SUPPORTING INFORMATION

Cumulative sub-millisecond all-atom simulations of the temperature-induced coil-to-globule transition of poly(*N*-vinylcaprolactam) in aqueous solution

Jonas Dittrich[†,‡], Michael Kather[#,Δ], Anna Holzberger[#,Δ], Andrij Pich[‡,#,Δ], Holger Gohlke*[,†,‡,§]

[†]Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

[‡]Bioeconomy Science Center (BioSC), Forschungszentrum Jülich, 52425 Jülich, Germany

[#]DWI-Leibniz-Institute for Interactive Materials, RWTH Aachen University, 52056 Aachen, Germany

[Δ]Institute of Technical and Macromolecular Chemistry, RWTH Aachen University, 52074 Aachen, Germany

[§]John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany.

Author ORCID

Jonas Dittrich: 0000-0003-2377-2268

Michael Kather: 0000-0001-6139-3483

Anna Holzberger: 0000-0002-9748-240X

Andrij Pich: 0000-0003-1825-7798

Holger Gohlke: 0000-0001-8613-1447

*Corresponding Author: Holger Gohlke

Address: Universitätsstr. 1, 40225 Düsseldorf, Germany.

Phone: (+49) 211 81 13662; Fax: (+49) 211 81 13847

E-mail: gohlke@uni-duesseldorf.de

Table of Contents

Supplementary Figures

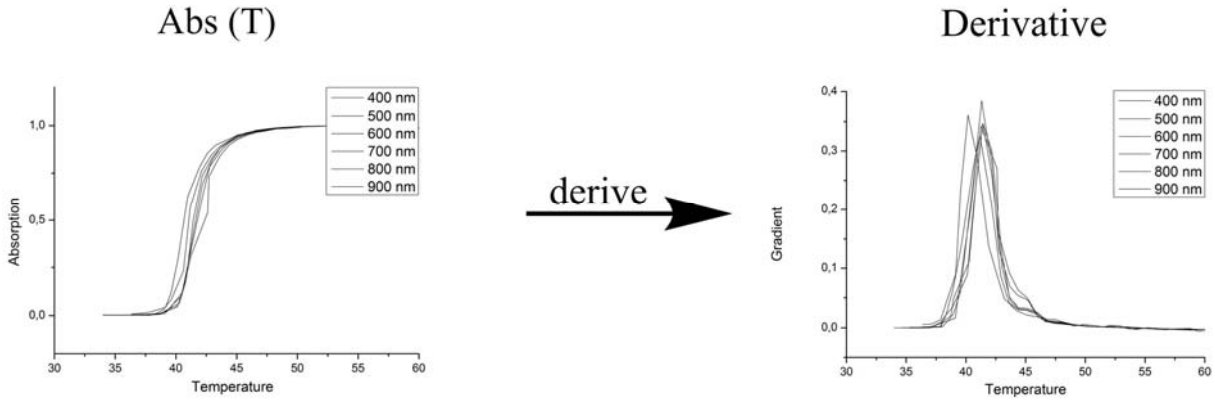

Figure S1. Cloud points were determined by measuring the absorption of the linear polymer in dependence of the temperature at various wavelengths. The first derivative was determined from these measurements, and the mean value of these peaks yields the respective cloud point. This example shows the cloud point determination of the 39mer at a polymer concentration of 0.3 wt.-%.

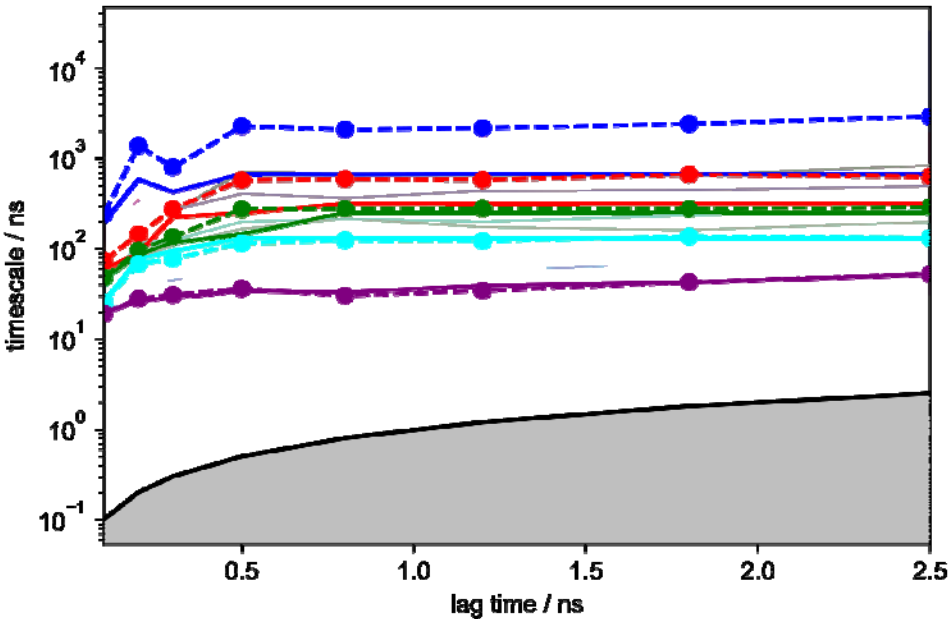

Figure S2. Implied timescale plot for HMMs constructed for five assumed hidden states (indicated by color) at different lag times up to 2.5 ns. The constructed HMM describes the coil-to-globule transition of a PNVCL 40mer at 293 K. It is constructed from 5x1 µs MD simulation. Errors are calculated using Bayesian error estimation and are depicted as corresponding error area. The maximum likelihood result is depicted as solid lines, and the dashed lines show the ensemble mean based on a Bayesian sampling procedure Calculations were performed with the PyEMMA[1] python library (v. 2.5.6).

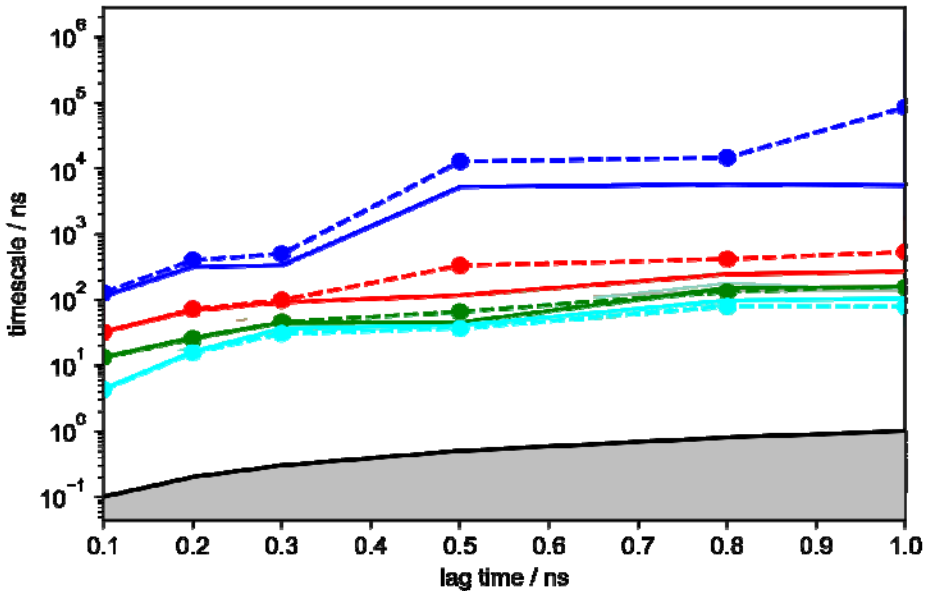

Figure S3. Implied timescale plot for HMMs constructed for five assumed hidden states (indicated by color) at different lag times up to 1 ns. The constructed HMM describes the coil-to-globule transition of a PNVCL 40mer at 313 K. The HMM is constructed from 5x1 µs MD simulation data. Errors are calculated using Bayesian error estimation and are depicted as corresponding error area. The maximum likelihood result is depicted as solid lines, and the dashed lines show the ensemble mean based on a Bayesian sampling procedure Calculations were performed with the PyEMMA[1] python library (v. 2.5.6).

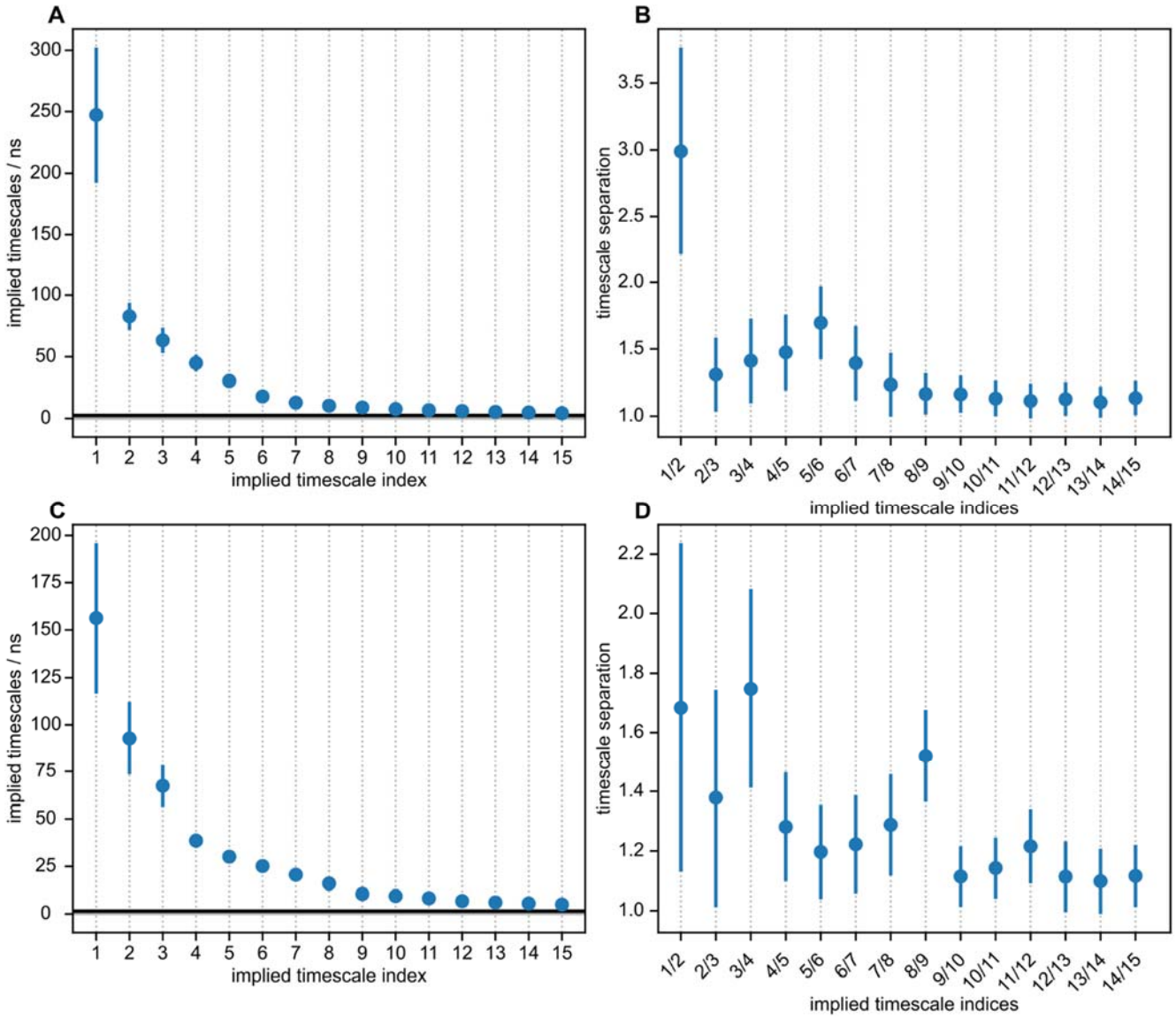

Figure S4: Implied timescales (**A**,**C**) and timescale separation (**B**,**D**) of MSMs constructed from 5x1 µs MD simulation of the coil-to-globule transition of a PNVCL 40mer at 293 K (**A**,**B**) and 313 K **(C,D)**. The MSMs were constructed with a lag time of 2.5 ns and 1 ns for 293 K and 313 K, respectively. We consider two/three hidden states as there are gaps between the first and second (**B**) and between the second and third (**D**) relaxation timescale for 293 K and 313 K, respectively. Calculations were performed with the PyEMMA[1] python library (v. 2.5.6).

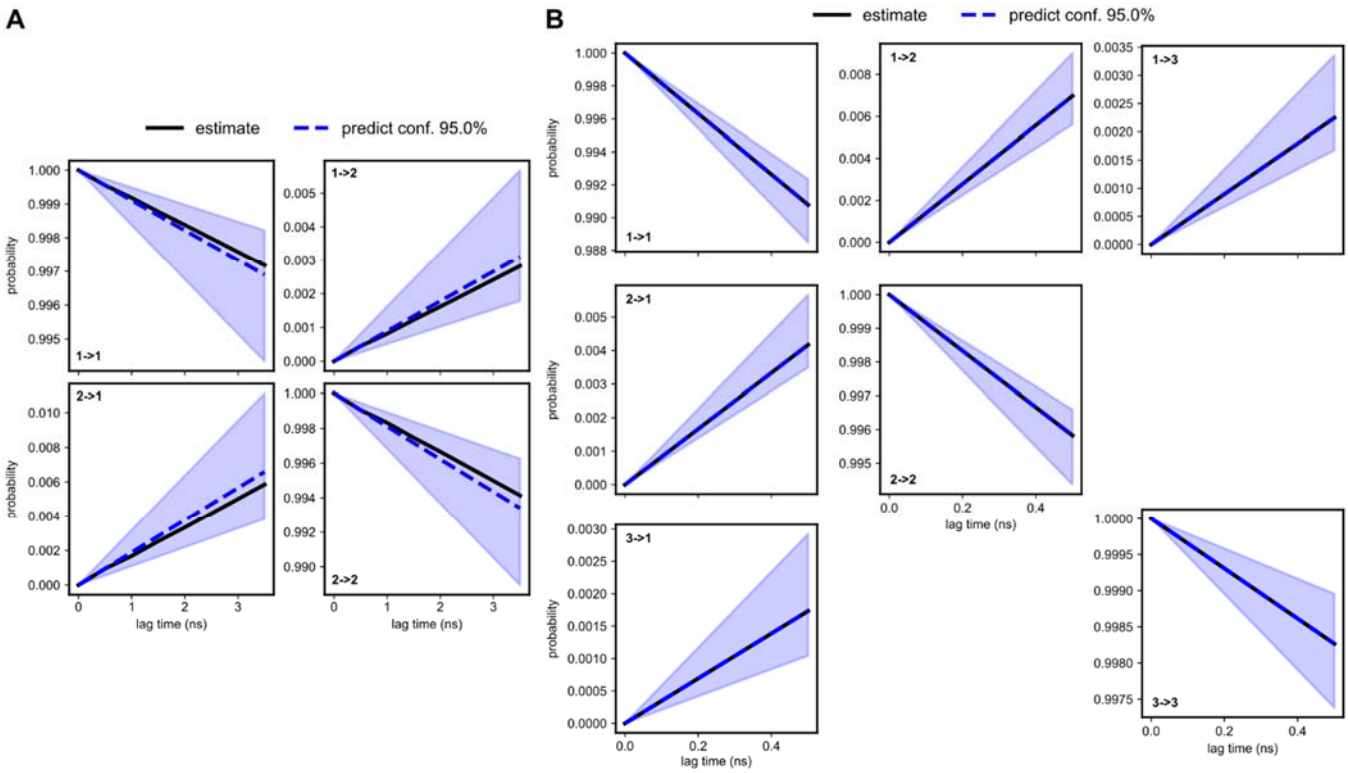

Figure S5. Chapman-Kolmogorov test for the constructed HMM describing the coil-to-globule transition of a PNVCL 40mer at 293 K **(A)** and 313 K **(B)**. The HMMs are constructed from 5x1 µs MD simulation data each. Estimated state transition probabilities and corresponding errors (at a 95 % confidence level) are depicted as dotted blue lines and error areas, respectively. Calculations were performed with the PyEMMA[1] python library (v. 2.5.6). In panel (B), black and blue lines lie on top of each other.

