Electronic Supplementary Information

A Molecular Dynamics Perspective on the Cyclization Efficiency for Poly(2-oxazoline)s and Poly(2-oxazine)s

Nick Huettner[a,b], Tim R. Dargaville[a,b*] and Neha S. Gandhi[a,c*]

[a] School of Chemistry and Physics, Faculty of Science, Queensland University of Technology (QUT), Brisbane, QLD 4000, Australia.

[b] Centre for Materials Science, Queensland University of Technology (QUT), Brisbane, QLD 4000, Australia.

[c] Centre for Genomics and Personalized Health, Queensland University of Technology (QUT), Brisbane, QLD 4000, Australia.

* Corresponding Authors (email: t.dargaville@qut.edu.au, neha.gandhi@qut.edu.au)

1. Methods

Model construction

Chemical structures of poly(2-*n*-propyl-2-oxazoline) (PnPropOx) and poly(2-*n*-propyl-2-oxazine) (PnPropOzi) polymers were obtained using the MarvinJS (Chemaxon internal, last accessed 27/02/2023) web application. These structures were then fed into Avogadro[1] (software version 1.2.0) to add hydrogen atoms before using the PolyParGen[2] online service to generate parameters for the "all-atom optimized parameters for liquid simulations" (OPLS-AA) forcefield[3] of each polymer. Finally, the GROMACS software package[4–10] (version 2019.3-foss-2019a, source code available[11]) was used for energy minimization, system equilibration and molecular dynamics simulations under periodic boundary conditions (PBC). Simulation time steps were chosen as 2 fs using the leap-frog integrator. All covalent bonds were constrained using the Linear Constraint Solver (LINCS) algorithm. Non-bonded interactions were considered using the Verlet cut-off scheme.The electrostatic interactions were treated using the Particle Mesh Ewald (PME) method. The cut-off radii for Coulomb and van der Waals interactions were set to 10.0 Å.

The through PolyParGen generated single polymer chains of PnPropOx or PnPropOzi were equilibrated under canonical (NVT) conditions for 10 ps, saving the trajectory every 1 ps. The 2 ps time point was

extracted and used for further simulation. This step was necessary to let the main and side chain groups on the polymer orient in space without chain collapse to avoid diverging forces in further simulation steps. The obtained polymer systems were then solvated using a dichloromethane (DCM) solvent system (OPLS-AA forcefield) in a cubic box. A detailed description of the solvent system development can be found below. Following solvation, the energy of the polymer-solvent system was minimized using the steepest descent minimization algorithm and stopped when the maximum force reached below 1000.0 kJ/mol/nm. The system was then equilibrated under NVT conditions for 1 ns to a temperature of 300 K using a modified Berendsen Thermostat[12]. Trajectories were saved every 10 ps. Using isobaric-isothermal (NPT) conditions the system pressure was kept at 1 bar and the density adjusted close to the experimental value of DCM. NPT equilibration was conducted for 10 ns with trajectories being recorded every 10 ps. Pressure was adjusted using the Parrinello-Rahman barostat[13,14]. Both, equilibration under NVT and NPT conditions were conducted to convergence of either temperature (NVT) or pressure and density (NPT). Finally, the production run was done under NVT conditions, without position restraint, for 300 ns with the trajectory being saved every 100 ps. The trajectories were post-processed by removing PBC and setting the center of mass of the polymer chain as the center of the box. For each single polymer-solvent system, a total number of three simulations were conducted. Parameters for each polymer-solvent system after equilibration can be found in Table S1.

The influence of multiple polymer chains on the solution properties of the simulated polymers was evaluated by creating systems containing two polymer chains. These have been constructed similarly to the single polymer systems. A single production run was conducted for these systems. Their parameter sets after equilibration can be found in Table S1.

Dichloromethane solvent system construction

The chemical structure of DCM was obtained using the MarvinJS web application. Hydrogen atoms were added using the Avogadro software. Finally, to obtain OPLS-AA forcefield parameters, the molecule was fed into the LigParGen[15–17] online service. Next, a system containing 1200 DCM molecules was constructed and energy minimized, followed by NVT and NPT equilibration until convergence of temperature, pressure, and system density. Finally, a production was over 100 ns under NPT conditions was conducted to result in the final DCM solvent system (box size: 4.93×4.93×4.93 nm). The DCM box was directly used for solvation of the above-mentioned polymers and their simulation.

