

Supplementary Information

Molecular origin of urea driven hydrophobic polymer collapse and unfolding depending on side chain chemistry

Divya Nayar, Angelina Folberth and Nico F. A. van der Vegt*

*Eduard-Zintl-Institut für Anorganische und Physikalische Chemie, Center of Smart Interfaces, Technische Universität Darmstadt, Alarich-Weiss-Strasse 10, 64287, Darmstadt, Germany;
E-mail: vandervegt@cpc.tu-darmstadt.de*

Table of Contents

Page S3	Thermodynamics of polymer collapse	Description of thermodynamic quantities computed.
Page S4	Observables computed	Description of observables computed.
Page S5	Error estimation	Description of error analysis.
Page S7	PDEA force-field validation	Checks for the PDEA force-field.
Page S8	Table S1	System setup details.
Page S9	Figure S1	Preferential binding coefficients of urea to PNiPAM.
Page S10	Figure S2	Preferential binding coefficients of urea to PDEA.
Page S11	Figure S3	Difference in preferential binding coefficients upon unfolding of the polymers as a function of urea concentration.
Page S12	Figure S4	Local-bulk partition coefficients of urea and water for the polymers.
Page S13	Figure S5	Change in internal polymer energies and electrostatic polymer solvation energy upon polymer unfolding in urea solutions.
Page S14	Figure S6	Probability distributions of radius of gyration of polymers sampled during MD simulations.
Page S15	Figure S7	Representative snapshots of collapsed and extended structures of polymers.
Page S16	Figure S8	Time evolution of radius of gyration of PDEA in water at different temperatures.
Page S17	Figure S9	Probability distributions of radius of gyration of PDEA in 2.7 M and 5.8 M aqueous urea.
Page S18	Figure S10	Correlation between change in polymer solvation energy and polymer-urea bivalent H-bonds upon PNiPAM unfolding.

Thermodynamics of polymer collapse

The polymers studied in this work undergo first-order collapse.¹ It has furthermore been experimentally shown that the coil-to-globule transition in PNiPAM is a two-state process for short chains.¹ In our work, we call these two (degenerate) states C (collapsed) and E (extended). Assuming a two-state conformational equilibrium between the collapsed (C) and extended (E) polymer conformations, $C \rightleftharpoons E$, the Gibbs free energy change associated with the unfolding of the polymer in aqueous urea solution relative to that in pure aqueous solution is defined by,

$$\Delta\Delta G_{C \rightarrow E} = \Delta\Delta U_{C \rightarrow E} - T\Delta\Delta S_{C \rightarrow E} \quad (1)$$

where,

$$\Delta\Delta G_{C \rightarrow E} = \Delta G_{C \rightarrow E}(\text{urea}) - \Delta G_{C \rightarrow E}(\text{water}) \quad (2)$$

$$\Delta\Delta U_{C \rightarrow E} = \Delta U_{C \rightarrow E}(\text{urea}) - \Delta U_{C \rightarrow E}(\text{water}) \quad (3)$$

$$\Delta\Delta S_{C \rightarrow E} = \Delta S_{C \rightarrow E}(\text{urea}) - \Delta S_{C \rightarrow E}(\text{water}) \quad (4)$$

are the Gibbs free energy, energy, and entropy of transferring the equilibrium $C \rightleftharpoons E$ from pure water to an aqueous urea solution. Eq. 1 should in principle further include a pressure-volume work term. This term contributes negligibly to $\Delta G_{C \rightarrow E}$ (and $\Delta\Delta G_{C \rightarrow E}$) at 1 atm. pressure and will therefore be ignored. We refer to $\Delta U_{C \rightarrow E}$ (and $\Delta\Delta U_{C \rightarrow E}$) as energies, which, due to the vanishing pressure-volume term, may also be referred to as enthalpies.

The above enthalpy and entropy changes may be obtained from the temperature-dependence of the Gibbs free energy and contain contributions from changes in polymer-solvent interactions, polymer-polymer (internal or intramolecular) interactions and solvent-solvent interactions (solvent-reorganization). Based on the Widom potential distribution theorem it can be shown that $\Delta G_{C \rightarrow E}$ (and $\Delta\Delta G_{C \rightarrow E}$) can also be expressed in terms of changes in polymer-solvent and polymer-polymer interactions only.^{2,3} Solvent reorganization contributions always cancel exactly in the Gibbs free energy changes (exact enthalpy-entropy compensation)^{2,3} and are therefore not considered in $\Delta U_{C \rightarrow E}$, $\Delta\Delta U_{C \rightarrow E}$, $\Delta S_{C \rightarrow E}$, and $\Delta\Delta S_{C \rightarrow E}$ analysed and discussed in this work. Solvent-solvent interactions of course affect the solvent structure around the polymer and as such contribute indirectly to all quantities evaluated.

Observables computed

Preferential binding coefficients

Cosolvent effects on the conformational equilibrium $C \rightleftharpoons E$ are analysed using the Wyman-Tanford relation⁴⁻⁶

$$\left(\frac{\partial \ln K}{\partial \ln a_u} \right)_{p,T} = \Delta \Gamma_{C \rightarrow E} \quad (5)$$

where K is the equilibrium constant, $\Delta \Gamma_{C \rightarrow E} = \Gamma_{Pu}^E - \Gamma_{Pu}^C$ the change in the (polymer-urea) preferential binding coefficient of the reaction, a_u is the cosolvent (urea) activity, p is the pressure and T the temperature. Urea acts as a denaturant if $\Delta \Gamma_{C \rightarrow E} > 0$, while acting as a protecting osmolyte if $\Delta \Gamma_{C \rightarrow E} < 0$. The relation between urea activity and urea concentration is provided by Kirkwood-Buff theory and requires to consider urea-urea (G_{uu}) and urea-water (G_{uw}) Kirkwood-Buff integrals (KBIs):

$$\left(\frac{\partial \ln a_u}{\partial \ln c_u} \right)_{p,T} = \frac{1}{1 + c_u(G_{uu} - G_{uw})} \quad (6)$$

Therefore, Eq. 5 can alternatively be written as,

$$\left(\frac{\partial \ln K}{\partial \ln c_u} \right)_{p,T} = \frac{\Delta \Gamma_{C \rightarrow E}}{1 + c_u(G_{uu} - G_{uw})} \quad (7)$$

The free energy change, $\Delta G_{C \rightarrow E} = -RT \ln K$ (with R the gas constant), of unfolding the collapsed state C of the polymer is obtained by integrating Eq. (7).

