

# Molecular Description of the LCST Behavior of an Elastin-Like Polypeptide

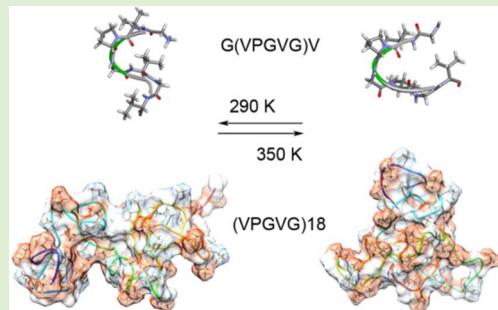
Nan K. Li,<sup>†</sup> Felipe García Quiroz,<sup>‡</sup> Carol K. Hall,<sup>§</sup> Ashutosh Chilkoti,<sup>‡</sup> and Yaroslava G. Yingling<sup>\*,†</sup>

<sup>†</sup>Department of Materials Science and Engineering and <sup>§</sup>Department of Chemical and Biomolecular Engineering, North Carolina State University, 911 Partners Way, Raleigh, North Carolina 27695, United States

<sup>‡</sup>Department of Biomedical Engineering, Duke University, P.O. Box 90281, Durham, North Carolina 27708, United States

## Supporting Information

**ABSTRACT:** Elastin-like polypeptides (ELPs) with the repeat sequence of VPGVG are widely used as a model system for investigation of lower critical solution temperature (LCST) transition behavior. In this paper, the effect of temperature on the structure, dynamics and association of (VPGVG)<sub>18</sub> in aqueous solution is investigated using atomistic molecular dynamics simulations. Our simulations show that as the temperature increases the ELP backbones undergo gradual conformational changes, which are attributed to the formation of more ordered secondary structures such as  $\beta$ -strands. In addition, increasing temperature changes the hydrophobicity of the ELP by exposure of hydrophobic valine-side chains to the solvent and hiding of proline residues. Based on our simulations, we conclude that the transition behavior of (VPGVG)<sub>18</sub> can be attributed to a combination of thermal disruption of the water network that surrounds the polypeptide, reduction of solvent accessible surface area of the polypeptide, and increase in its hydrophobicity. Simulations of the association of two (VPGVG)<sub>18</sub> molecules demonstrated that the observed gradual changes in the structural properties of the single polypeptide chain are enough to cause the aggregation of polypeptides above the LCST. These results lead us to propose that the LCST phase behavior of poly(VPGVG) is a collective phenomenon that originates from the correlated gradual changes in single polypeptide structure and the abrupt change in properties of hydration water around the peptide and is a result of a competition between peptide–peptide and peptide–water interactions. This is a computational study of an important intrinsically disordered peptide system that provides an atomic-level description of structural features and interactions that are relevant in the LCST phase behavior.



## INTRODUCTION

Elastin is a polymeric extracellular matrix protein that is responsible for the extensibility and elastic recoil exhibited by many vertebrate tissues, including skin, lungs, and larger blood vessels. Despite the great diversity in elastin's structures, some common features seem to have been preserved in elastin during evolution, such as the presence of significant amounts of glycine, proline and several aliphatic residues.<sup>1–4</sup> Elastin-like polypeptides (ELPs) are peptide polymers derived from a portion of the hydrophobic domain of elastin, which is defined by a pseudoperiodic, low complexity sequence with several types of repeat motifs. Most importantly ELPs are known to exhibit inverse temperature or lower critical solution temperature (LCST) phase behavior in aqueous solutions, that is, they undergo a first-order phase transition into polypeptide-rich and water-rich phases upon heating above the so-called cloud point temperature ( $T_{LCST}$ ).<sup>5,6</sup> This transition is essentially temperature-triggered coacervation and is reversible.<sup>7</sup> The lower critical solution temperature  $T_{LCST}$  of ELPs is dependent on pH,<sup>8</sup> salt,<sup>9</sup> polypeptide hydrophobicity,<sup>10,11</sup> and polypeptide length.<sup>12</sup> ELPs are also inherently biodegradable and biocompatible with higher sensitivity to salt than synthetic polymers such as poly(*N*-isopropylacrylamide) (PNIPAM).<sup>13</sup>