Analysis of the molecular dynamics trajectories

The polymer conformations in DCM were visualized using Visual Molecular Dynamics (VMD)[18] (version 1.9.4a53, University of Illinois at Urbana-Chaimpaign). Polymer chains after solvent removal and non-covalent interactions within the polymer were visualized using the Chimera software package[19] (version 1.16, University of California, San Francisco). Carbon, hydrogen, oxygen and nitrogen atoms are shown in black, white, red and blue, respectively. Analysis of dihedral angles of the polymer side chains, radius of gyration of the polymers, radial distribution functions (RDF), and end-to-end distance of terminal polymer atoms (terminal hydrogen atom of the alkyne and terminal nitrogen atom of the azide terminus) was conducted using various post-processing modules in the GROMACS software. Applicable analysis methods were conducted on three replicate simulations for single polymer systems and a single simulation for two-polymer systems. Analysis of dihedral angles within the polymer chain was conducted using the trajectory of the first simulation replicate. All figures show, unless noted otherwise, exemplary the data of the first simulation replicate. The obtained data files were plotted using SigmaPlot (version 14.0, SPSS Inc.). The number of events, in which the polymer chain termini reach distances below 2 and 1 nm were determined using Microsoft Excel (version 2211, Microsoft). RDF functions were derived between the oxygen atoms of the polymer chain and the carbon atoms of the

surrounding DCM solvent molecules. Intramolecular non-bonded interactions within the polymer chain of $PnPropOzi_{20}$ have been determined using the Chimera software package.

2. Supplementary Data

Table S1 Polymer-solvent system parameters after NVT and NPT equilibration.

Polymer	Run	NVT			NPT		
		T_{DCM} [K]	T_{Poly} [K]	Box Volume [nm^3]	System Density [g/cm^{-3}]	Pressure [bar]	Box Volume [nm^3]
$PnPropOx_{20}$	1	299.96	299.94	528.48	1.385	1.08	513.10
	2	299.98	300.04	528.48	1.384	0.33	516.47
	3	299.96	300.09	528.48	1.383	0.43	515.29
$PnPropOzi_{20}$	1	299.97	299.93	1237.53	1.395	1.09	1220.38
	2	299.96	300.26	1237.53	1.388	0.11	1225.87
	3	299.96	299.94	1237.53	1.388	0.06	1225.42
$PnPropOzi_{15}$	1	300.02	300.51	511.88	1.383	0.18	502.41
	2	299.96	299.97	511.88	1.383	-0.02	505.06
	3	300.00	299.98	511.88	1.383	0.16	502.59
2x $PnPropOx_{20}$	1	300.00	300.12	528.58	1.382	0.85	513.16
2x $PnPropOzi_{20}$	1	299.98	299.90	1704.69	1.392	0.93	1667.83