The preferential binding coefficients for the collapsed and extended states were calculated using⁶

$$\Gamma_{Pu} = \left\langle n_u(r) - \frac{N_u - n_u(r)}{N_w - n_w(r)} n_w(r) \right\rangle \quad (8)$$

where, $n_x(r)$ is the number of urea or water molecules at proximal distance r from the polymer surface calculated as the minimum distance from the center of mass of the solvent molecule to the polymer surface. N_x is the total number of water or urea molecules in the system. The computation of average Γ_{Pu} and the associated errors are described below in the subsection on error estimation. The quantity $\Delta \Gamma_{C \rightarrow E} = (\Gamma_{Pu}^E - \Gamma_{Pu}^C)$ was computed as a function of concentration for both the polymers as shown in Figure S3. Since, the Γ_{Pu}^E of PDEA at 5.8 M shows poor convergence at large distance from the polymer, we do not

consider the $\Delta\Gamma_{C\rightarrow E}$ at 5.8 M urea concentration for computation of $\Delta\Delta G_{C\rightarrow E}$ of PDEA.

Figure S1 and S2 show the preferential binding parameters Γ_{Pu}^E and Γ_{Pu}^C for PNiPAM and PDEA at different urea concentrations. The data in Figure S1 show *positive* preferential binding of urea to the two polymers in, both, C and E states at all urea concentrations considered. The convergence of the data, achieved by virtue of conformational averaging, is much better than that previously achieved and thus allows to quantitatively estimate the free energy change of unfolding. The error bars on Γ_{Pu} are higher in case of PDEA than those for PNiPAM. This indicates that the PDEA chain conformations are more sensitive to the solvent binding, which may be due to the larger size of C and E states of PDEA than those for PNiPAM. Figure S2 shows that $\Delta\Gamma_{C\rightarrow E} > 0$ for PDEA and $\Delta\Gamma_{C\rightarrow E} < 0$ for PNiPAM. This observation implies that urea unfolds PDEA while it collapses PNiPAM, in agreement with the experimental data.⁷

Local-bulk Partition coefficients

The concentration of the urea molecules relative to that of water in the solvation shells of the polymers can also be computed using the partition coefficient, K_p .

$$K_p = \frac{\langle n_u(r) \rangle / \langle n_w(r) \rangle}{(N_u/N_w)} \quad (9)$$

where the variables in the equation are same as described above. The K_p can be related to the Γ_{Pu} as,

$$\Gamma_{Pu} = \langle n_u(r) \rangle \left(1 - \frac{1}{K_p} \right) \quad (10)$$

with,

$$\Delta\Gamma_{C\rightarrow E} = \Delta \langle n_u(r) \rangle - \Delta \left(\frac{\langle n_u(r) \rangle}{K_p} \right) \quad (11)$$

Figure S4 shows urea clouding in the first solvation shell of the polymers in both collapsed and extended states, supporting the implications from the preferential binding coefficients. Higher urea accumulation is observed, however, on the collapsed polymer surface for both the polymers at all urea concentrations. In case of PDEA, despite higher urea accumulation on the collapsed surface, the $\Delta\Gamma_{C\rightarrow E}$ can be positive since $\Delta \langle n_u(r) \rangle$ is dominant.

Solvent accessible surface area

The hydrophobic solvent accessible surface area of polymers is computed using a solvent probe radius of 1.4 nm. It is calculated using the double cubic lattice method based on the approach suggested by Shrake and Rupley,^{8,9} as implemented in Gromacs 4.6.7 analysis tool.

Error estimation

The errors reported in different observables are structure-averaged errors that have been computed in a way to reflect the fluctuations in the observable over different conformations of the polymer in the C or E state. This is unlike the usual computation of the errors by time-averaging over simulated MD trajectories of a single coil or globule, where the effect of structural variation of the polymer (within the highly degenerate C and E states) on the computed observable is not considered (reversible coil-globule transitions of a PNiPAM 40-mer at 300 K occur on 100 ns time scales). For the analysis used in this work, two-dimensional histograms of the observable and the radius of gyration (Rg) of the polymers were created for both C and E structural ensembles. Averages of the observable (polymer-solvent energetics, H-bonds, SASA) were then computed from these histograms together with the corresponding standard error ($\sigma/\sqrt{(N-1)}$, σ the standard deviation).