The stimulus-responsive character of ELPs has led to their use in a variety of applications, including stimuli-triggered molecular actuators for recombinant protein purification,<sup>14,15</sup> drug delivery,<sup>16,17</sup> and as stimuli-responsive materials for tissue engineering.<sup>18,19</sup> However, despite the proliferation of applications of ELPs in recent years, the origin of their LCST behavior is still a matter of controversy.<sup>20–22</sup> A detailed understanding of the physical underpinnings of this intriguing phenomenon is sorely needed to enable further progress into the development of new elastin-based materials.

ELPs with the canonical (VPGVG)<sub>n</sub> sequence seen in animal tropoelastin are widely used as the model system to investigate LCST behavior. Moreover, the LCST behavior of poly-(VPGVG) has been experimentally observed using a variety of methods including microscopy,<sup>7,10,23</sup> differential scanning calorimetry,<sup>24,25</sup> dielectric relaxation,<sup>26</sup> and different spectroscopic techniques (CD, NMR, FT-IR).<sup>27–31</sup> The fourth amino acid valine (Val) is often termed to be the “guest” residue, as other amino acid residues except proline can be substituted for

Received: May 7, 2014

Revised: August 15, 2014

it. The identity of this guest provides a precise molecular parameter to tune the  $T_{LCST}$ , as hydrophobic residues depress the  $T_{LCST}$ , while polar and charged residues raise the  $T_{LCST}$ , as compared to the parent poly(VPGVG).<sup>11</sup>

The LCST phase behavior of poly(VPGVG) is a complex and multistep phenomenon, which has been proposed by Urry et al.<sup>32</sup> to include structural transitions, loss of hydration, expulsion of water molecules and physical association of chains. However, the interplay and relevance of these factors as drivers of the LCST phase behavior of poly(VPGVG) is a matter of controversy, particularly as it relates to the structural changes accompanying the phase transition. Urry et al. carried out extensive studies on synthetic poly(VPGVG) and created a model called the  $\beta$ -spiral model that involves one type-II  $\beta$ -turn per VPGVG pentamer stabilized by intra- and interspiral and interturn hydrophobic contacts.<sup>8,10,33</sup> They proposed that upon heating  $(VPGVG)_n$  undergoes conformational changes from random coil to  $\beta$ -spiral, exposing hydrophobic amino acid side chains to the solvent, which leads to aggregation of ELPs due to hydrophobic interactions.<sup>33</sup> However, a more recent solid-state NMR study indicated that this polypeptide is in a well hydrated and hence soluble state at temperatures below  $T_{LCST}$  with relatively well-defined type-II  $\beta$ -turns centered around the Pro-Gly pair. Thus, Urry's model says that  $(VPGVG)_n$  assumes a random coil structure below the LCST and a  $\beta$ -turn rich conformation ( $\beta$ -spiral) above the  $T_{LCST}$ , but the solid state NMR data says the opposite.<sup>31,34</sup> The conformational state of an ELP below its transition temperature was also elucidated by CD spectroscopy, which showed a significant amount of type-II  $\beta$ -turns below the LCST.<sup>35</sup> In addition, a study of  $(GIVGVP)_6$  by Gross et al.<sup>36</sup> suggested that the polypeptide could adopt a  $\beta$ -sheet structure instead of a  $\beta$ -spiral in the folded state. A solid-state NMR study by Yao and Hong<sup>37</sup> suggested that Urry's model missed the more populated distorted  $\beta$ -strand structure. Clearly, there is considerable disagreement about the nature and degree of the structural changes and ordering upon heating of ELPs, which provides one of the motivations for this study.