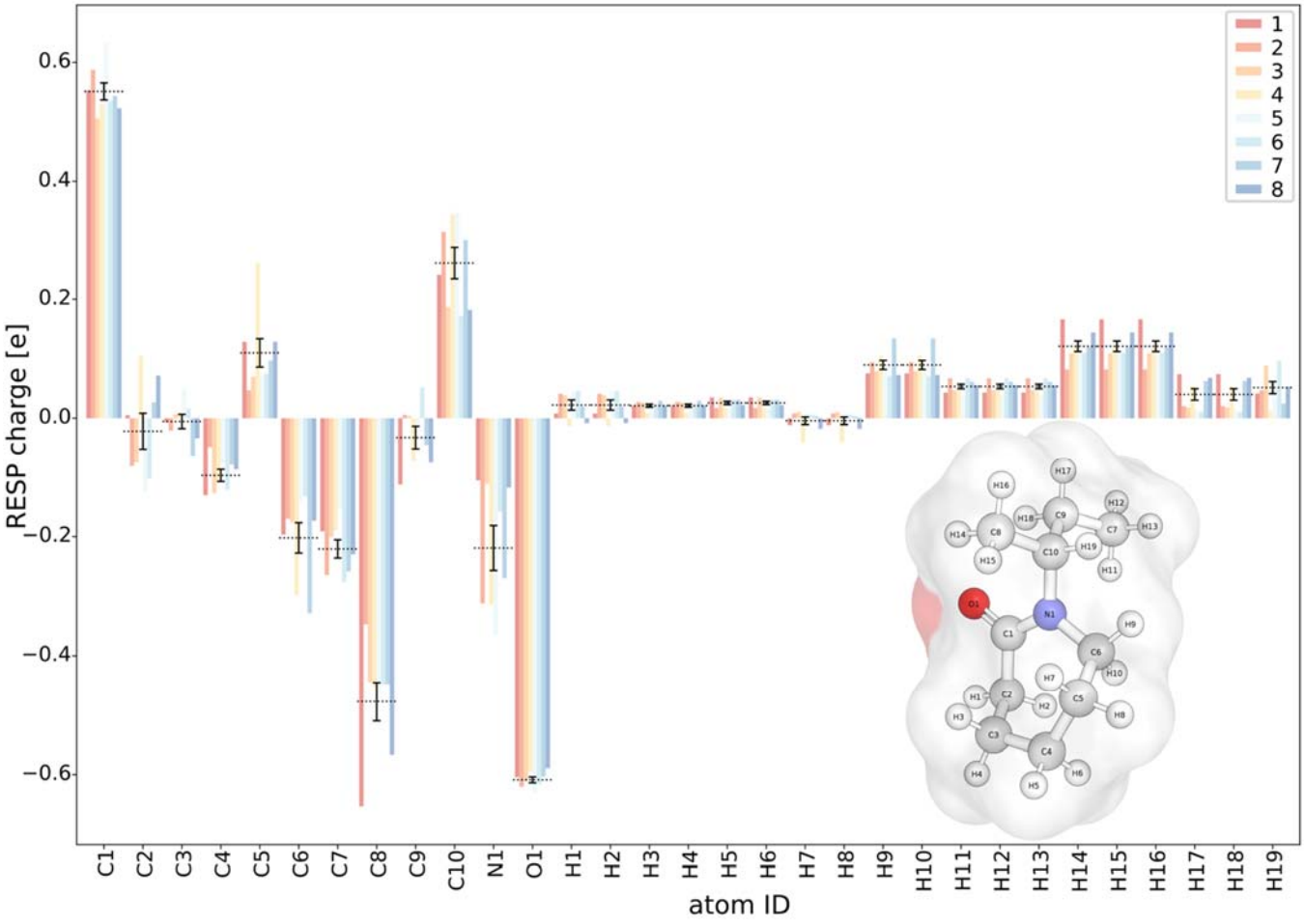

Figure S6. Multi-conformer single point charge fit for methyl-capped *N*-(vinylcaprolactam) monomeric units. The charges were calculated at the HF/6-31G* level using Gaussian09[2]. Charge fitting was done using the RESP charge fitting procedure implemented in antechamber.[3] The mean value is depicted as dotted horizontal line, and the standard error is shown by error bars. The inset depicts an *N*-(vinylcaprolactam) monomeric unit with partial charges mapped onto the molecular surface.

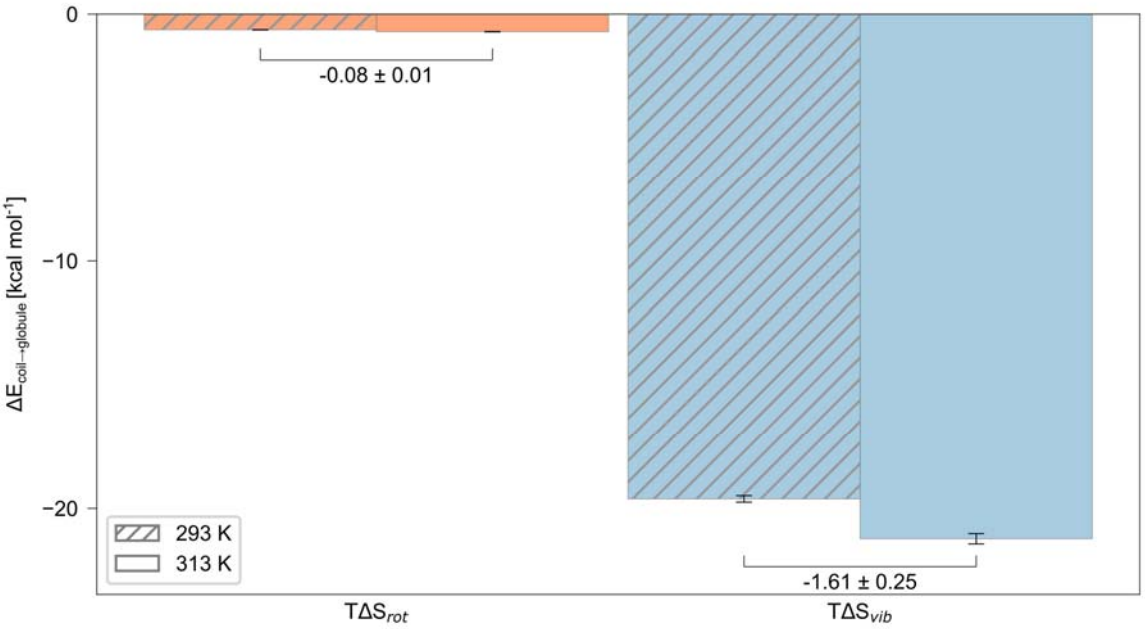

Figure S7. Energy components of $T\Delta S_{config}$ computed with MM-PBSA for the coil-to-globule transition of a PNVCL 40mer in water at 293 K (hatched) and at 313 K (solid). The bar depicts $TS_{\{rot,\ vib\},globule}$ - $TS_{\{rot,\ vib\},coil}$. The difference in entropy for the coil-to-globule transition is mainly dominated by the loss in vibrational entropy. It is a result of both a higher number of configurational degrees of freedom of the coil state and the presence of a compact state of PNVCL at 313 K. Mean values were computed from ten MM-PBSA calculations of 100 MD snapshots of PNVCL in either coil or globule conformation; the error is calculated following the laws of error propagation using the respective standard error of the mean. Calculations were performed using MMPBSA.py[4].

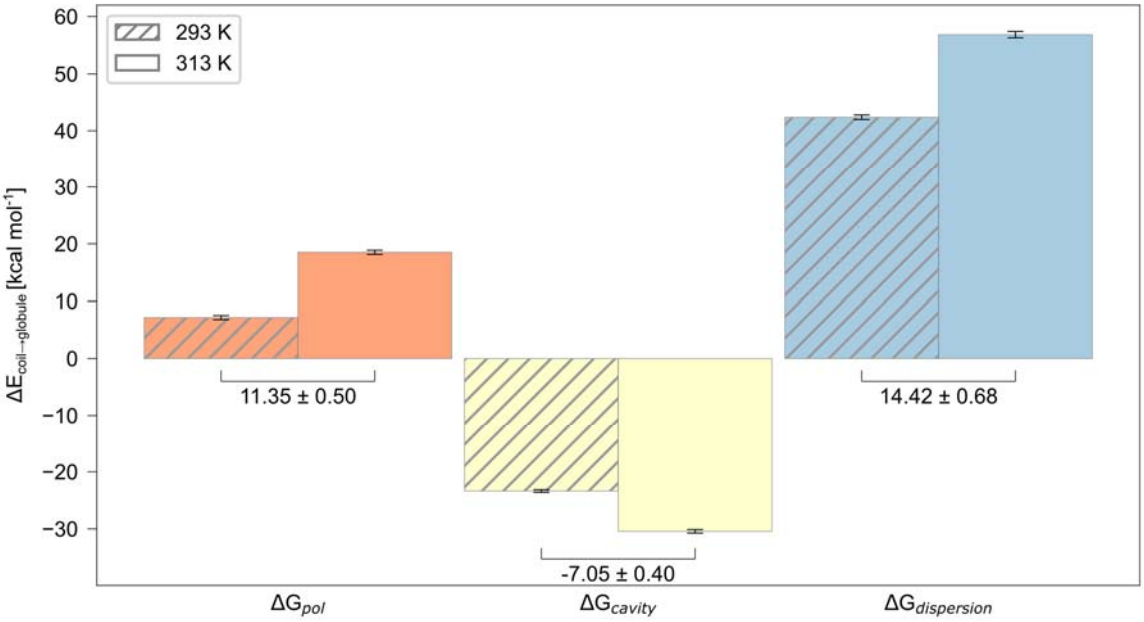

Figure S8. Energy components of ΔG_{solv} computed with MM-PBSA for the coil-to-globule transition of a PNVCL 40mer in water at 293 K (hatched) and at 313 K (solid). The bar depicts $G_{\{pol,\ cavity,\ dispersion\},globule}$ - $G_{\{pol,\ cavity,\ dispersion\},coil}$. Mean values were computed from ten MM-PBSA calculations of 100 MD snapshots of PNVCL in either coil or globule conformation; the error is calculated following the laws of error propagation using the respective standard error of the mean. Calculations were performed using MMPBSA.py[4].

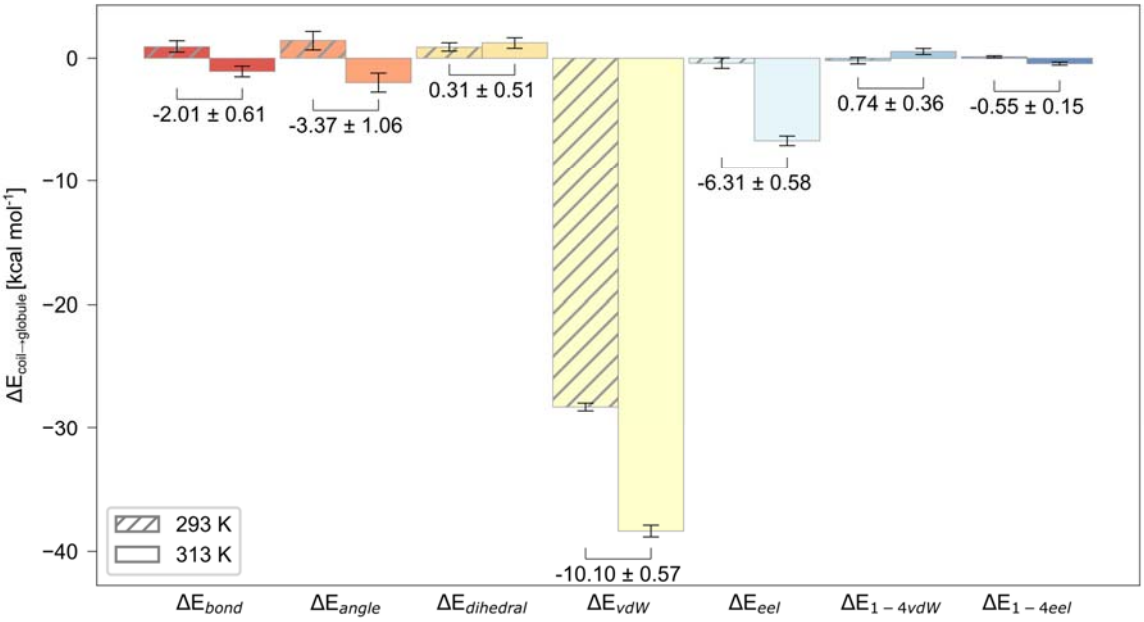

Figure S9. Energy components of ΔE_{MM} computed with MM-PBSA for the coil-to-globule transition of a PNVCL 40mer in water at 293 K (hatched) and at 313 K (solid). The bar depicts $E_{\{\text{bond, angle, dihedral, vdW, eel, 1-4 vdW, 1-4 eel}\},\text{globule}} - E_{\{\text{bond, angle, dihedral, vdW, eel, 1-4 vdW, 1-4eel}\},\text{coil}}$. Mean values were computed from ten MM-PBSA calculations of 100 MD snapshots of PNVCL in either coil or globule conformation; the error is calculated following the laws of error propagation using the respective standard error of the mean. Calculations were performed using MMPBSA.py [4].