The preferential binding coefficients (Γ_{Pu}) are very sensitive to the change in polymer conformations. For computing the error in Γ_{Pu} , averaging was performed over the polymer conformations histogrammed over different Rg bins to obtain time-averaged $\Gamma_{Pu}(r, Rg)$ profiles, with r being the proximal distance from the polymer surface. From 200 MD simulations 200 histograms $\Gamma_{Pu}(r, Rg)$ are obtained. As the next step, we reduced this information to obtain 5 pre-averaged $\Gamma_{Pu}(r, Rg)$ histograms, with each of those 5 histograms corresponding to an average over 40 $\Gamma_{Pu}(r, Rg)$ histograms taking out of the 200 samples. The associated errors were computed accordingly. The remaining 5 histograms were finally averaged to obtain $\Gamma_{Pu}(r, Rg)$, and $\Gamma_{Pu}^E(r)$ and $\Gamma_{Pu}^C(r)$ (by averaging within C and E states using a cut off criterion for Rg; see Figure S6). The errors on the final $\Gamma_{Pu}^E(r)$ and $\Gamma_{Pu}^C(r)$ profiles were obtained using the Gaussian error propagation method.¹⁰ The thermodynamic limiting value of Γ_{Pu} , which can be obtained at large distance from the polymer, was computed by averaging over a distance of 1.5-2.0 nm from the polymer. In case of 0.6 M and 1.3 M urea concentrations for PNiPAM, averaging has been done over distance of 1.2-1.5 nm.

PDEA force-field validation

To validate the OPLS-AA force-field model of PDEA, the conformational behaviour of PDEA was examined in water and aqueous urea solutions. As the first step, MD simulations initiating with different extended polymer conformations were carried out in SPC/E water at temperatures below (280 K) and above (320 K) the experimental LCST (308 K) and the radius of gyration (R_g) of PDEA was monitored as shown in Figure S8(a). The correct conformational behaviour was predicted over a run length of 50 ns. As expected, the extended conformation collapsed after 10 ns at the higher temperature, whereas, no collapse was observed at 280 K. The R_g of PDEA was monitored for another set of simulations at 300 K (close to LCST), where the simulations were initiated with collapsed and extended structures, as shown in Figure S8(b). The extended state did not collapse over 50 ns. The simulation initiated with the collapsed state, however, sampled both collapsed and uncollapsed structures, as expected at 300 K. The conformational preferences of PDEA model were also checked in aqueous urea solutions at low (2.7 M) and high (5.8 M) concentrations. 20 MD simulations starting with different extended conformations, each for 50 ns, were carried out each at 300 K and at the respective LCST of urea solutions. The probability distributions of R_g obtained (Figure S9 (a)) show that at low urea concentrations, along with the extended states, the simulations also sample the collapsed state but with lower probability. The probability of sampling the collapsed structures increases at the LCST. Similar behaviour is observed for PDEA in 5.8 M urea solution at the LCST (Figure S9(b)). At 300 K, only the extended state is sampled, as is expected for PDEA in high urea concentrations. Therefore, the OPLS-AA model of PDEA is able to predict the expected conformational behaviour in water and aqueous urea solutions.

System setup details

Table S1: System setup of PNiPAM and PDEA in urea-water mixtures. N_w and N_u represent the number of water and urea molecules in the system. ρ is the average density of the system in g cm^{-3} and l denotes the box length of the system in nm, as obtained from 50 ns NpT simulations.

$[c]_u(\text{M})$	PNiPAM				PDEA			
	N_w	N_u	$\langle\rho\rangle$	$\langle l\rangle$	N_w	N_u	$\langle\rho\rangle$	$\langle l\rangle$
0.0	8548	0	0.998	6.41	23971	0	0.996	9.00
0.3	-	-	-	-	23490	136	1.001	8.98
0.6	8131	100	1.007	6.37	23013	272	1.006	8.96
1.3	8033	200	1.018	6.41	21912	590	1.017	8.92
1.9	7700	300	1.028	6.39	20992	862	1.028	8.89
2.7	6836	400	1.040	6.23	19734	1224	1.041	8.85
3.3	6643	500	1.050	6.25	18858	1496	1.052	8.82
3.8	7135	600	1.056	6.43	18112	1723	1.061	8.80
4.5	6300	700	1.069	6.29	17120	2040	1.074	8.77
5.8	6156	800	1.079	6.31	15305	2630	1.105	8.71