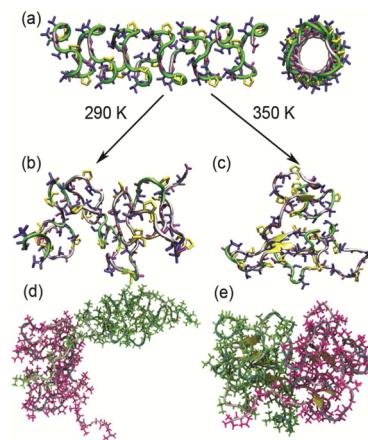
Computer simulations have also been employed to explore the LCST mechanism of poly(VPGVG) in aqueous solution. As detailed observation of phase separation is not yet possible for these systems, molecular dynamics (MD) simulation studies have been focused on the temperature-induced conformational changes of a single ELP in water. A single ELP was chosen due to current computational limitations. Simulations have been performed on short peptides, GVG(VPGVG)<sup>38</sup> and GVG-(VPGVG)<sub>3</sub>,<sup>39</sup> and a polypentapeptide  $(VPGVG)_{18}$ .<sup>40</sup> For the peptide GVG(VPGVG), extended structural conformations dominate at all temperatures and the presence of some compact structures is observed at high temperature.<sup>38</sup> The dominance of the extended conformations may be due to the very short length of this peptide, which would make it difficult to maintain the intrachain interactions required for structural compaction. Interestingly, a study of a peptide that is twice as long, GVG(VPGVG)<sub>3</sub>, over a long simulation time challenged the idea of ELP folding upon heating. It was argued that, below the transition temperature, ELP chains adopt a rigid conformational state, with no resemblance to a random coil; but at high temperatures, the peptide is highly flexible with the presence of local ordered structural elements.<sup>39</sup> In order to determine if the LCST phase behavior is an intrinsic property of the pentameric sequence or is the result of a co-operative effect of many pentamers in a polypeptide, detailed simulations of longer

chains are needed. Simulations of a much longer ELP,  $(VPGVG)_{18}$ , but for a relatively short time scale of 9 ns proposed that at high temperature ELP adopts a compact structure with distorted  $\beta$ -strands, fluctuating turns, buried hydrophobic residues, and main-chain polar atoms that form hydrogen bonds with water.<sup>40</sup> Unfortunately, the length of this simulation was too short to observe and characterize the dynamic processes of the polypeptide backbone.

This paper presents a comprehensive analysis of the temperature dependence of the secondary structure, torsion angles, hydration, and dynamics of  $(VPGVG)_{18}$  in explicit solvent by very long atomistic MD simulations. Notably, this paper is the first to examine the effect of temperature on the association between two ELP chains, which is the necessary first step in understanding the molecular determinants of the LCST behavior of polypeptides.

## MATERIALS AND METHODS

**Single ELP Simulation.** Atomistic MD simulations were performed using Amber 11<sup>41</sup> and the ff99SB force field for proteins. An initial model of Urry's  $\beta$ -spiral for  $(VPGVG)_{18}$  was solvated in explicit water using the TIP3P water model<sup>42</sup> (Figure 1a). The



**Figure 1.** (a) Side and front view of the  $\beta$ -spiral structure for  $(VPGVG)_{18}$ . Final simulation snapshots of (b, c) a single  $(VPGVG)_{18}$  structure and (d, e) the interaction between two ELPs at (b, d) 290 K and (c, e) 350 K. The backbone is represented as secondary structure ribbons where turn motif is colored in green,  $\beta$ -sheet is yellow, and coil is silver. All heavy atoms are depicted as sticks. In panels a–c, the side chains are colored by amino acid (Gly in purple, Pro in yellow, and Val in blue). In panels d and e, the side chains are colored by molecule.

simulation box of about  $5.9 \text{ nm} \times 7.7 \text{ nm} \times 5.0 \text{ nm}$  size contained a single peptide and 6856 water molecules. The size of the box was chosen to be large enough so that no peptide-peptide interactions through periodic boundary conditions can occur, which represent an infinitely dilute state. The simulations indicated that the initial configuration of the idealized  $\beta$ -spiral model is not stable at all temperatures, which is in agreement with previous simulations<sup>39,40</sup> (Figure 1). In order to span the temperature range that is relevant to the experimental study of LCST phase transitions in aqueous solutions, simulations were performed at 10 temperatures between 290 and 350 K.