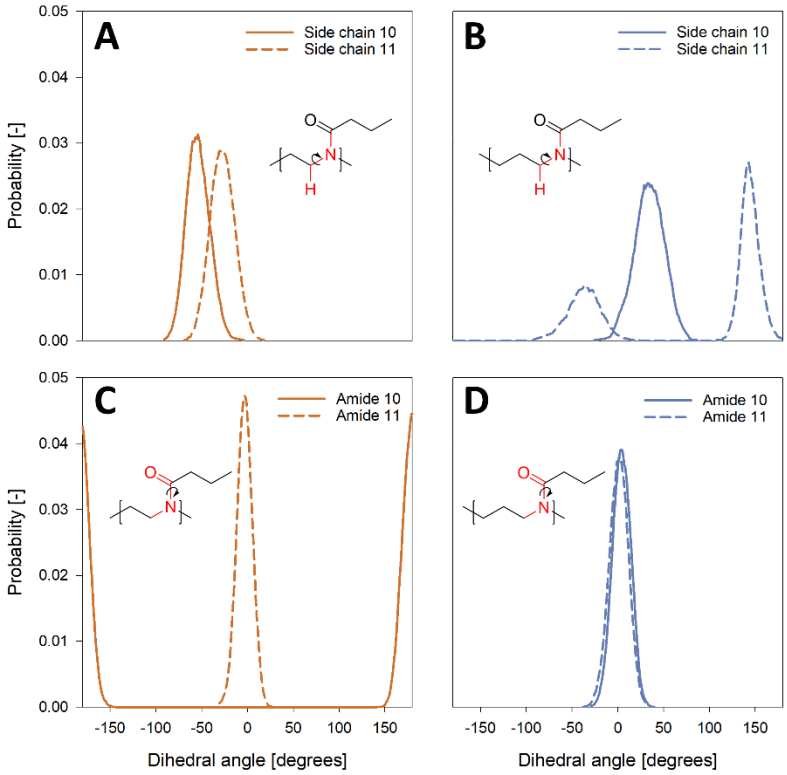

Fig. S1 Plot showing distributions of dihedral angles determining the orientation of the polymer side chain (A and B) and the carbonyl group (C and D) in PnPropOx$_{20}$ (B and D, orange) and PnPropOzi$_{20}$ (A and C, blue). For each set, groups at repeating units 10 (solid line) and 11 (dashed line) have been evaluated. Hydrogen atoms are not shown unless they were relevant for dihedral angle calculation. Investigated atoms for dihedral angle determination have been marked in red. Data is presented as a running average and was obtained from the first simulation of three replicates.

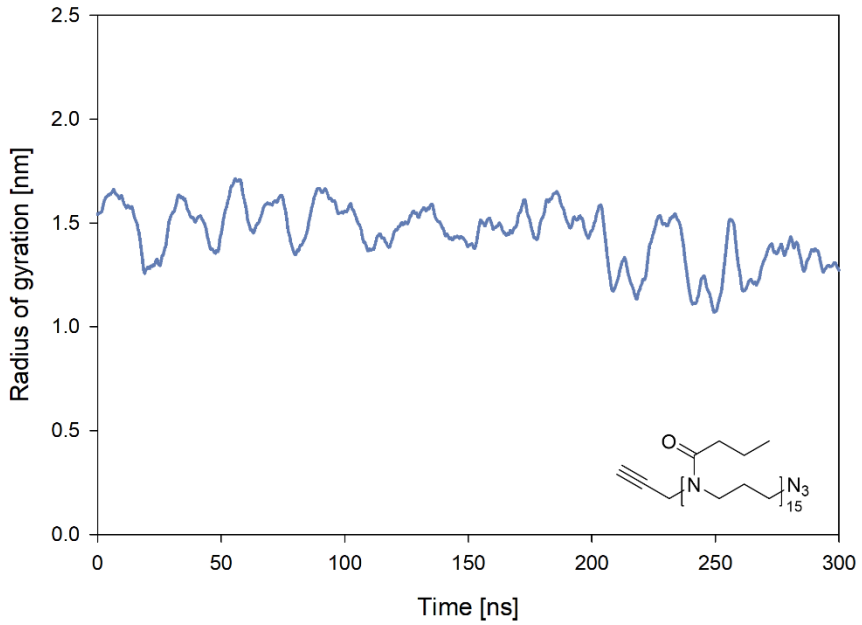

Fig. S2 Running average of the radius of gyration of PnPropOzi$_{15}$ as a function of time. Data was obtained during a 300 ns production run under NVT conditions and represents the first simulation of three replicates.

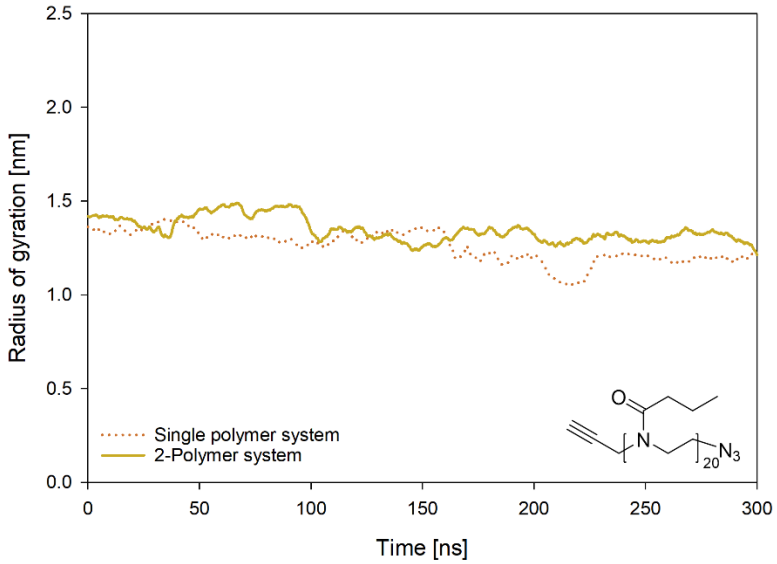

Fig. S3 Radius of gyration of $PnPropOx_{20}$ as a function of time. The graph compares the running averages of a polymer-solvent system containing either a single (dotted line, orange) or two polymers (solid line, yellow). Data was obtained during a 300 ns production run under NVT conditions. The data from the single polymer system represents the first simulation of three replicates.