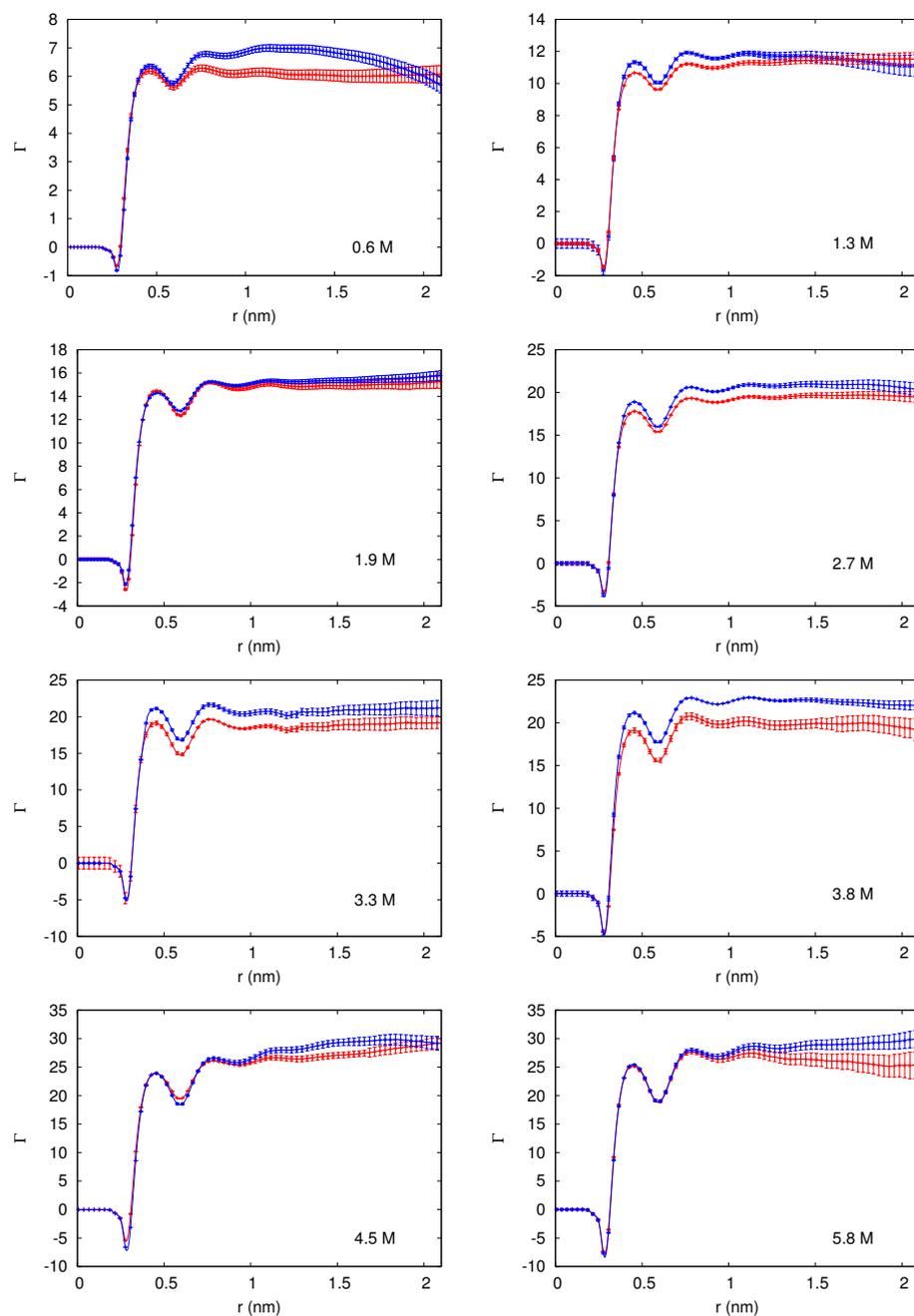


Figure S1: Preferential binding coefficients Γ_{Pu} of urea for PNiPAM as a function of the proximal distance with respect to the the polymer surface. The blue and red lines represent the collapsed and extended states of the polymers, respectively, at different urea concentrations.

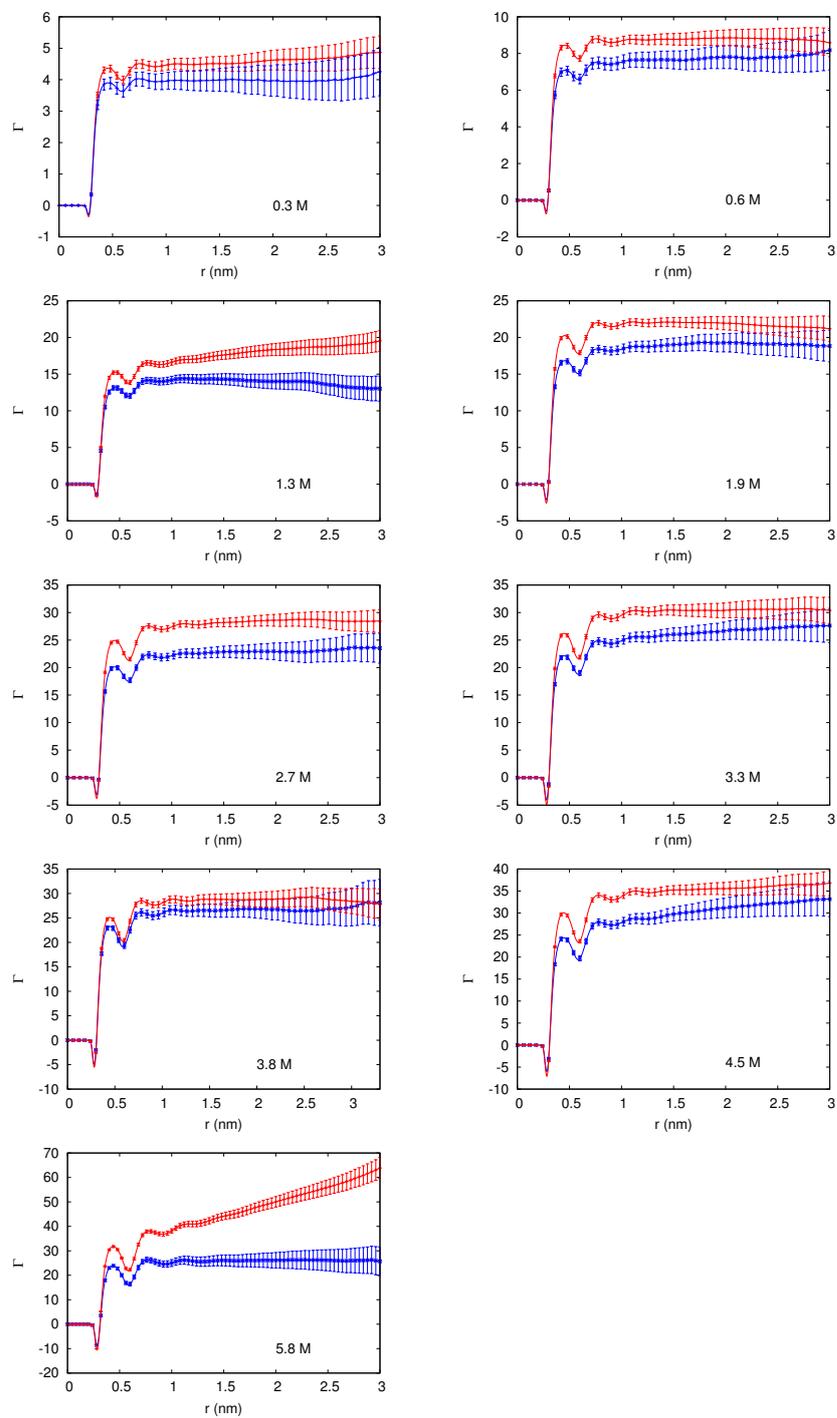


Figure S2: Preferential binding coefficients Γ_{Pu} of urea for PDEA as a function of the proximal distance with respect to the the polymer surface. The blue and red lines represent the collapsed and extended states of the polymers, respectively, at different urea concentrations.