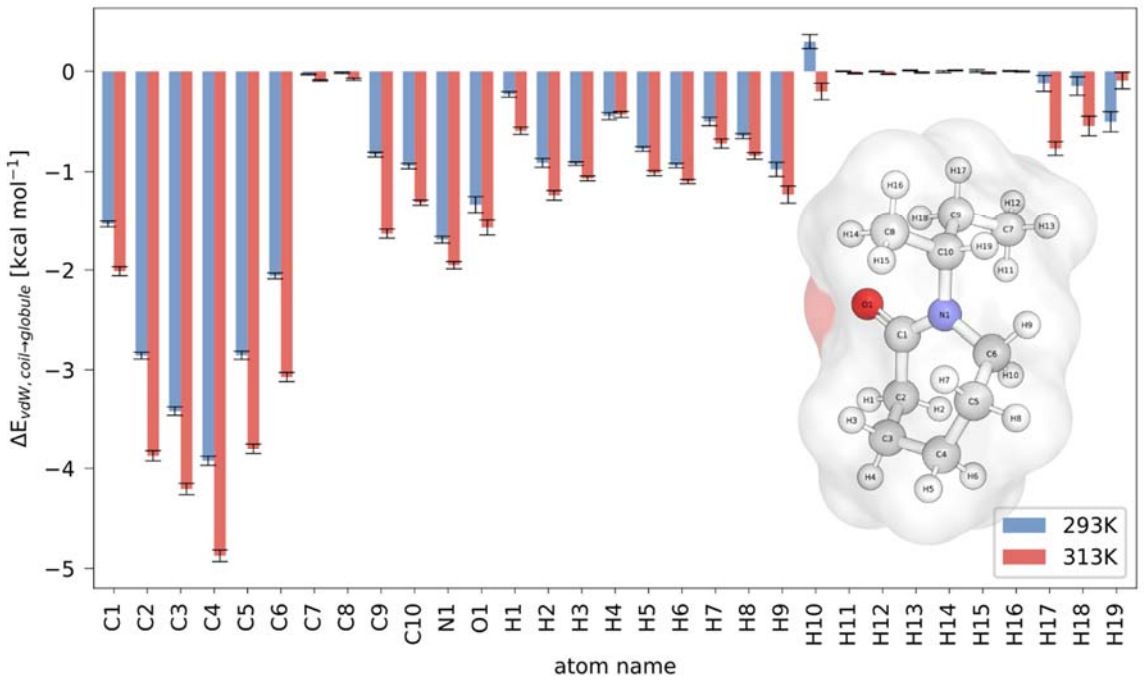

Figure S10. Atomwise energy decomposition of ΔE_{vdW} (see Figure S9) for the coil-to-globule transition of an atactic PNVCL 40mer simulated under NVT conditions at 293 K (blue) and 313 K (red). The carbon atoms of the caprolactam ring, especially C_3 and C_4, contribute to the favorable increase in ΔE_{vdW} to a large extent, fostering the coil-to-globule transition at 293 K and even more at 313 K.

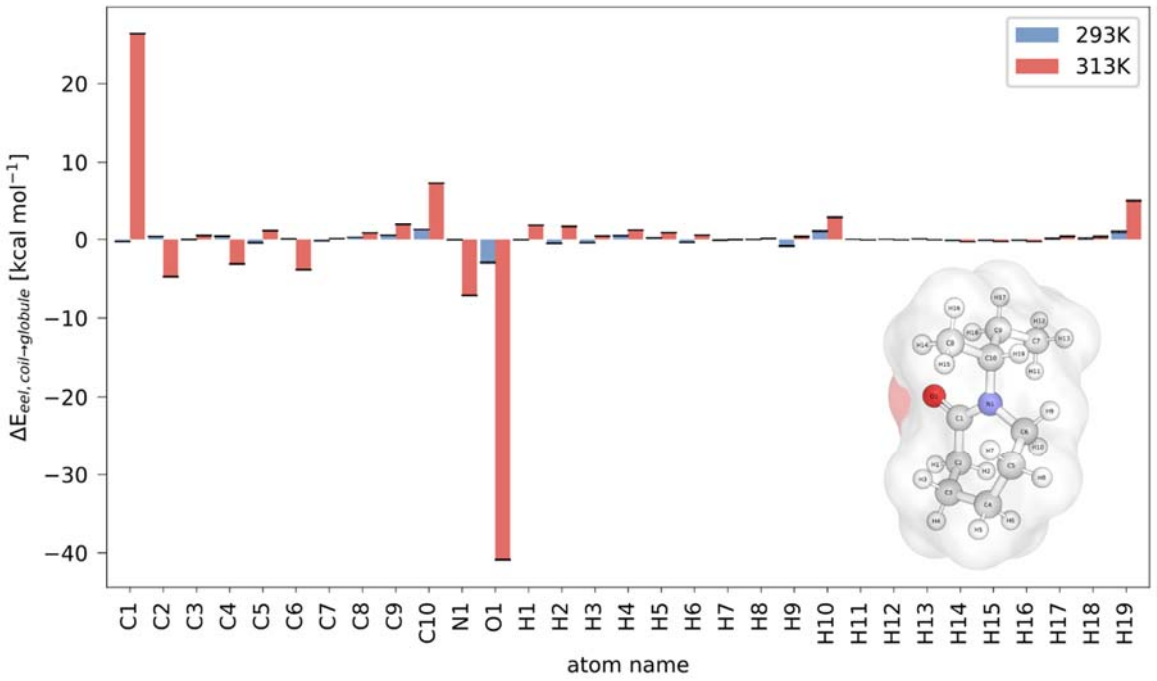

Figure S11. Atomwise energy decomposition of ΔE_{eel} (see Figure S9) for the coil-to-globule transition of an atactic PNVCL 40mer simulated under NVT conditions at 293 K (blue) and 313 K (red). The unfavorable increase in the ΔE_{eel} term for the lactam carbonyl carbon is outbalanced by a favorable interaction of the carbonyl oxygen, leading to a favorable contribution to the coil-to-globule transition at 313 K.

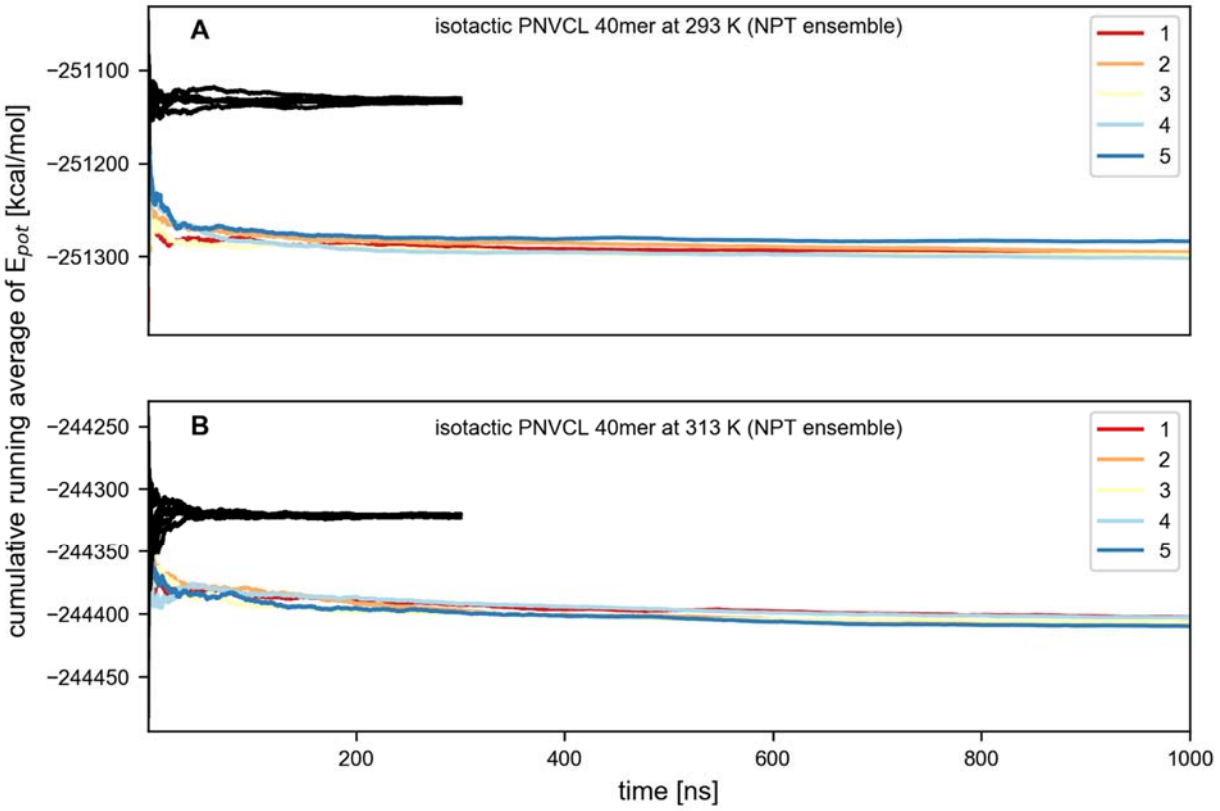

Figure S12. Running cumulative average of the potential energy of the 5x1 µs NPT simulations (depicted by color) of the isotactic PNVCL 40mer at 293 K (**A**) and 313 K (**B**). The running cumulative averages of the potential energy of the pure water system (5x300 ns) are shown as black lines. The enthalpy of the polymer part of the polymer/water system is -161.9 ± 3.3 kcal mol^{-1} and -85.1 ± 1.4 kcal mol^{-1} at 293 K and 313 K, respectively. Therefore, the transition enthalpy ΔH = 76.9 ± 3.6 kcal mol^{-1} for the 40mer, which is equivalent to 1.92± 0.56 kcal mol^{-1} per repeating unit.

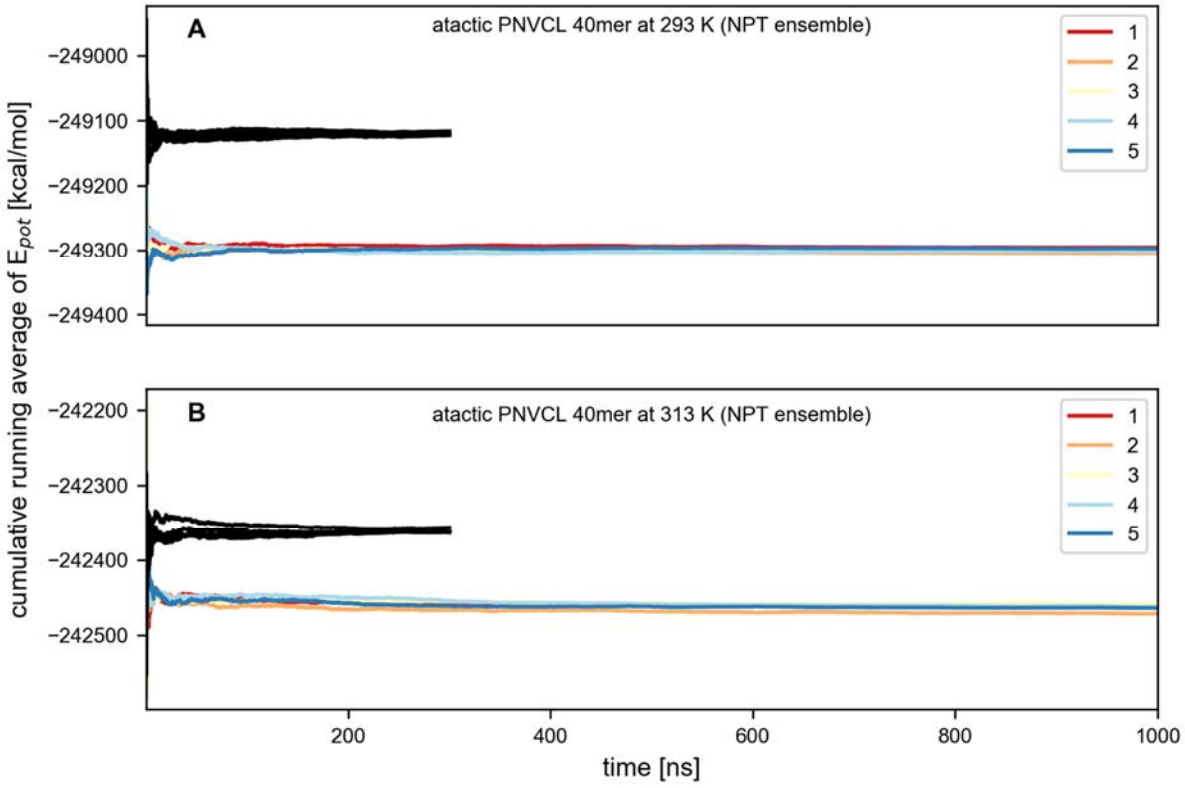

Figure S13. Running cumulative average of the potential energy of the 5x1 µs NPT simulations (depicted by color) of the atactic PNVCL 40mer at 293 K (**A**) and 313 K (**B**). The running cumulative averages of the potential energy of the pure water system (5x300 ns) are shown as black lines. The enthalpy of the polymer part of the polymer/water system is -179.7 ± 1.7 kcal mol^{-1} and -101.7 ± 2.6 kcal mol^{-1} at 293 K and 313 K, respectively. Therefore, the transition enthalpy ΔH = 78.0 ± 3.1 kcal mol^{-1} for the 40mer, which is equivalent to 1.95 ± 0.51 kcal mol^{-1} per repeating unit.