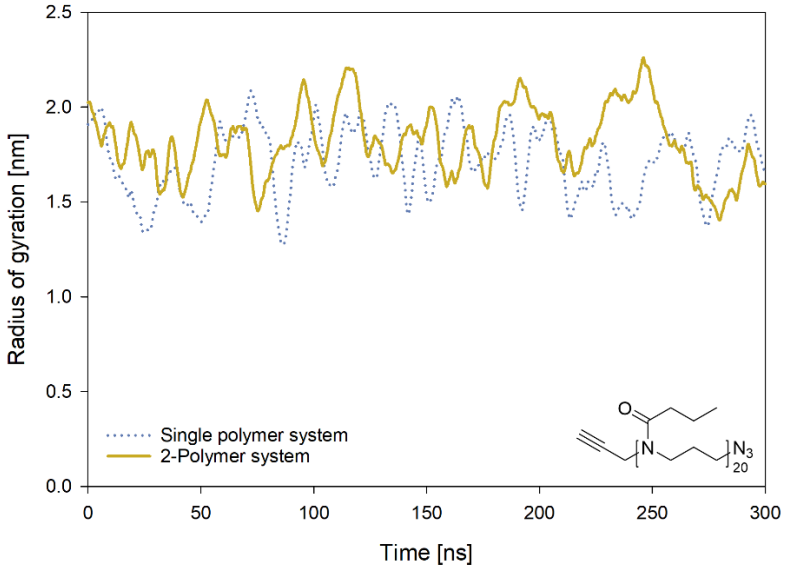

Fig. S4 Radius of gyration of $PnPropOzi_{20}$ as a function of time. The graph compares the running average of a polymer-solvent system containing either a single (dotted line, blue) or two polymers (solid line, yellow). Data was obtained during a 300 ns production run under NVT conditions. The data from the single polymer system represents the first simulation of three replicates.

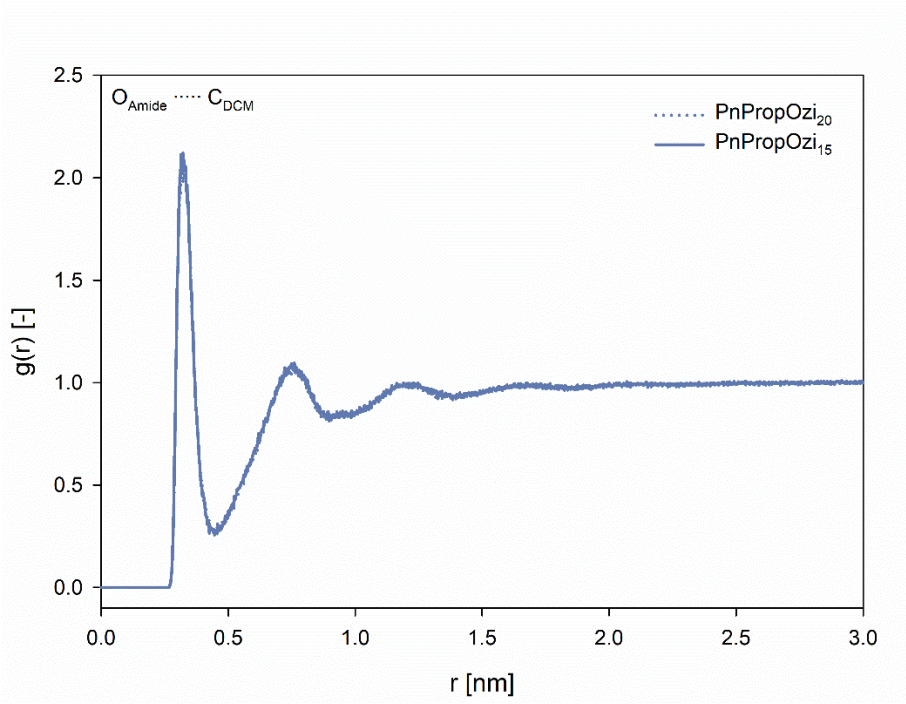

Fig. S5 Comparison of radial distribution functions calculated between oxygen atoms of the amide groups in $PnPropOzi_{20}$ (dotted line) or $PnPropOzi_{15}$ (solid line) and the carbon atoms of the surrounding DCM solvent molecules. Data was obtained during a 300 ns production run under NVT conditions and is presented for the first simulation of three replicates.