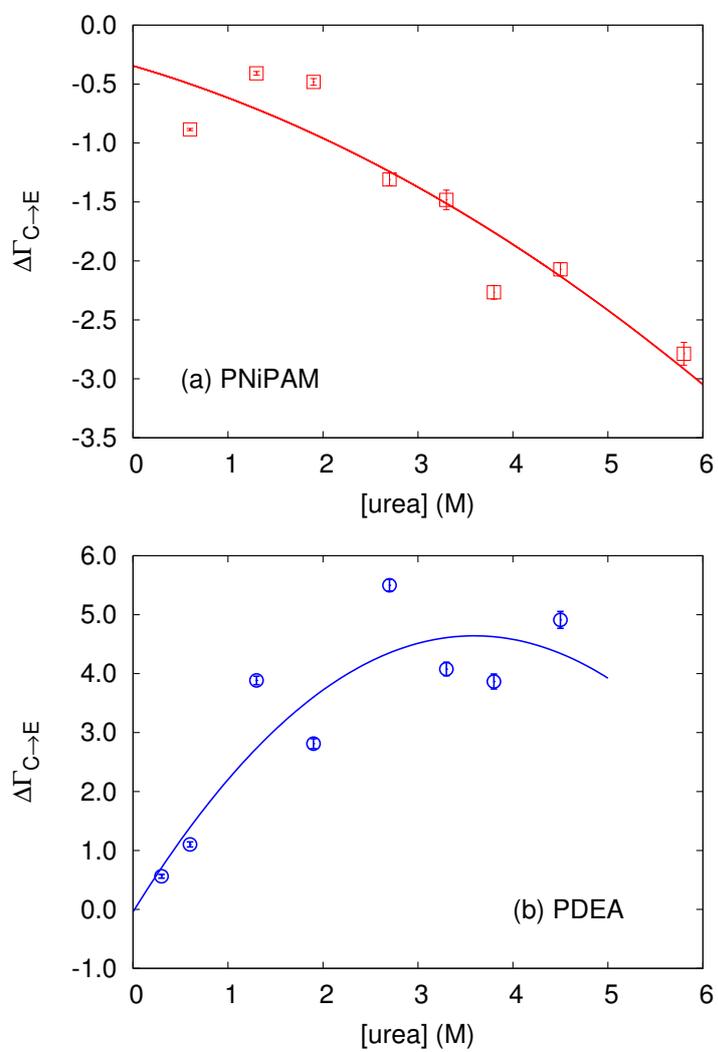


Figure S3: The difference in preferential binding coefficients on unfolding of (a) PNiPAM and (b) PDEA at different urea concentrations. The points are fitted to a quadratic polynomial of the form $ax^2 + bx + c$ where, $a = -0.035$, $b = -0.2354$, $c = -0.03451$ for PNiPAM and $a = -0.3629$, $b = 2.6076$, $c = -0.0428$ for PDEA.

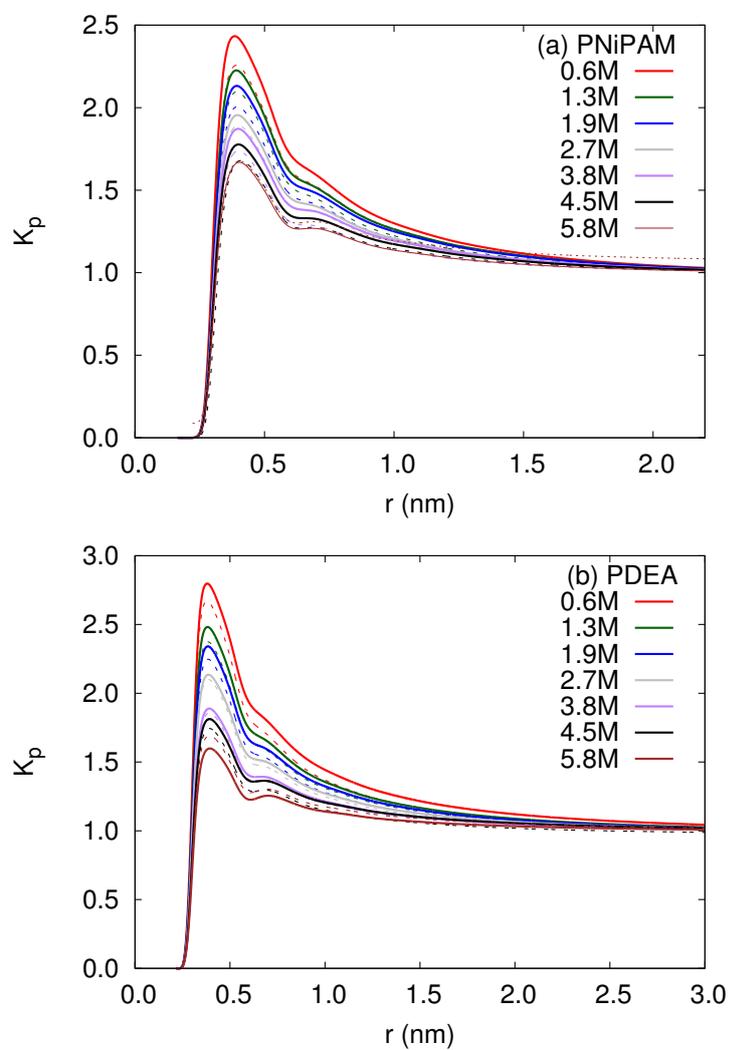


Figure S4: Local-bulk partition coefficients of urea for (a) PNiPAM and (b) PDEA as a function of the proximal distance with respect to the the polymer surface. The solid and dotted lines represent the collapsed and extended states of the polymers, respectively, in different urea concentrations.