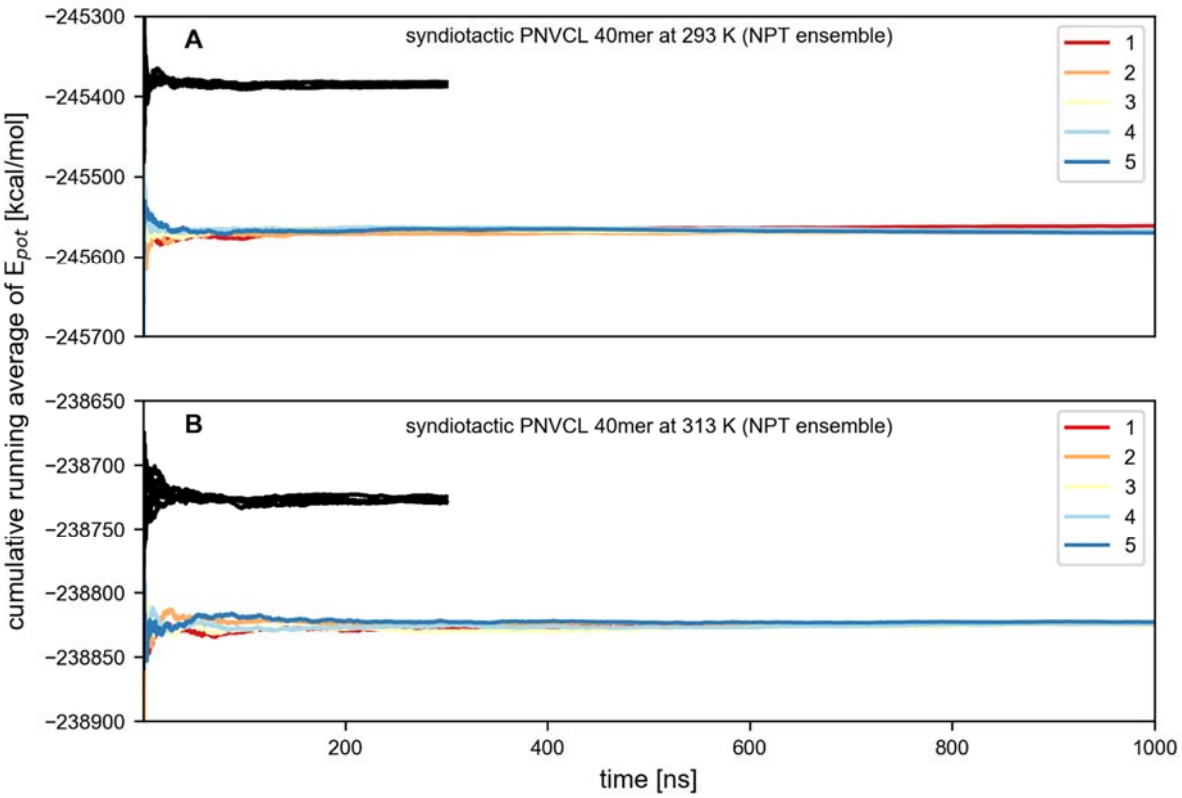

Figure S14. Running cumulative average of the potential energy of the 5x1 µs NPT simulations (depicted by color) of the syndiotactic PNVCL 40mer at 293 K (**A**) and 313 K (**B**). The running cumulative averages of the potential energy of the pure water system (5x300 ns) are shown as black lines. The enthalpy of the polymer part of the polymer/water system is -183.1 ± 1.9 kcal mol^{-1} and -96.4 ± 0.9 kcal mol^{-1} at 293 K and 313 K, respectively. Therefore, the transition enthalpy ΔH = 86.6 ± 2.1 kcal mol^{-1}for the 40mer, which is equivalent to 2.17 ± 0.33 kcal mol^{-1} per repeating unit.

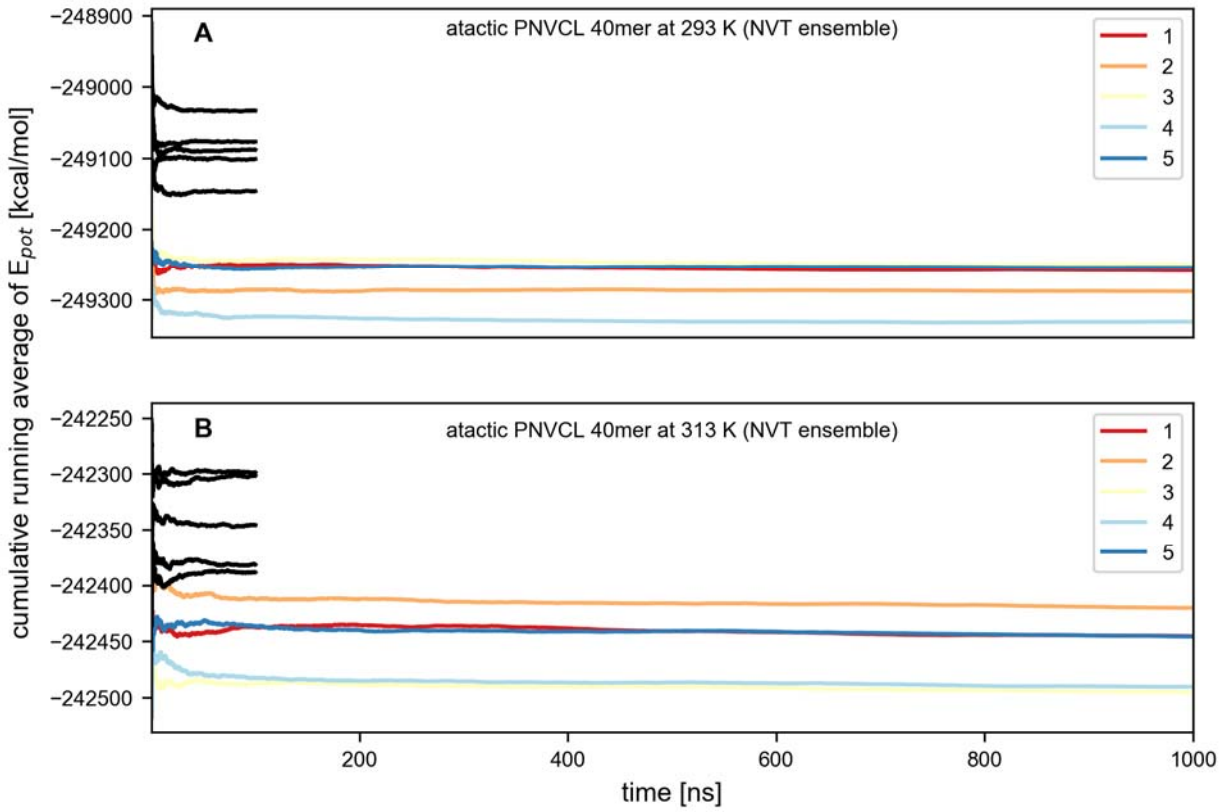

Figure S 15. Running cumulative average of the potential energy of the 5x1 µs NVT simulations (depicted by color) of the atactic PNVCL 40mer at 293 K (**A**) and 313 K **(B)**. The running cumulative averages of the potential energy of the pure water system (5x100 ns) are depicted as black lines. The potential energy of the system is linearly dependent on the system's volume, indicating that the volume was not completely adjusted to the equilibrium value in the NPT step during thermalization.

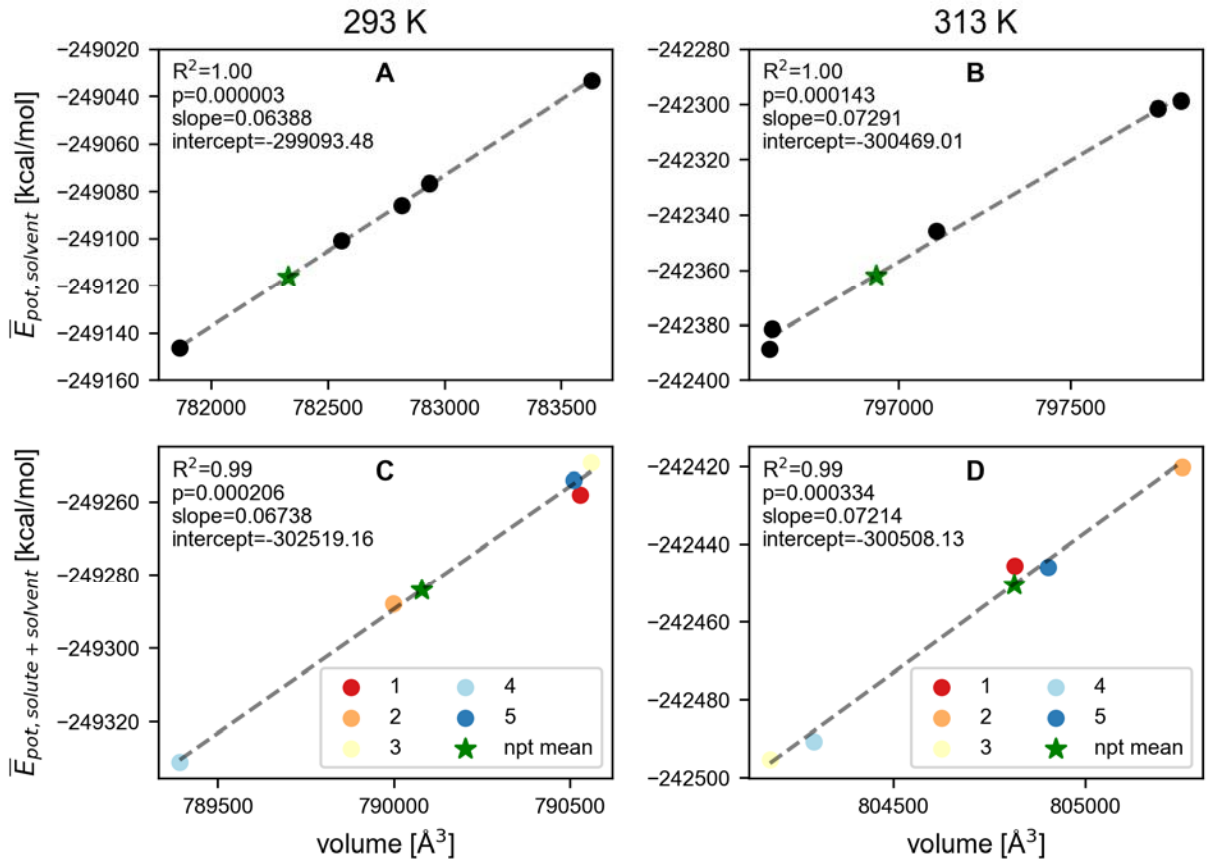

Figure S16. Mean potential energies of pure water systems (**A**,**B**) and the atactic PNVCL 40mer (**C**,**D**) in water systems for MD simulations in the NVT ensemble at 293 K (**A**,**C**) and 313 K(**B**,**D**) as a function of the simulation box size for five independent replicas (depicted by color for the PNVCL). Green stars depict the interpolated potential energy of the system using the linear regression based on the five mean potential energies of the NVT simulations (determined parameters and statistics for the regression are displayed in each panel) and the mean volumes of the corresponding NPT simulations. The transition enthalpy is calculated according to eq. (7) of the main text and yields ΔH = 1.98 kcal mol^{-1} per repeating unit for the NVT ensemble.

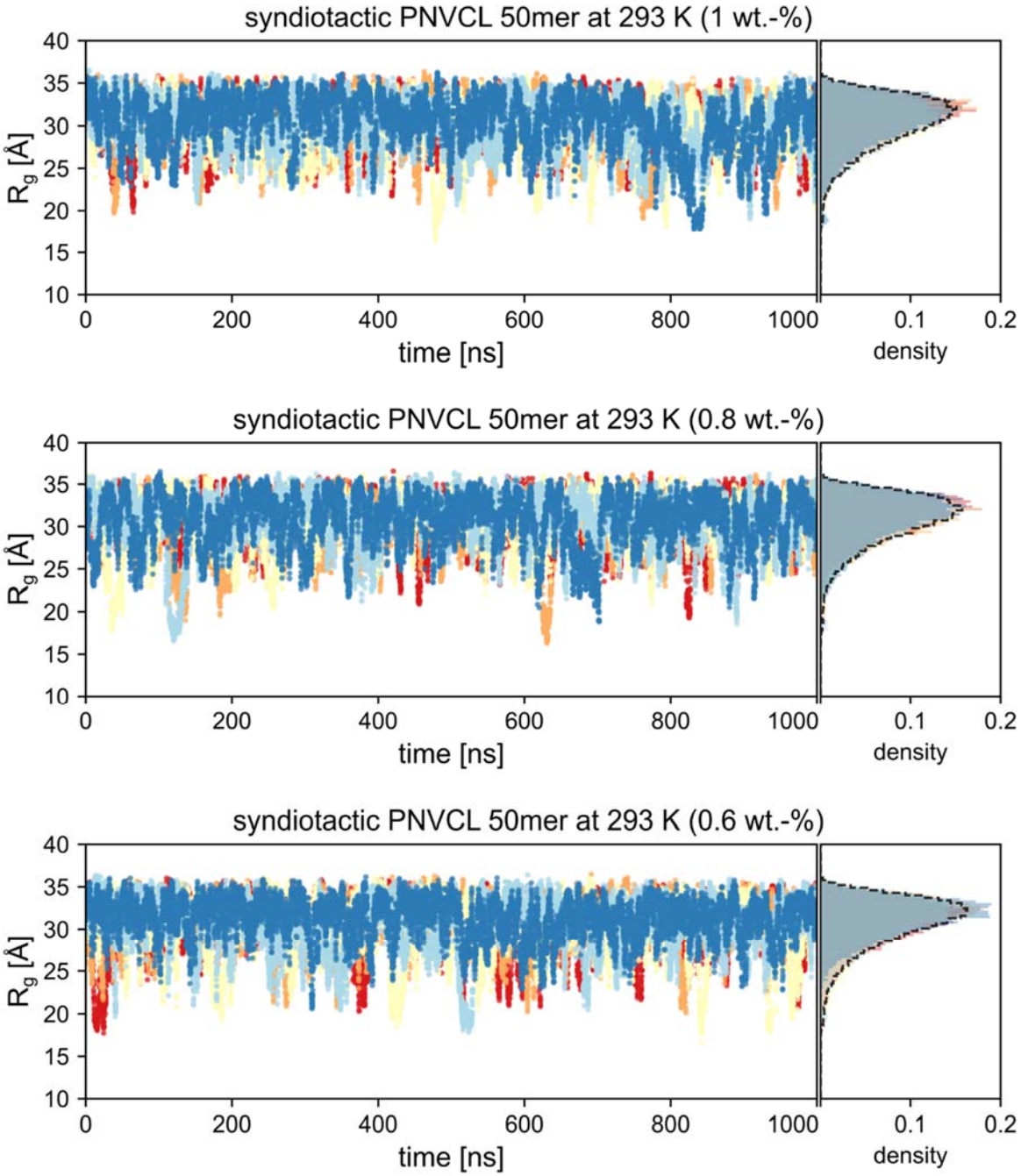

Figure S17. Radius of gyration (R_g) during five·MD simulations of 1 µs length of a syndiotactic PNVCL 50mer at 293 K for polymer concentrations of 1 wt.-%, 0.8 wt.-%, and 0.6 wt.-% (top to bottom). Corresponding frequency distributions are shown next to the time series in matching color, a frequency distribution of all data is shown as dashed black line.