In all cases, each system was equilibrated in eight stages starting from a solvent minimization using steepest descent method for 10000 steps while keeping the all-peptide atoms restrained with 200 kcal/mol. The system was then gradually heated to an assigned temperature for 10 ps while maintaining the 200 kcal/mol constraint on the peptide. A short 40 ps isothermal-isobaric (NPT) ensemble MD run was then performed, again with the peptide restraint maintained at 200

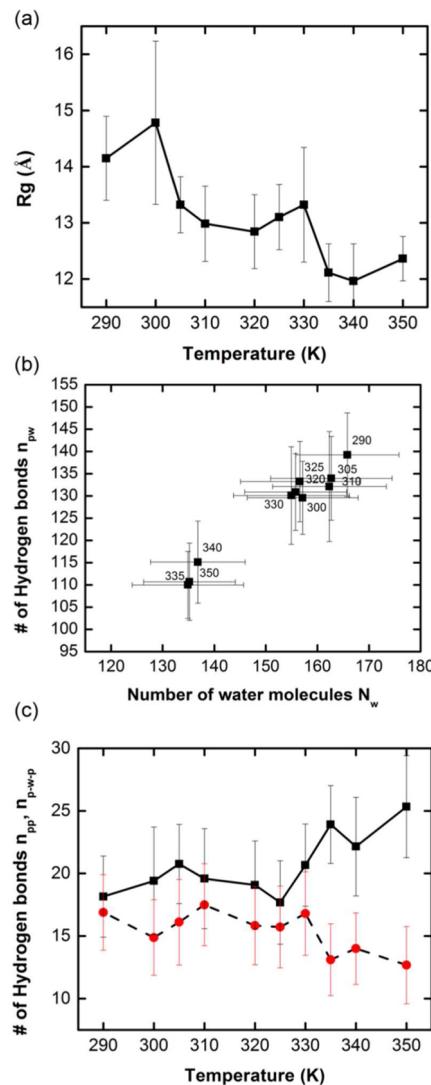
kcal/mol. Another minimization step followed for 10000 steps with the restraint of 25 kcal/mol. A second NPT MD run was performed at 25 kcal/mol restraint for 20 ps. Subsequently, a final unconstrained minimization of 1000 cycles was performed before reheating the system to the assigned temperature at constant volume for 40 ps. The NPT ensemble was adopted for the equilibration and production MD run to ensure uniformity in solvent density. The long-range electrostatic interactions were calculated by particle mesh Ewald (PME) summation<sup>43</sup> and the nonbonded interactions were truncated at a 9 Å cutoff along with a 0.00001 tolerance for Ewald convergence. The temperature was maintained using a Berendsen thermostat.<sup>44</sup> The SHAKE algorithm was used to constrain bonds involving hydrogen atoms.<sup>45</sup> The production simulations were performed for at least 70 ns with a 2 fs time step. Only the last 40 ns of the trajectories from each case were considered for statistical analysis. The convergence of the simulations was assured by the data presented in the Supporting Information (Figures S3b, S4, S5, and S7). MD trajectories were processed using in-house scripts along with the standard tool suite accompanying Amber11.0. The interaction energy was calculated using the molecular mechanics energy function in NAMD 2.7.<sup>46</sup> The hydrogen bond analysis was performed using a distance cutoff of 3.5 Å and an angle cutoff of 30°.