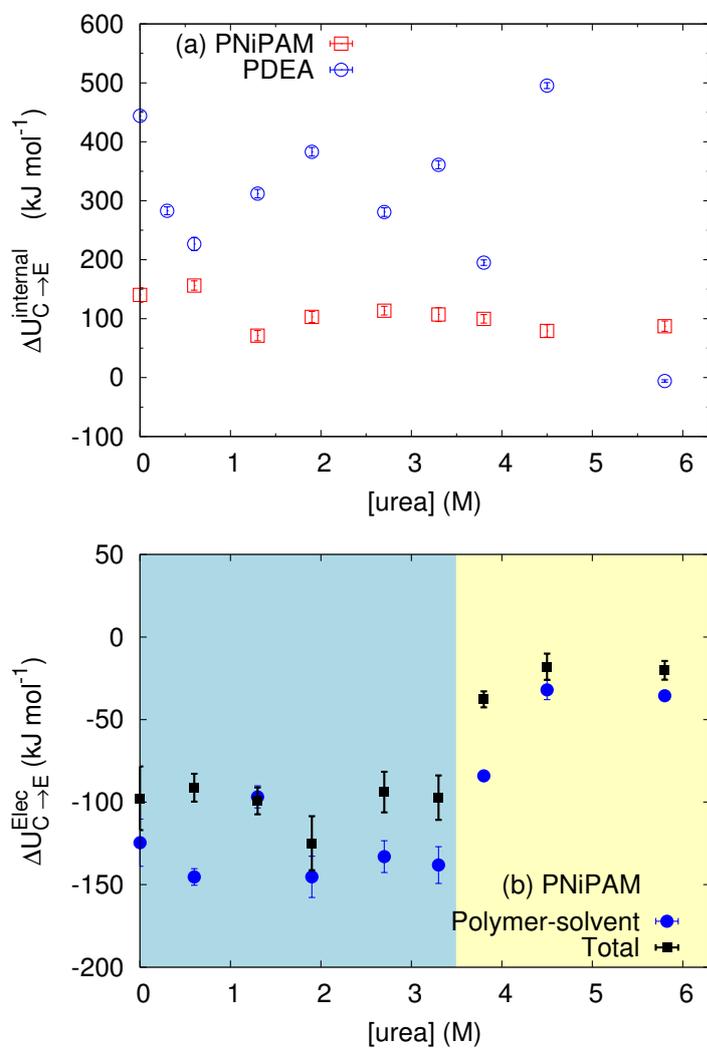


Figure S5: (a) Change in internal energies of PNiPAM and PDEA on unfolding in solution with different urea concentrations and (b) change in electrostatic polymer solvation energy upon unfolding of PNiPAM in solution with different urea concentrations. The internal polymer energies comprise of angular and dihedral contributions as well as non-bonded van der Waals and electrostatic contributions.