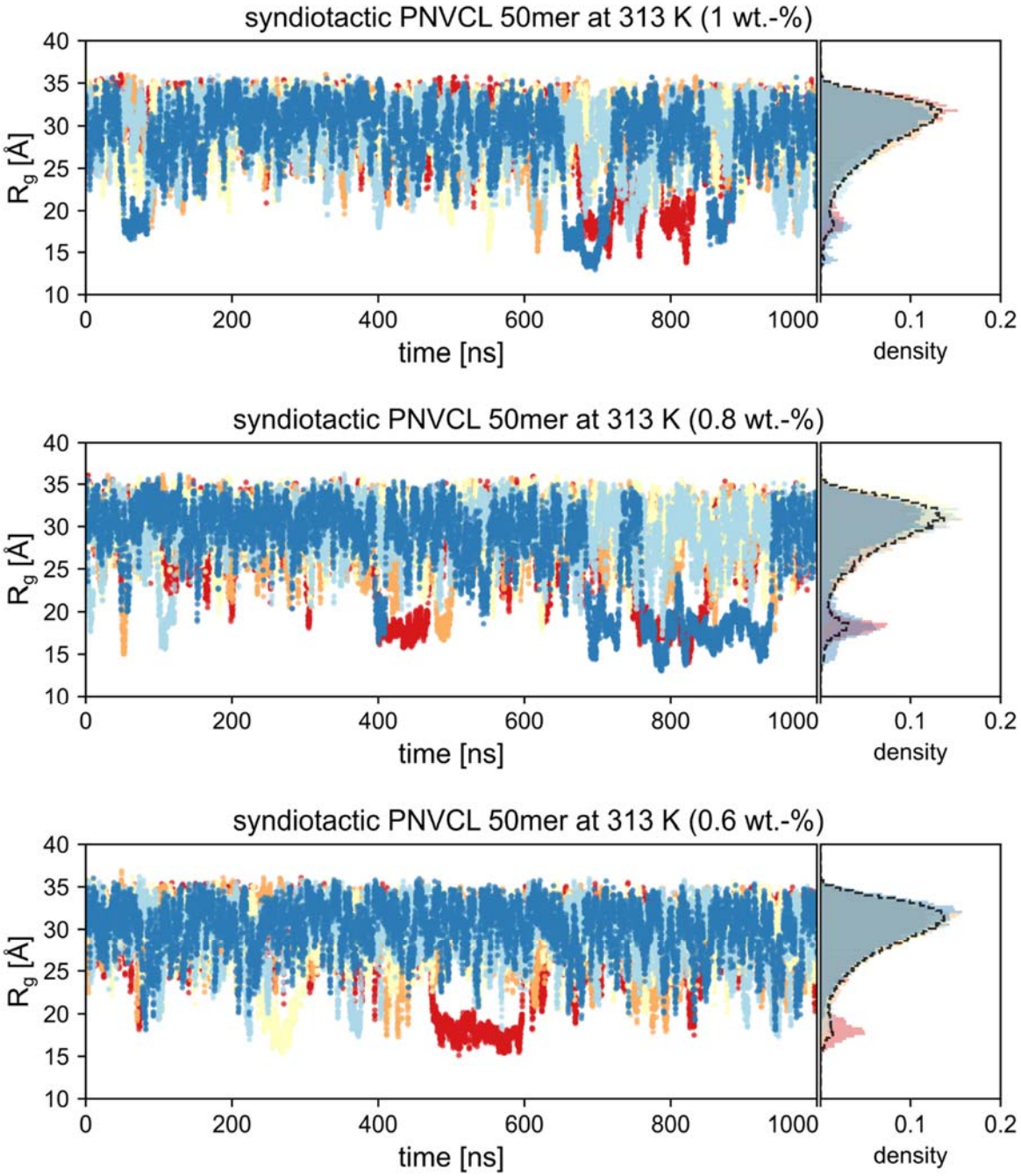

Figure S18. Radius of gyration (R_g) during five·MD simulations of 1 µs length of a syndiotactic PNVCL 50mer at 313 K for polymer concentrations of 1 wt.-%, 0.8 wt.-%, and 0.6 wt.-% (top to bottom). Corresponding frequency distributions are shown next to the time series in matching color, a frequency distribution of all data is shown as dashed black line.

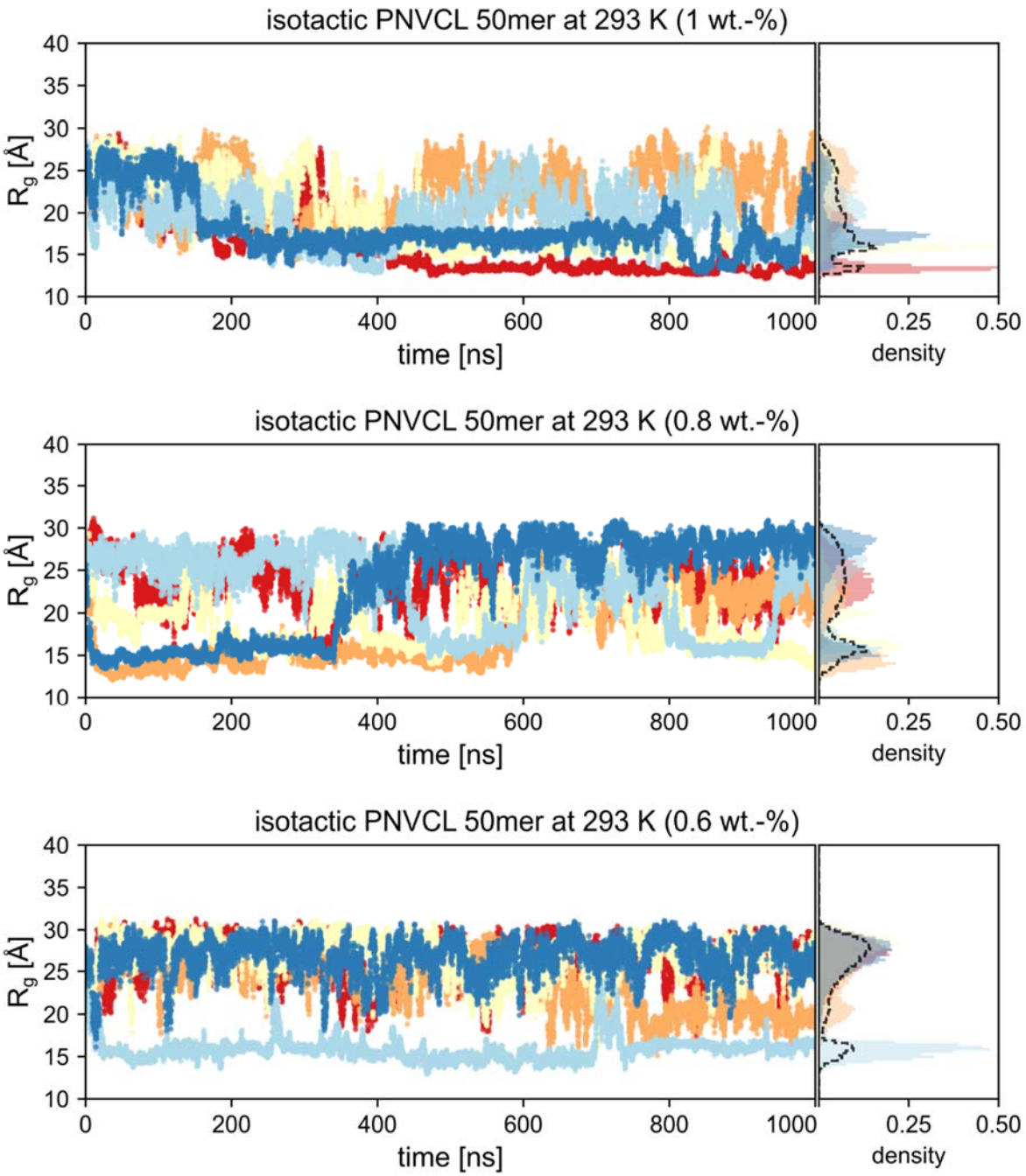

Figure S19. Radius of gyration (R_g) during five·MD simulations of 1 µs length of an isotactic PNVCL 50mer at 293 K for polymer concentrations of 1 wt.-%, 0.8 wt.-%, and 0.6 wt.-% (top to bottom). Corresponding frequency distributions are shown next to the time series in matching color, a frequency distribution of all data is shown as dashed black line.

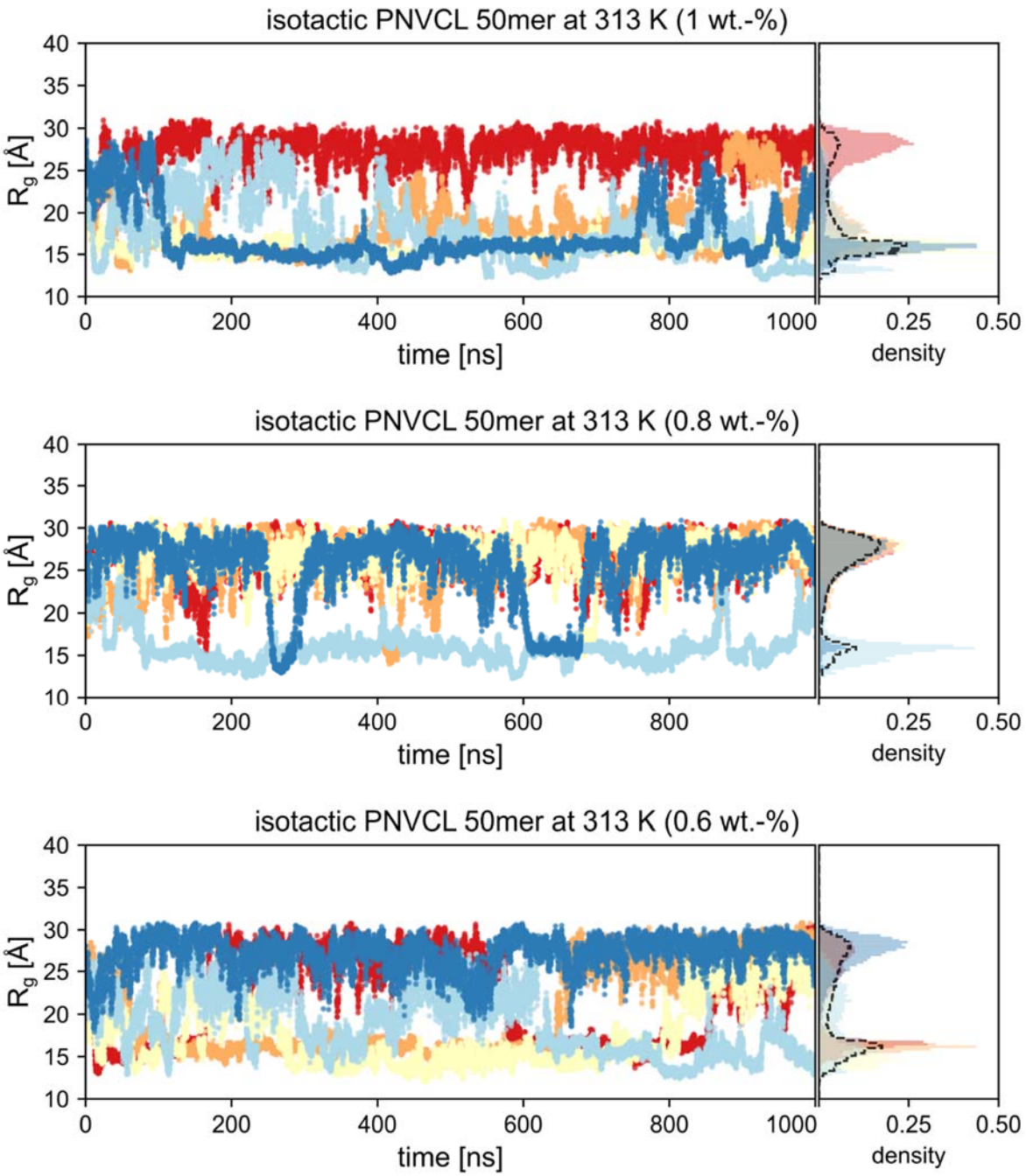

Figure S20. Radius of gyration (R_g) during five·MD simulations of 1 µs length of an isotactic PNVCL 50mer at 313 K for polymer concentrations of 1 wt.-%, 0.8 wt.-%, and 0.6 wt.-% (top to bottom). Corresponding frequency distributions are shown next to the time series in matching color, a frequency distribution of all data is shown as dashed black line.

