L. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635-2643.

7. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.

8. L. C. Pierce, R. Salomon-Ferrer, F. d. O. C. Augusto, J. A. McCammon and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 2997-3002.

9. D. Hamelberg, J. Mongan and J. A. McCammon, *J. Chem. Phys.*, 2004, **120**, 11919-11929.

10. D. Hamelberg, C. A. F. de Oliveira and J. A. McCammon, *J. Chem. Phys.*, 2007, **127**, 155102.

11. J. D. Hunter, *Comput. Sci. Eng.*, 2007, **9**, 90-95.

12. T. E. Oliphant, *Comput. Sci. Eng.*, 2007, **9**, 10-20.

13. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph. Model.*, 1996, **14**, 33-38, 27-28-33-38, 27-28.