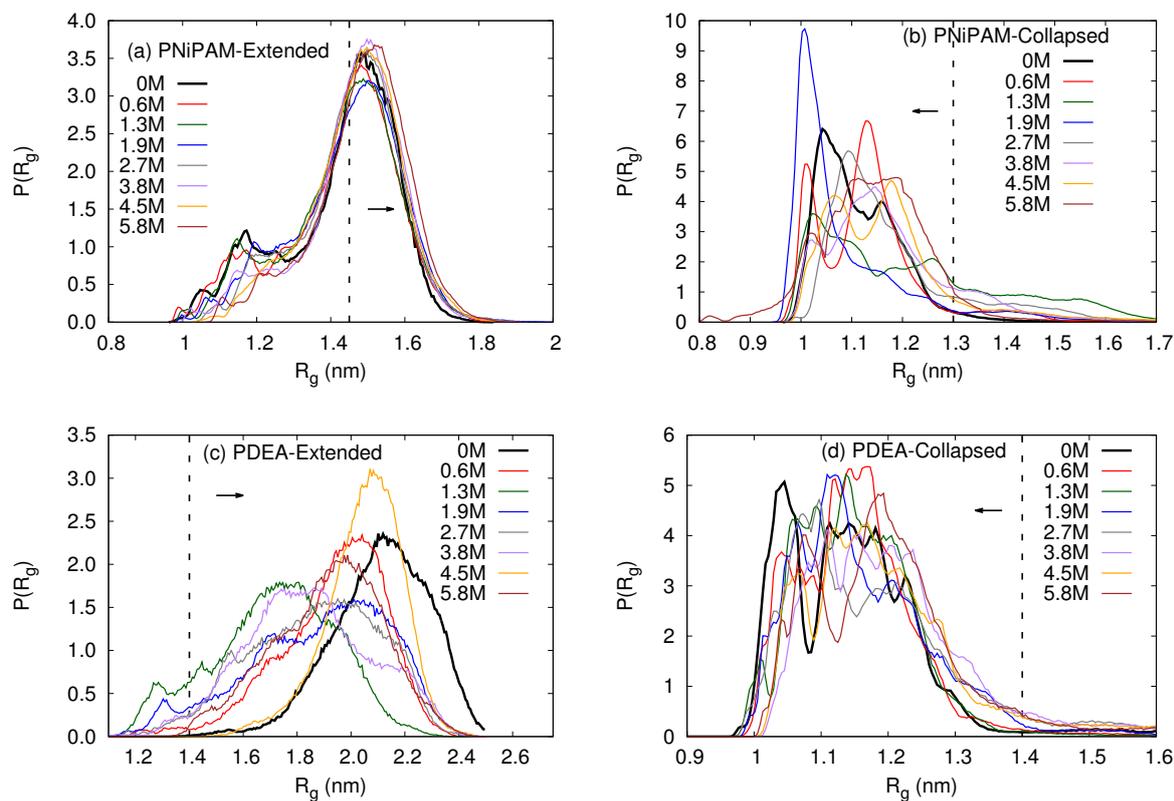


Figure S6: Normalized probability distributions of the radius of gyration (R_g) obtained from large-scale conformational sampling of extended (a,c) and collapsed (b,d) PNIPAM and PDEA chains at different urea concentrations. Each distribution is based on 100 independent 20-50 ns MD simulations. The vertical dashes lines correspond to the R_g -values that were used as selection criterion for distinguishing the initially collapsed and extended chains. Only conformations within the R_g -windows corresponding to the direction of the arrows have been used to compute average energetic and structural properties of collapsed and extended states.

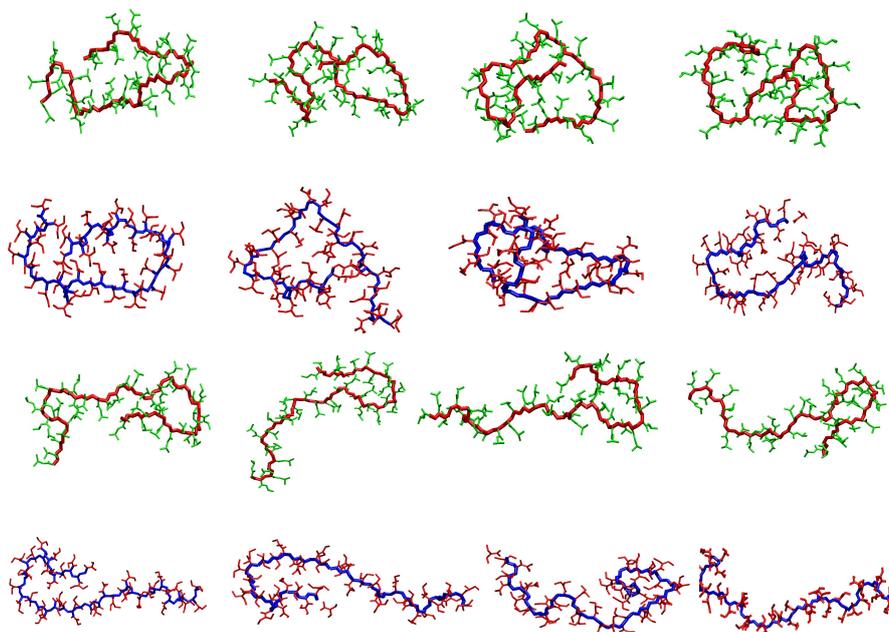


Figure S7: Representative snapshots of the collapsed (first and second rows) and extended (third and fourth rows) structures of PNiPAM (red backbone and green side chains) and PDEA (blue backbone and red side chains) as collected from simulations in different urea concentrations.

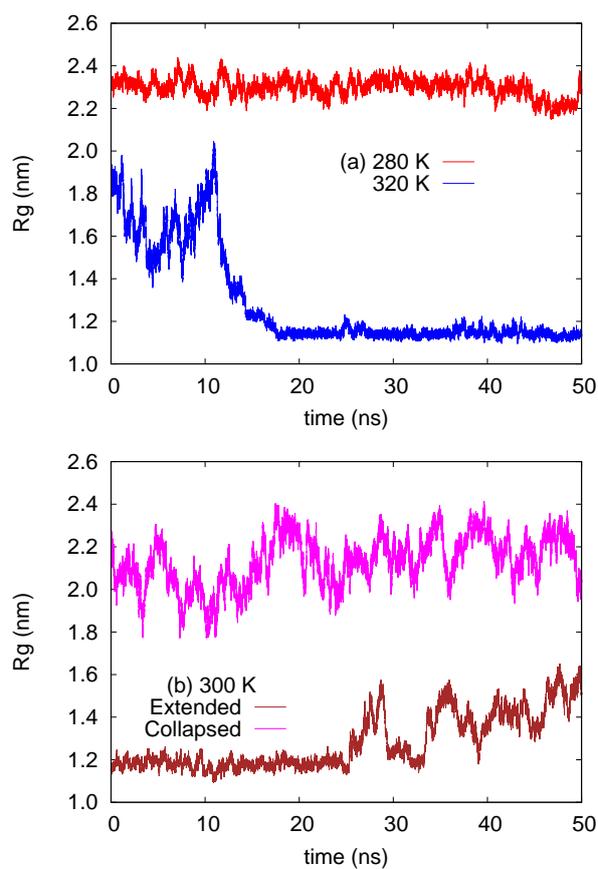


Figure S8: The time evolution of radius of gyration (R_g) of PDEA in SPC/E water. (a) shows the R_g at temperatures below (280 K) and above (320 K) the LCST of PDEA in water (≈ 308 K) starting with extended conformations, demonstrating the unfolding and collapse respectively. (b) shows the R_g at 300 K starting from extended and collapsed states of the polymer.