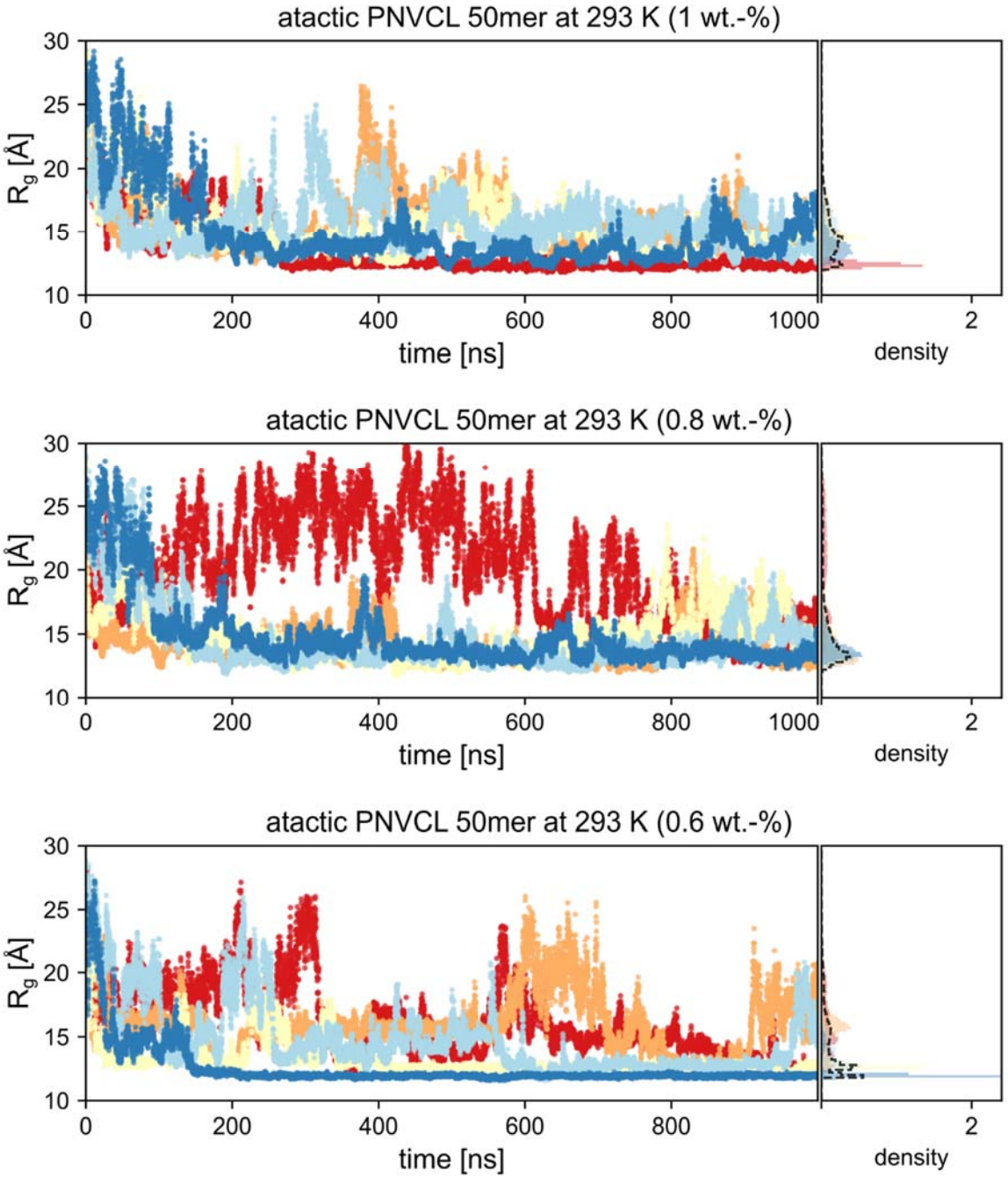

Figure S21. Radius of gyration (R_g) during five·MD simulations of 1 µs length of an atactic PNVCL 50mer at 293 K for polymer concentrations of 1 wt.-%, 0.8 wt.-%, and 0.6 wt.-% (top to bottom). Corresponding frequency distributions are shown next to the time series in matching color, a frequency distribution of all data is shown as dashed black line.

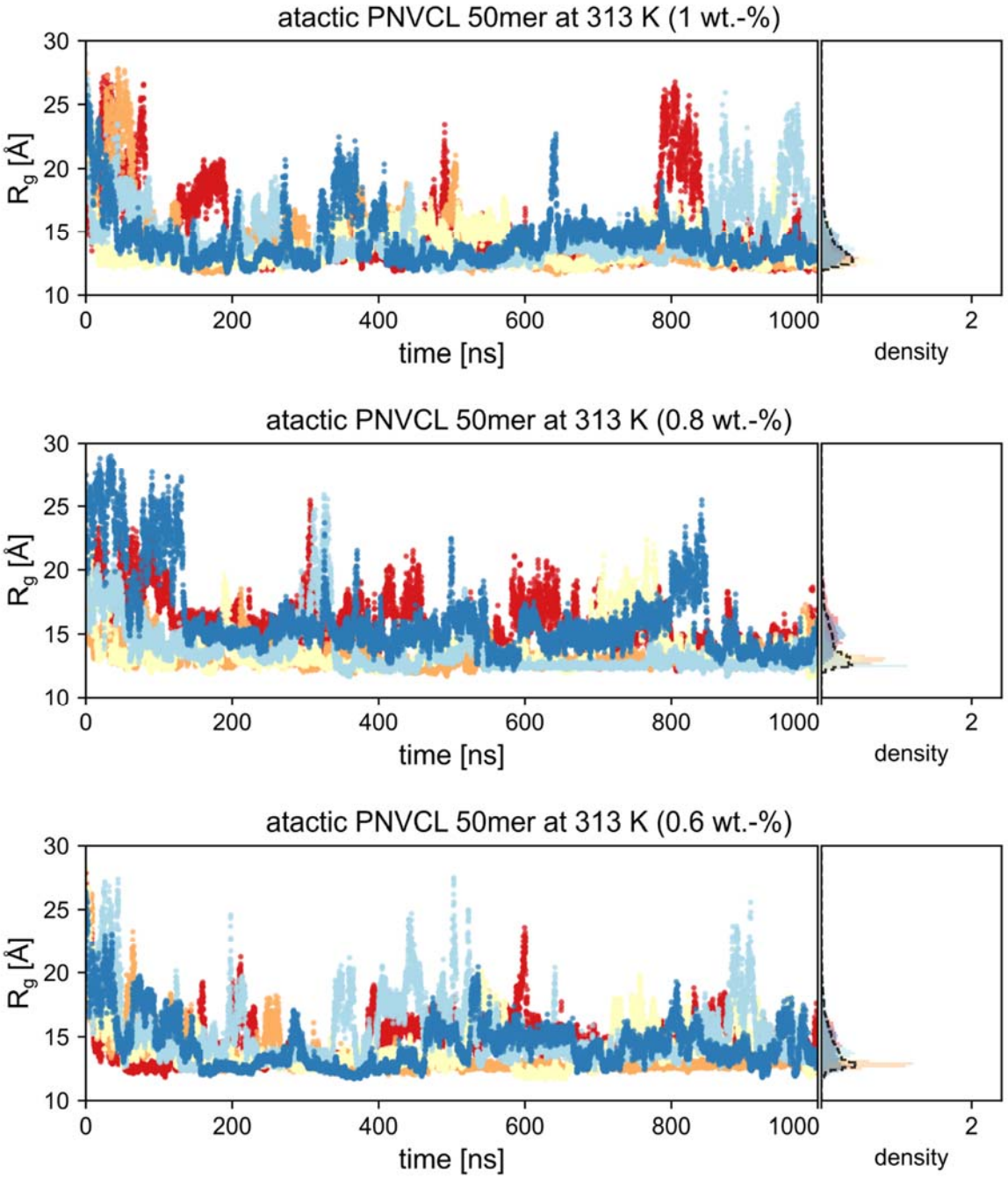

Figure S22. Radius of gyration (R_g) during five·MD simulations of 1 µs length of an atactic PNVCL 50mer at 313 K for polymer concentrations of 1 wt.-%, 0.8 wt.-%, and 0.6 wt.-% (top to bottom). Corresponding frequency distributions are shown next to the time series in matching color, a frequency distribution of all data is shown as dashed black line.

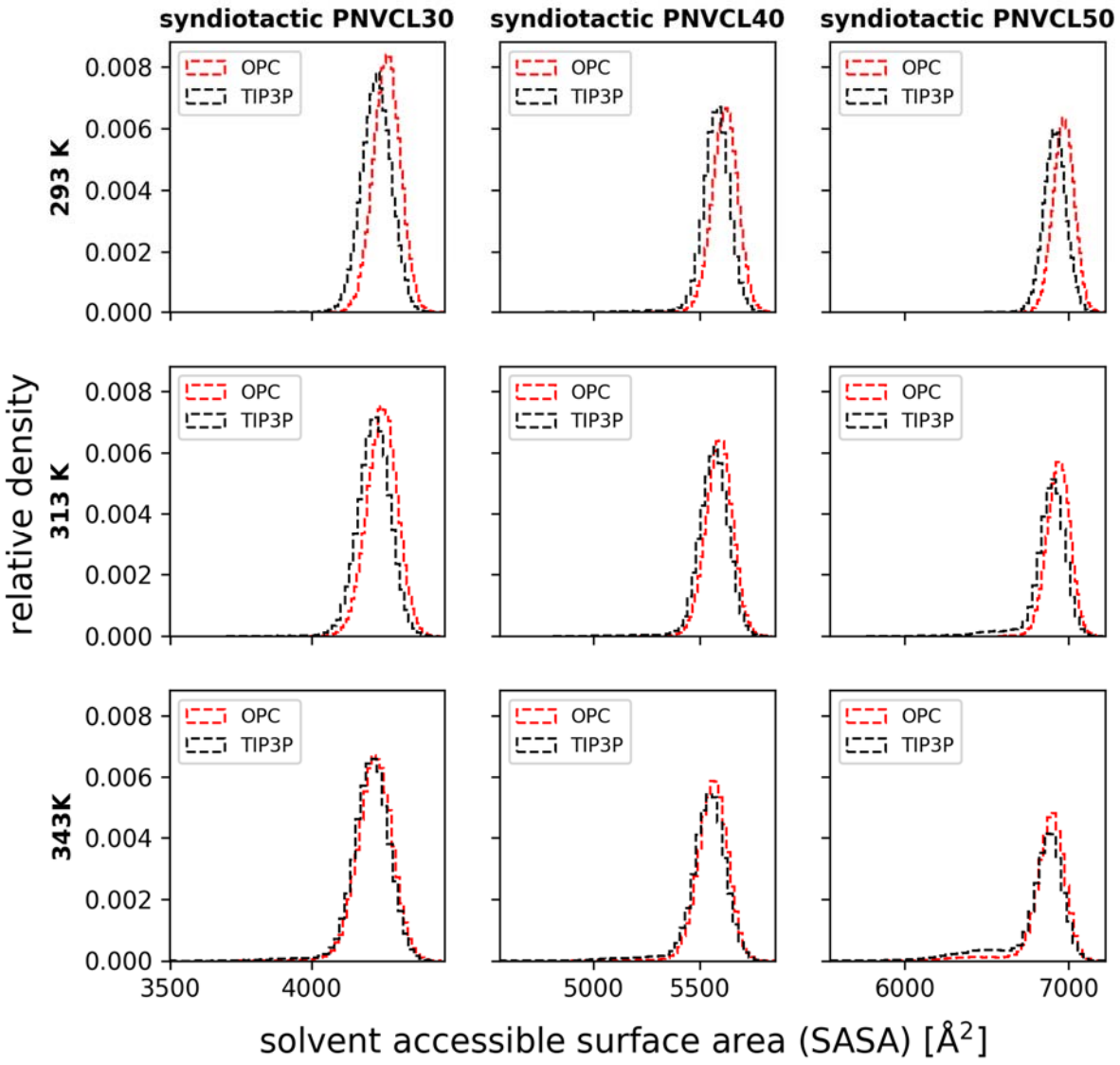

Figure S23. Influence of different water models on the solvent-accessible surface area (SASA) of syndiotactic PNVCL oligomers. The histograms depict the SASA observed in five independent 1 µs MD simulations for systems of PNVCL oligomers of varying length (30, 40, and 50 repeating units, from left to right) at different temperatures (293 K, 313K, and 343 K, from top to bottom) using the TIP3P water model (black) or the OPC water model (red).

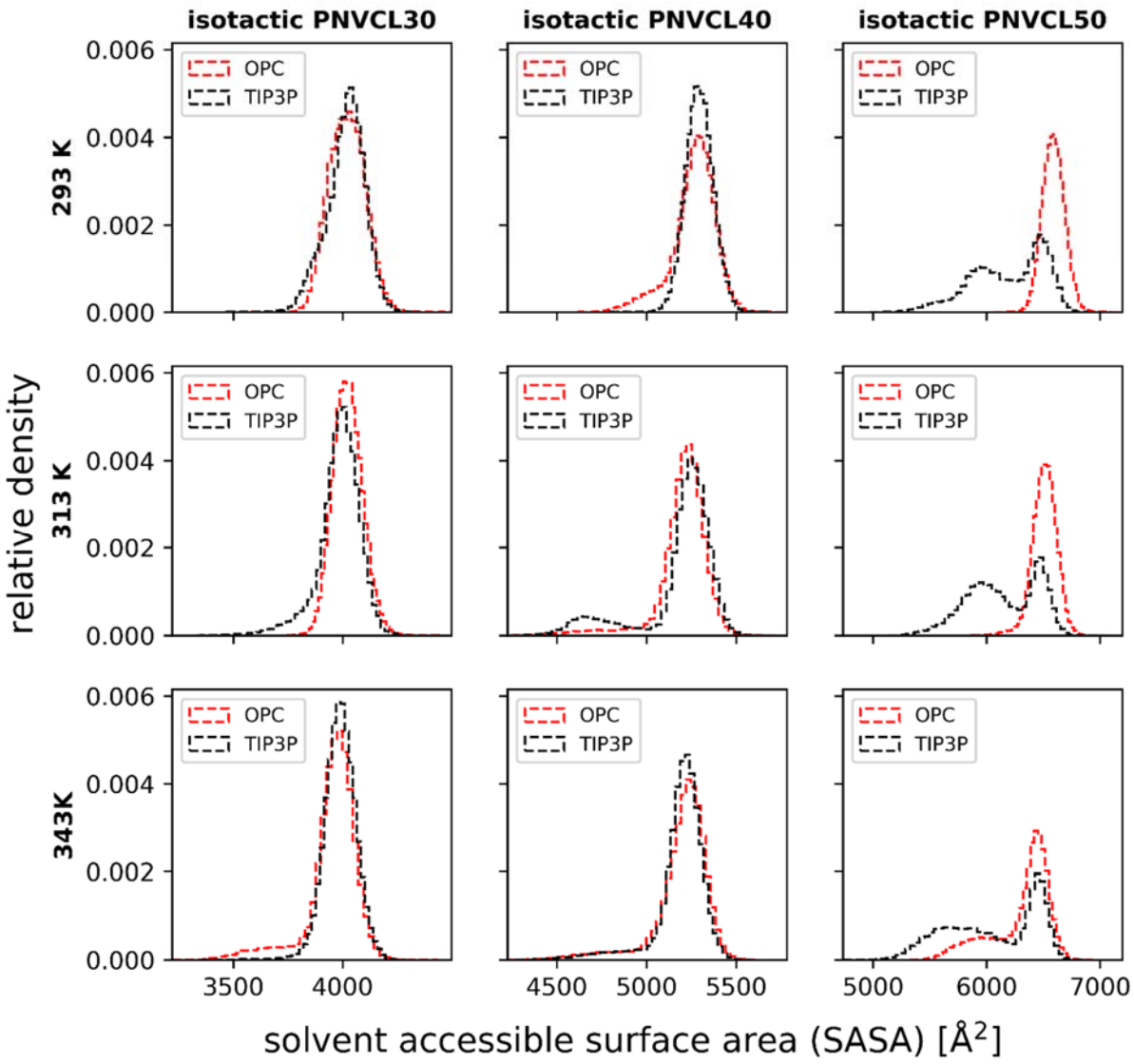

Figure S24. Influence of different water models on the solvent-accessible surface area (SASA) of isotactic PNVCL oligomers. The histograms depict the SASA observed in five independent 1 μs MD simulations for systems of PNVCL oligomers of varying length (30, 40, and 50 repeating units, from left to right) at different temperatures (293 K, 313K, and 343 K, from top to bottom) using the TIP3P water model (black) or the OPC water model (red).