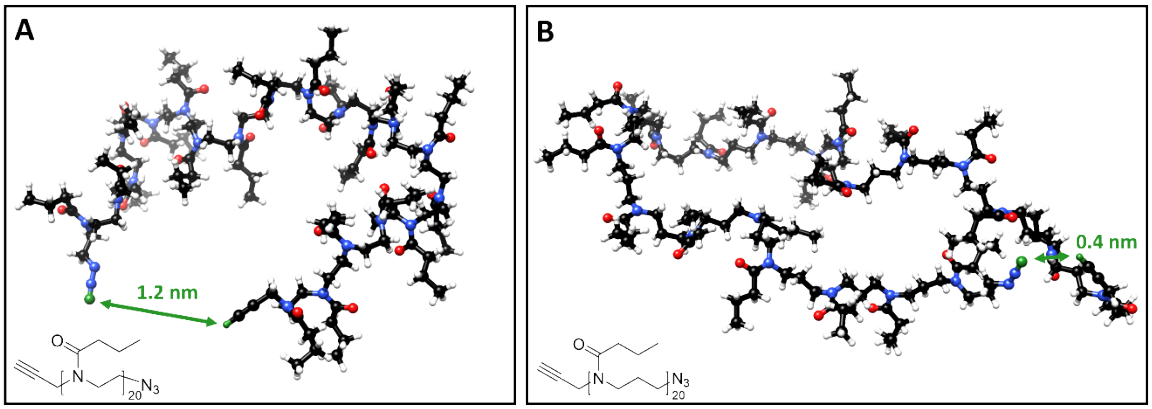

Fig. S6 Comparison of the conformations of (A) $PnPropOx_{20}$ and (B) $PnPropOzi_{20}$ during the production run in DCM, at which the polymer termini (green) are in closest proximity to each other. Snapshots were taken at 164.6 ns ($PnPropOx_{20}$) and 234.0 ns ($PnPropOzi_{20}$), respectively. Polymer conformations are shown for the first simulation of three replicates.

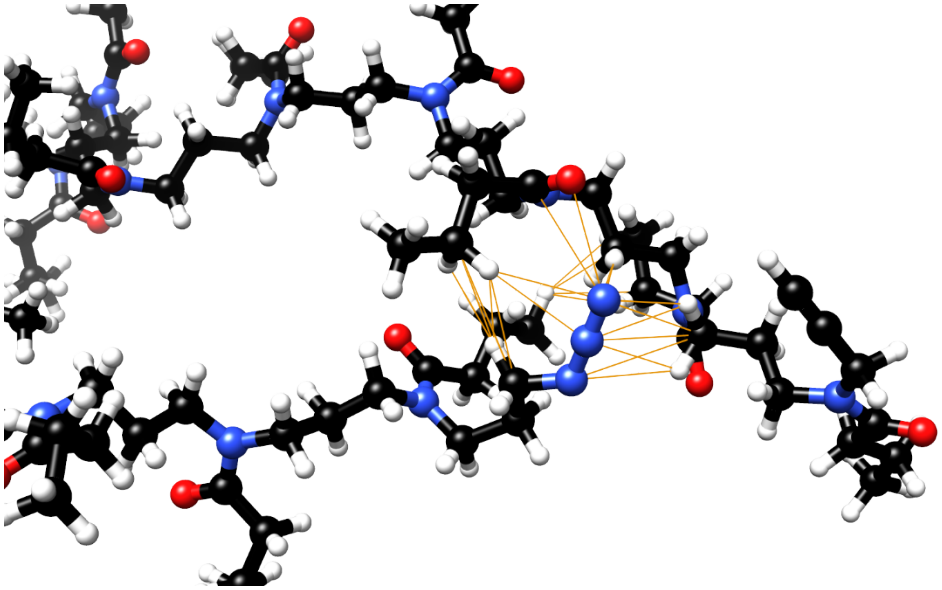

Fig. S7 Magnification into the closed conformation representative of PnPropOzi$_{20}$ at 234.0 ns. Due to the close proximity of the polymer chain ends to each other, non-bonded interactions (orange, van der Waals overlap ≥ -0.4 Å) between the azide group and the polymer backbone/side chains of the opposing terminus are possible. Further interactions found are between opposite backbone and side chains moieties. The polymer conformation is shown for the first simulation of three replicates.