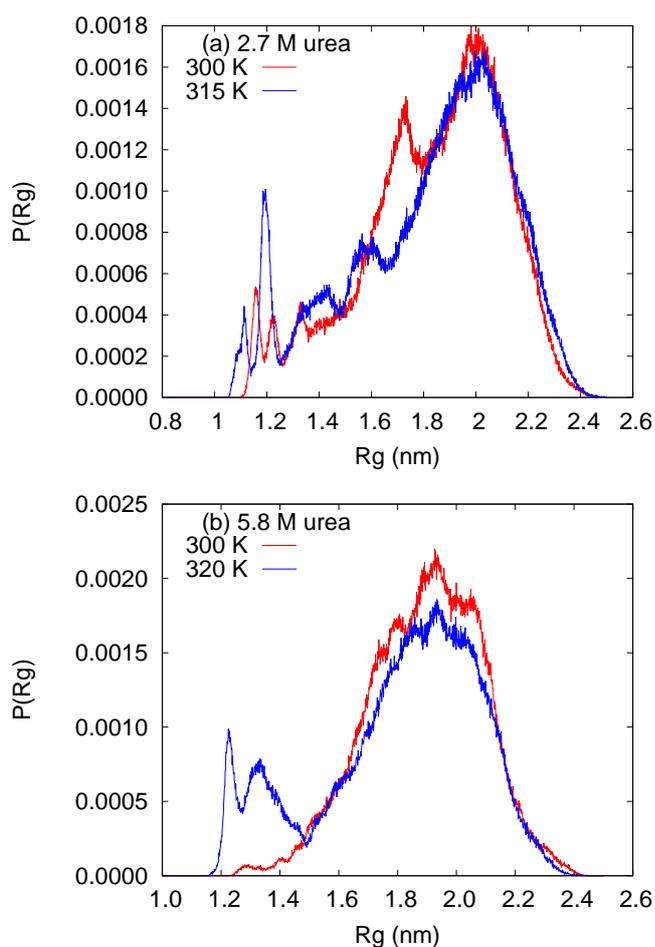


Figure S9: Normalized probability distributions of the radius of gyration (R_g) of PDEA in (a) 2.7 M urea and (b) 5.8 M urea at 300 K and the respective LCST (315 K for 2.7 M urea and 320 K for 5.8 M urea). The distributions have been plotted for the conformations sampled from 20 MD simulations starting with different extended structures, with each simulation of 50 ns run length.

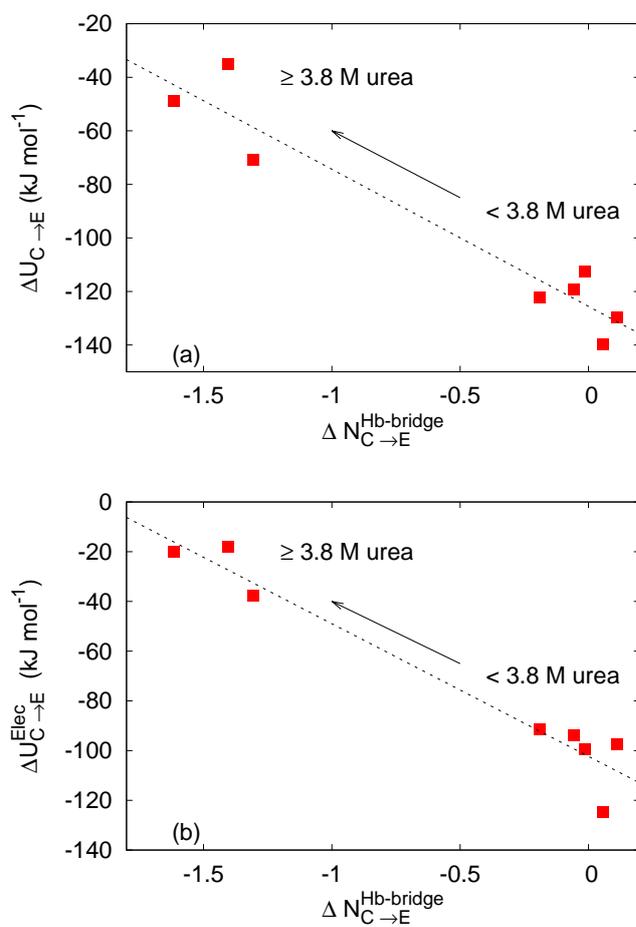


Figure S10: Correlation between the change in number of polymer-urea bridging hydrogen bonds and (a) total and (b) electrostatic polymer solvation energy upon PNiPAM unfolding in different urea concentrations. The arrow indicates the increasing urea concentration.

References

- [1] G. Graziano, *Int. J. Biol. Macromol.*, 2000, **27**, 89–97.
- [2] N. F. A. van der Vegt, D. Trzesniak, B. Kasumaj and W. F. van Gunsteren, *Chem. Phys. Chem.*, 2004, **5**, 144–147.
- [3] D. Ben-Amotz, *Annu. Rev. Phys. Chem.*, 2016, **67**, 617–638.
- [4] J. Wyman, *Adv. Protein Chem.*, 1964, **19**, 223–286.
- [5] C. Tanford, *J. Mol. Biol.*, 1969, **39**, 539–544.
- [6] V. Pierce, M. Kang, M. Aburi, S. Weerasinghe and P. E. Smith, *Cell Biochem. Biophys.*, 2008, **50**, 1–22.
- [7] J. Wang, B. Liu, G. Ru, J. Bai and J. Feng, *Macromolecules*, 2016, **49**, 234–243.
- [8] A. Shrake and J. A. Rupley, *J. Mol. Biol.*, 1973, **79**, 351–371.
- [9] F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander and M. Scharf, *J. Comp. Chem.*, 1995, **16**, 273–284.
- [10] H. H. Ku, *J. Res. Nat. Stand.*, 1966, **70C**, 263–273.