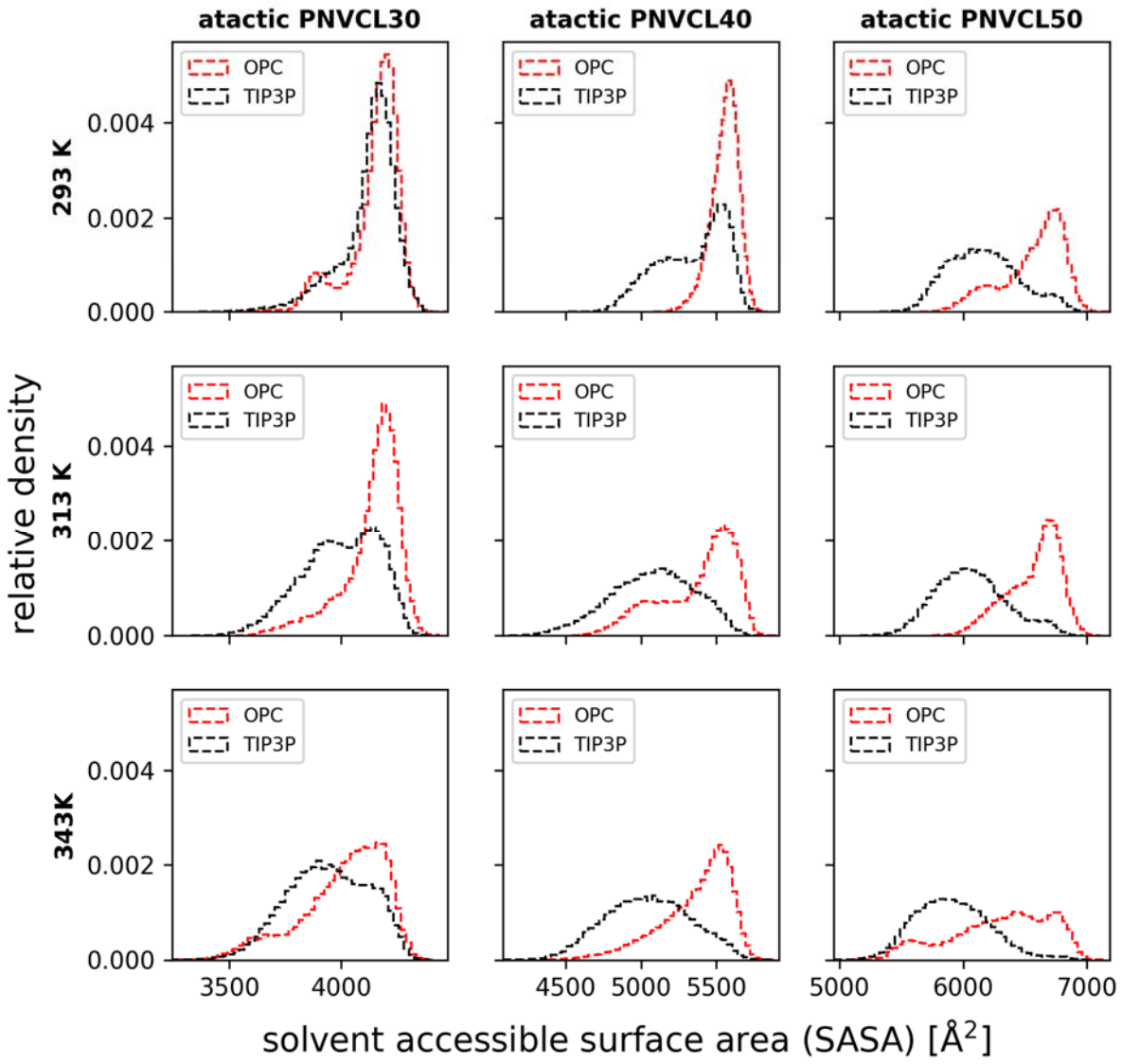

Figure S25. Influence of different water models on the solvent-accessible surface area (SASA) of atactic PNVCL oligomers. The histograms depict the SASA observed in five independent 1 µs MD simulations for systems of PNVCL oligomers of varying length (30, 40, and 50 repeating units, from left to right) at different temperatures (293 K, 313K, and 343 K, from top to bottom) using the TIP3P water model (black) or the OPC water model (red).

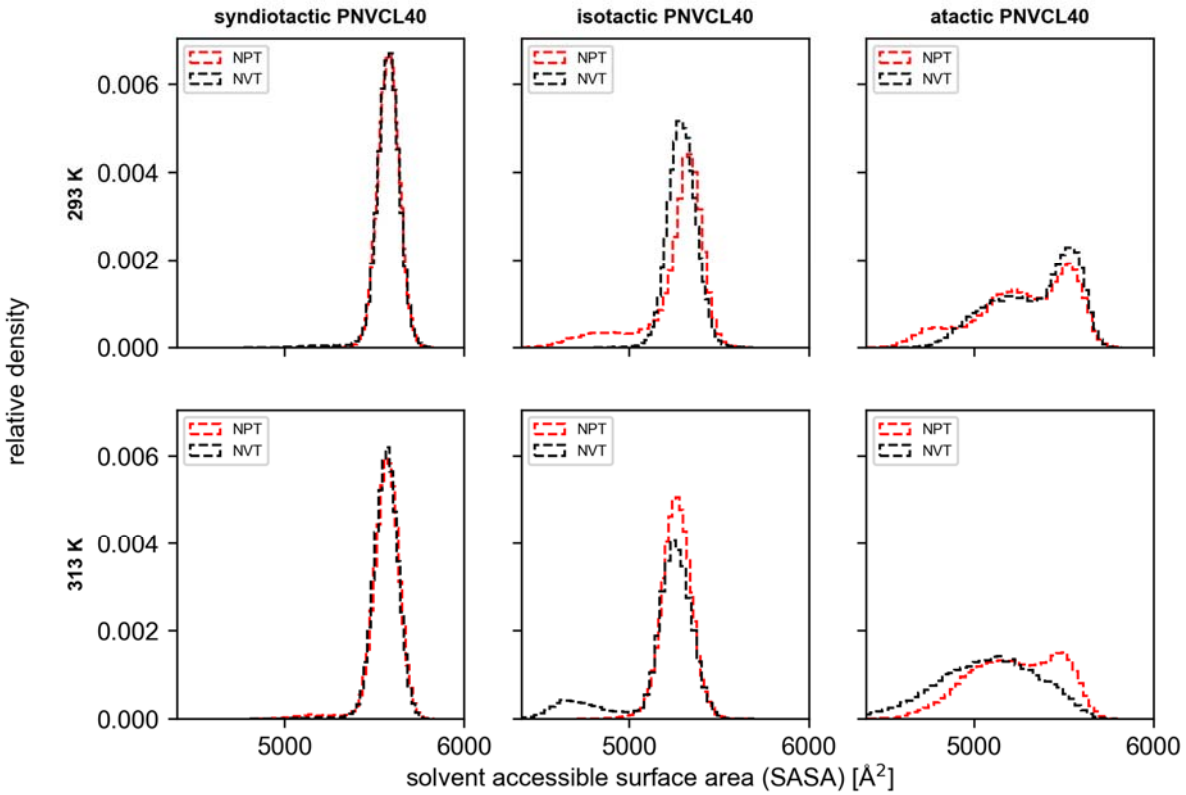

Figure S26. Influence of the MD ensemble (NVT / NPT) on the solvent-accessible surface area (SASA) of the iso-, syndio-, and atactic PNVCL 40mer. The histograms depict the SASA observed in five independent 1 µs MD simulations for each tacticity of PNVCL (from left to right) at different temperatures (293 K and 313K, from top to bottom) using the NVT (black) or NPT (red) ensemble.

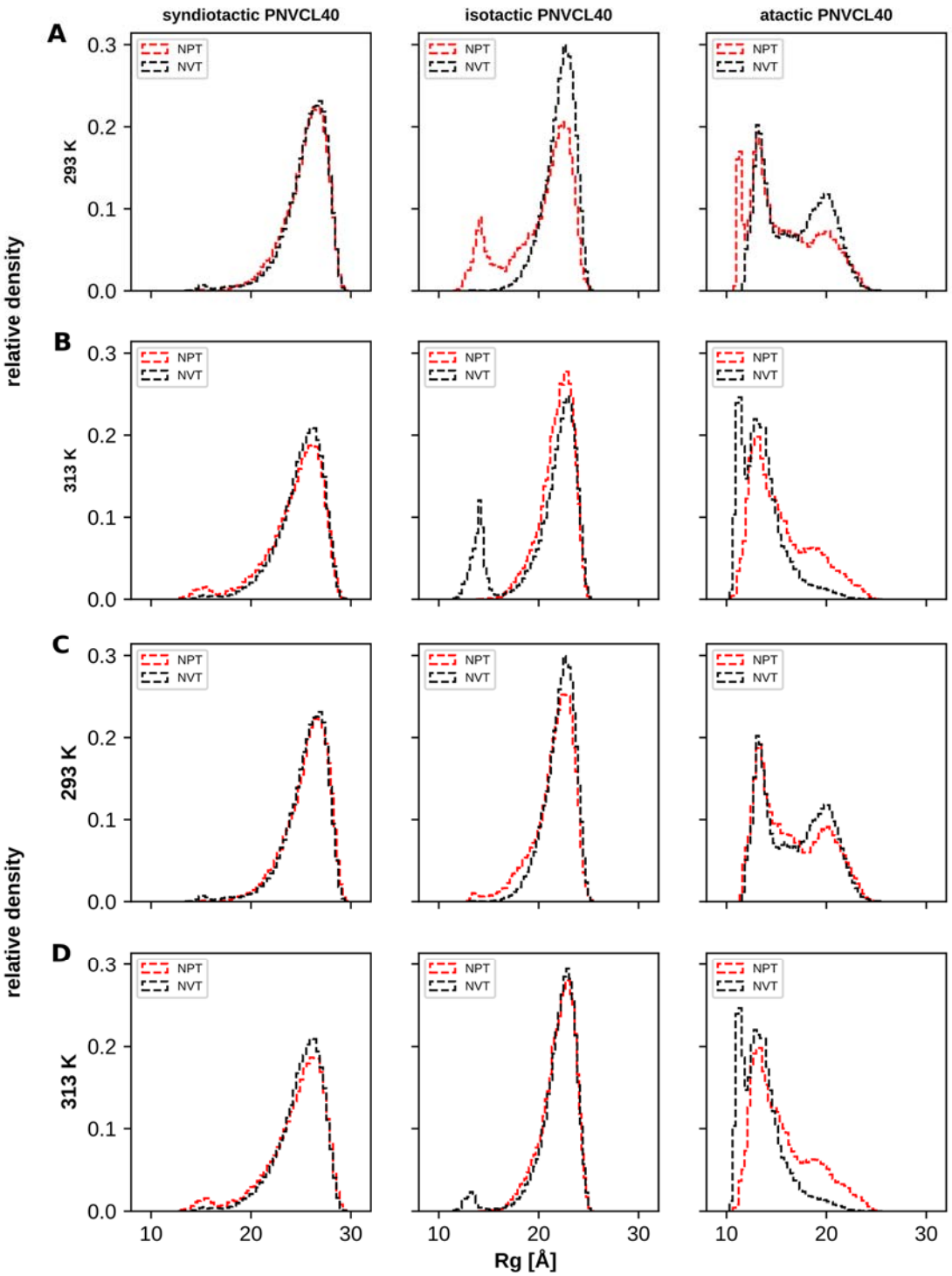

Figure S27. Influence of the MD ensemble (NVT / NPT) on the radius of gyration (R_g) of the iso-, syndio-, and atactic PNVCL 40mer. (**A, B**)The histograms depict the R_g observed in five independent 1 μs MD simulations for each tacticity of PNVCL (from left to right) at different temperatures (293 K (A) and 313K (B), from top to bottom) using the NVT (black) or NPT (red) ensemble. (**C, D**) Histograms are obtained by removing one deviating trajectory from the five trajectories of the three systems showing notable deviations in panel A and B, respectively (atactic PNVCL (293 K, NPT), isotactic PNVCL (293 K, NPT), isotactic PNVCL (313 K, NVT)).