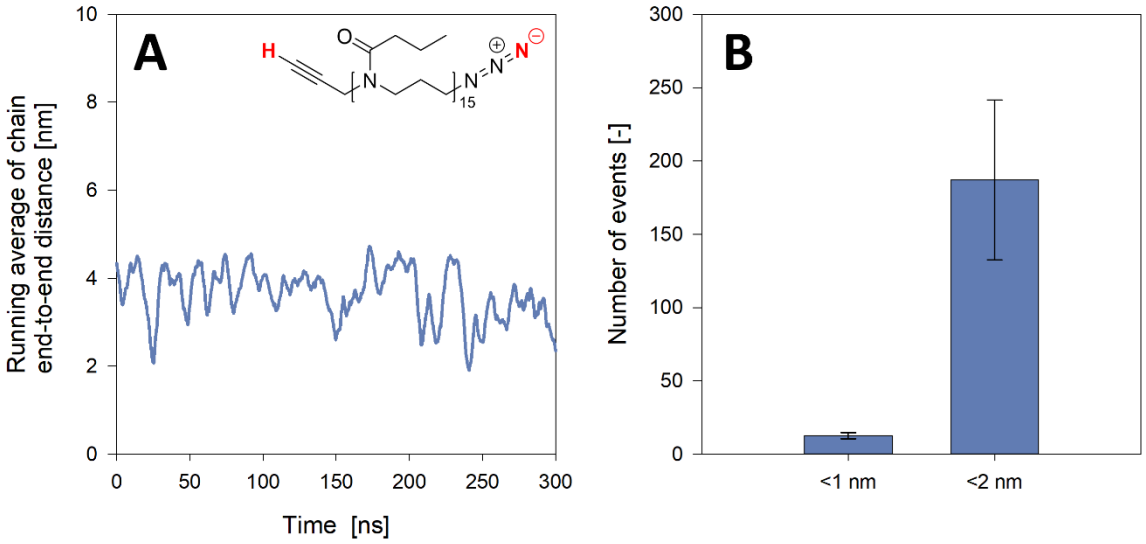

Fig. S8 (A) Running average of the PnPropOzi$_{15}$ chain end-to-end distance as a function of time. Polymer termini (red) are defined as the terminal hydrogen atom of the alkyne group and the terminal nitrogen atom on the azide moiety. The data represents the first simulation of three replicates. (B) Number of events at which the distance between polymer termini reaches below 2 (187 ± 55 times) and 1 nm (13 ± 2 times) distance. This value is an average of the events that occurred during three replicate simulations.

3. References

1 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminform.*, 2012, **4**, 17.

2 M. YABE, K. MORI, K. UEDA and M. TAKEDA, *J. Comput. Chem. Japan -International Ed.*, , DOI:10.2477/jccjie.2018-0034.

3 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

4 B. Hess, C. Kutzner, D. Van Der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.

5 E. Lindahl, B. Hess and D. van der Spoel, *J. Mol. Model.*, 2001, **7**, 306–317.

6 S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. Van Der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845–854.

7 H. J. C. Berendsen, D. van der Spoel and R. van Drunen, *Comput. Phys. Commun.*, 1995, **91**, 43–56.

8 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.

9 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**–**2**, 19–25.

10 S. Páll, M. J. Abraham, C. Kutzner, B. Hess and E. Lindahl, eds. S. Markidis and E. Laure, Springer International Publishing, Cham, 2015, pp. 3–27.

11 E. Lindahl, M. J. Abraham, B. Hess and D. Van Der Spoel, , DOI:10.5281/ZENODO.3243833.

12 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.

13 S. Nosé and M. L. Klein, *Mol. Phys.*, 1983, **50**, 1055–1076.

14 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.

15 W. L. Jorgensen and J. Tirado-Rives, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6665–6670.

16 L. S. Dodda, I. C. De Vaca, J. Tirado-Rives and W. L. Jorgensen, *Nucleic Acids Res.*, 2017, **45**, W331–W336.

17 L. S. Dodda, J. Z. Vilseck, J. Tirado-Rives and W. L. Jorgensen, *J. Phys. Chem. B*, 2017, **121**, 3864–3870.

18 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

19 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.