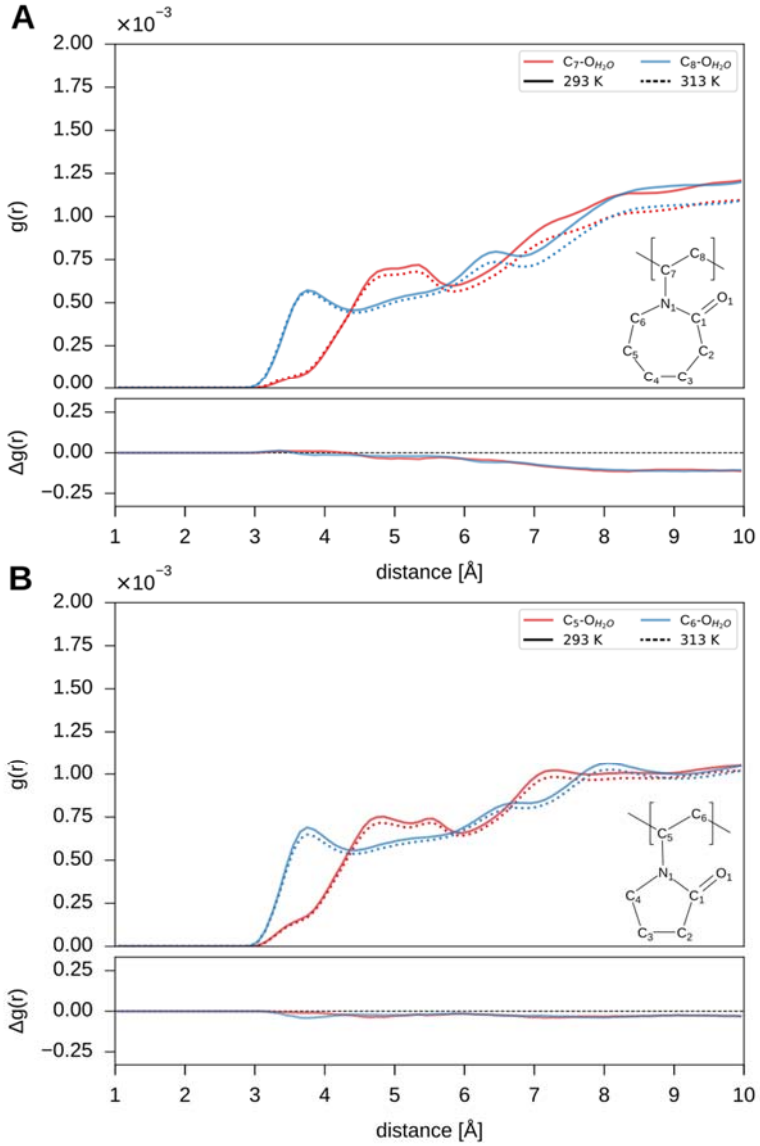

Figure S28. Radial distribution functions (RDFs) for the polymer backbone carbons atoms of PNVCL **(A)** and PVP **(B)** and the oxygen of the surrounding water molecules. Atom pairings are depicted in respective colors, and RDFs calculated from trajectories at 293 K, and 313 K are shown as solid and dotted lines, respectively. Differences in the RDFs, $\Delta g(r) = g(r)_{313} - g(r)_{293}$, are depicted in the plots below. Calculations were performed using CPPTRAJ[5].

Supplementary Tables

Table S1. Overview of performed simulations on poly(*N*-vinylcaprolactam) (PNVCL).

System							
Polymer	**Repeating units**	**Tacticity**	**MD ensemble/ water model**	**Temperature [a]**	**Length[b]**	**Replicas**	**Total [b]**
			NVT				
PNVCL	5	isotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	5	syndiotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	5	atactic	TIP3P	293, 313, 343	1	5	15
PNVCL	10	isotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	10	syndiotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	10	atactic	TIP3P	293, 313, 343	1	5	15
PNVCL	15	isotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	15	syndiotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	15	atactic	TIP3P	293, 313, 343	1	5	15
PNVCL	20	isotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	20	syndiotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	20	atactic	TIP3P	293, 313, 343	1	5	15
PNVCL	25	isotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	25	syndiotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	25	atactic	TIP3P	293, 313, 343	1	5	15
PNVCL	30	isotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	30	syndiotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	30	atactic	TIP3P	293, 313, 343	1	5	15
PNVCL	40	isotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	40	syndiotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	40	atactic	TIP3P	293, 313, 343	1	5	15
PNVCL	50	isotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	50	syndiotactic	TIP3P	293, 313, 343	1	5	15
PNVCL	50	atactic	TIP3P	293, 313, 343	1	5	15
Sum							**360**
			NPT				
PNVCL	40	isotactic	TIP3P	293, 313	1	5	10
PNVCL	40	syndiotactic	TIP3P	293, 313	1	5	10
PNVCL	40	atactic	TIP3P	293, 313	1	5	10
Sum							**30**
			NVT				
PNVCL	30	isotactic	OPC	293, 313, 343	1	5	15
PNVCL	30	syndiotactic	OPC	293, 313, 343	1	5	15
PNVCL	30	atactic	OPC	293, 313, 343	1	5	15
PNVCL	40	isotactic	OPC	293, 313, 343	1	5	15
PNVCL	40	syndiotactic	OPC	293, 313, 343	1	5	15
PNVCL	40	atactic	OPC	293, 313, 343	1	5	15
PNVCL	50	isotactic	OPC	293, 313, 343	1	5	15
PNVCL	50	syndiotactic	OPC	293, 313, 343	1	5	15
PNVCL	50	atactic	OPC	293, 313, 343	1	5	15
Sum							**135**

Table S1 continued

			TIP3P				
PNVCL	50	isotactic	0.6 wt.-%	293, 313	1	5	10
PNVCL	50	syndiotactic	0.6 wt.-%	293, 313	1	5	10
PNVCL	50	atactic	0.6 wt.-%	293, 313	1	5	10
PNVCL	50	isotactic	0.8 wt.-%	293, 313	1	5	10
PNVCL	50	syndiotactic	0.8 wt.-%	293, 313	1	5	10
PNVCL	50	atactic	0.8 wt.-%	293, 313	1	5	10
Sum							**60**
Cumulative simulation time							**585**

[a] In [K].

[b] In [μs].

Table S2. Overview of performed simulations on poly(*N*-vinylpyrrolidone) (PVP):

System							
Polymer	**Repeating units**	**Tacticity**	**MD ensemble/ water model**	**Temperature[a]**	**Length[b]**	**Replicas**	**Total[b]**
			NVT				
PVP	30	isotactic	TIP3P	293, 313	0.5	5	5
PVP	30	syndiotactic	TIP3P	293, 313	0.5	5	5
PVP	30	atactic	TIP3P	293, 313	0.5	5·3	15
PVP	40	isotactic	TIP3P	293, 313	0.5	5	5
PVP	40	syndiotactic	TIP3P	293, 313	0.5	5	5
PVP	40	atactic	TIP3P	293, 313	0.5	5·3	15
PVP	50	isotactic	TIP3P	293, 313	0.5	5	5
PVP	50	syndiotactic	TIP3P	293, 313	0.5	5	5
PVP	50	atactic	TIP3P	293, 313	0.5	5·3	15
Cumulative simulation time							**75**

[a] In K.

[b] In μs.

Table S3. Number of atoms within each investigated system. Polymer (PNVCL/PVP) and simulation conditions (NVT/NPT and TIP3P/OPC) are provided for each section, number of repeating units and tacticity are given per row.

System		
PNVCL, NVT, TIP3P		**Number of atoms**
Repeating units	**Tacticity**	
5	isotactic	5073
5	atactic	4902
5	syndiotactic	4692
10	isotactic	10165
10	atactic	9400
10	syndiotactic	9559
15	isotactic	17369
15	atactic	16478
15	syndiotactic	17153
20	isotactic	26211
20	atactic	24459
20	syndiotactic	25215
25	isotactic	35779
25	atactic	33151
25	syndiotactic	33355
30	isotactic	46514
30	atactic	45164
30	syndiotactic	44609
40	isotactic	79159
40	atactic	78532
40	syndiotactic	77368
50	isotactic	123504
50	atactic	122310
50	syndiotactic	123150
PNVCL, NVT, OPC		
30	isotactic	61662
30	atactic	60254
30	syndiotactic	59478
40	isotactic	105548
40	atactic	104752
40	syndiotactic	103232
50	isotactic	164098
50	atactic	163098
50	syndiotactic	163670
PNVCL, NPT, TIP3P		
40	isotactic	79159
40	atactic	78532
40	syndiotactic	77368
PNVCL, NVT, TIP3P, 0.8 wt.-%		
50	isotactic	169311
50	atactic	167559
50	syndiotactic	168747
PNVCL, NVT, TIP3P, 0.6 wt.-%		
50	isotactic	227754
50	atactic	225354
50	syndiotactic	227061

Table S3 continued

System		Number of atoms
PVP, NVT, TIP3P		
Repeating units	Tacticity	
30	isotactic	53798
30	atactic	53804
30	atactic	51803
30	atactic	51782
30	syndiotactic	53696
40	isotactic	97348
40	atactic	97372
40	atactic	95350
40	atactic	97447
40	syndiotactic	97279
50	isotactic	154881
50	atactic	153132
50	atactic	152988
50	atactic	153057
50	syndiotactic	155220

Table S4. Residues containing cabon atoms with inverted chirality for atactic structures of PNVCL and PVP.

	Position in the chain																																																	
					5					**10**					**15**					**20**					**25**					**30**					**35**					**40**					**45**					**50**
PNVCL, NVT, TIP3P																																																		
5mer					x																																													
10mer	x	x		x	x		x	x	x	x																																								
15mer	x		x	x			x	x	x		x	x	x		x																																			
20mer		x		x				x					x	x	x	x		x	x																															
25mer	x	x		x			x		x			x		x		x	x	x	x			x	x		x																									
30mer		x					x		x	x	x			x		x	x	x	x		x	x		x	x	x	x	x	x																					
40mer								x		x				x	x	x	x				x	x	x		x	x	x		x	x			x	x				x	x											
50mer										x				x	x		x	x			x		x					x	x	x			x			x	x	x	x	x	x							x		
PNVCL, NPT, TIP3P																																																		
40mer								x		x				x	x	x	x				x	x	x		x	x	x		x	x			x	x				x	x											
PNVCL, NVT, OPC																																																		
30mer		x					x		x	x	x			x		x	x	x	x		x	x		x	x	x	x	x	x																					
40mer								x		x				x	x	x	x				x	x	x		x	x	x		x	x			x	x				x	x											
50mer										x				x	x		x	x			x		x					x	x	x			x			x	x	x	x	x	x							x		
PNVCL, NVT, TIP3P, different concentrations																																																		
50mer										x				x	x		x	x			x		x					x	x	x			x			x	x	x	x	x	x							x		
PVP, NVT, TIP3P																																																		
30mer				x	x		x	x		x	x	x	x	x				x	x	x	x			x			x	x	x																					
	x		x		x	x	x	x			x		x	x			x					x			x		x		x	x																				
		x	x	x	x	x			x		x	x	x	x			x	x		x	x		x	x					x	x																				
40mer		x	x			x			x		x	x	x		x	x		x	x	x					x	x		x	x	x	x	x		x			x													
		x	x			x	x	x		x					x				x			x	x			x		x	x					x		x	x	x		x										
	x	x			x	x		x			x				x			x	x		x	x	x		x		x			x	x	x	x		x	x	x	x	x											
50mer	x		x		x	x	x	x	x	x	x	x				x	x		x	x				x	x	x					x	x		x	x		x					x		x		x	x	x	x	x
	x	x						x		x		x	x				x		x	x			x			x			x			x	x	x					x	x	x			x		x	x			x
	x		x		x		x	x	x	x	x	x					x				x		x		x	x		x				x	x	x		x			x		x	x		x	x	x	x	x	x	x

Table S5. Partial charges and coordinates of the atoms of the PNVCL repeating unit. The repeating unit possesses two open valences/connect records (C9 and C10), its net charge is zero.

Atom number	Atom name[a]	X	Y	Z	Charge[b]
1	C10	3.540003	1.419779	0.000002	0.1586
2	N1	2.072742	1.385823	-0.113022	-0.1702
3	C1	1.361680	2.430967	0.394038	0.4927
4	C2	0.147772	2.431003	0.206963	-0.1076
5	C3	0.882646	1.277252	0.907811	0.0136
6	C4	0.873659	-0.037083	0.126603	-0.1299
7	C5	0.516471	-0.608169	-0.149407	0.0697
8	C6	1.454795	0.332093	-0.912610	-0.1441
9	H9	2.259787	-0.257933	-1.322608	0.0656
10	H10	0.934927	0.757022	-1.767444	0.0656
11	H7	0.993614	-0.902292	0.781933	0.0050
12	H8	0.402425	-1.514410	-0.740563	0.0050
13	H5	1.458537	-0.775905	0.666469	0.0292
14	H6	1.383829	0.119987	-0.823742	0.0292
15	H4	1.913757	1.579298	1.058816	0.0210
16	H3	0.460752	1.129275	1.898607	0.0210
17	H1	0.473614	3.375599	0.616797	0.0418
18	H2	0.392762	2.437013	-0.853078	0.0418
19	O1	1.880690	3.339947	0.991038	-0.6639
20	H19	3.751995	2.301651	0.578912	0.0885
21	C9	4.098076	0.210499	0.763167	0.0483
22	H17	5.182318	0.252471	0.698154	0.0096
23	H18	3.806962	-0.716729	0.274041	0.0096

[a] See Figure S6 for the structure.

[b] In e.

Table S6. Partial charges and coordinates of the atoms of the PVP repeating unit. The repeating unit possesses two open valences/connect records (C4 and C5); its net charge is zero.

Atom number	Atom name	X	Y	Z	Charge[a]
1	C4	3.540003	1.419779	0.000002	0.1324
2	N1	2.147718	1.340814	-0.422081	-0.1473
3	C2	1.262838	0.488983	0.154826	0.5274
4	C1	-0.049847	0.596013	-0.599976	-0.2009
5	C8	0.093183	1.888002	-1.406950	-0.0742
6	C3	1.612482	2.031987	-1.581003	-0.0750
7	H3	1.951550	1.570864	-2.506283	0.0560
8	H4	1.926608	3.068291	-1.597996	0.0560
9	H14	-0.428687	1.867039	-2.355686	0.0455
10	H15	-0.287779	2.726984	-0.83595	0.0455
11	H1	-0.884155	0.583035	0.08728	0.0767
12	H2	-0.135916	-0.278231	-1.24017	0.0767
13	O1	1.481789	-0.232962	1.08673	-0.6768
14	H5	3.613974	0.735868	0.83388	0.0888
15	C5	4.507986	0.953756	-1.09498	0.0573
16	H6	5.518135	1.052730	-0.70591	0.0059
17	H7	4.449030	1.624475	-1.94961	0.0059

[a] In e.

References

1. Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Pérez-Hernández, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J.-H.; Noé, F., PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* **2015,** *11* (11), 5525-5542.

2. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford CT, 2016.

3. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006,** *25* (2), 247-260.

4. Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E., MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012,** *8* (9), 3314-3321.

5. Roe, D. R.; Cheatham, T. E., PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013,** *9* (7), 3084-3095.