**Interactions between Two ELPs.** In order to get a better understanding of how the temperature-dependent single peptide properties can cause the aggregation between peptides, we simulated the interaction between two temperature-equilibrated peptides. The initial configurations of peptides used in these simulations were chosen through the application to the MD simulation trajectories described above of a hierarchical RMSD-based cluster algorithms<sup>47</sup> following by the energetic analysis (Figure S1, Table S1). Specifically, the last 40 ns of a trajectory from the single-peptide MD simulations are clustered to produce three structural clusters using the pairwise RMSD between frames as a metric comparing the atoms named CA with a critical distance of 12 Å. Six representative structures from the most populated clusters at 290 and 350 K are shown in Figure S1. The NAMD energy function were used on these six structures to identify the lowest energy structure, which is shown on Figure S1(b) for 290 K and Figure S1(f) for 350 K. Then, the two chosen single polypeptide structures were placed side-by-side with a 32.3–32.4 Å distance between the centers of mass of the polypeptides and a 7 Å distance between the two closest surfaces. The structures were then solvated in explicit solvent with the closest distance between any solute atom and the edge of the periodic box to be 8 Å and simulated at 290 and 350 K. The simulation box contained 11453 water molecules with the size of 5.6 nm × 7.9 nm × 8.3 nm at 290 K and 8291 water molecules with size of 5.0 nm × 7.0 nm × 8.2 nm at 350 K. Equilibration protocols and MD simulations were the same as described in the section above. The production simulations were performed for 20 ns with a 2 fs time step.

## RESULTS

**Conformation of a Single (VPGVG)<sub>18</sub> Polypeptide.** Our simulations show that in aqueous solution, a single ELP molecule adopts a collapsed state over a temperature range between 290 and 350 K (Figure 1b,c and Figure S2). Previous simulations of the LCST behavior of ELP<sup>40</sup> and PNIPAM<sup>48</sup> showed that the value of the radius of gyration ( $R_g$ ) dramatically drops around a critical temperature; we expect that this is most likely an artifact of the short simulation time, particularly at low temperature. In contrast, our simulations of a single (VPGVG)<sub>18</sub> molecule indicate a gradual decrease in the average value of  $R_g$  as the temperature increases (Figure 2a). The temporal evolution of  $R_g$  at various temperatures is displayed in Figure S3 and indicates that it takes much longer for a polypeptide to assume a stable conformation at lower temperatures than that at high temperatures.

The change in  $R_g$  is related to the change in interactions between chemical groups in the polypeptide and the solvent—

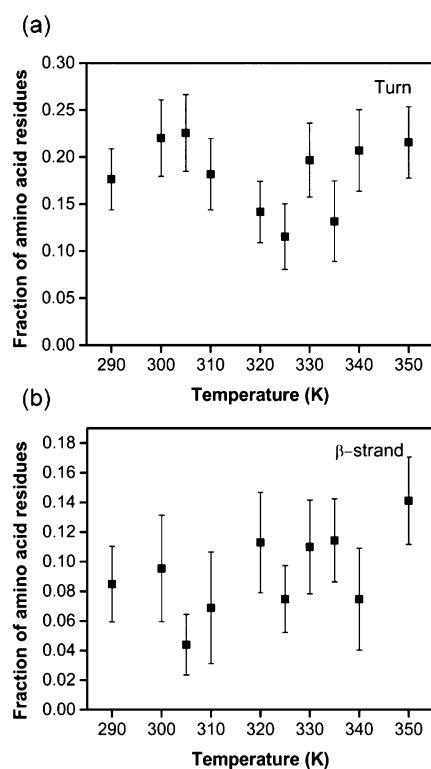


**Figure 2.** (a) Temperature dependence of the radius of gyration ( $R_g$ ) of (VPGVG)<sub>18</sub>. (b) Correlation between  $N_w$ , the number of water molecules in the second watershell of backbone, and  $n_{pw}$ , the number of peptide–water hydrogen bonds. The temperature is labeled at the data points. (c) Number of the intramolecular intrapeptide hydrogen bonds (black) and intramolecular water-mediated peptide hydrogen bonds (red). The error bars in these plots represent the standard deviation.

water molecules. To examine the contribution of H-bonding at different temperatures, we examine the relationship between the number of polypeptide–water H-bonds,  $n_{pw}$ , and the number of waters of hydration  $N_w$  in Figure 2b. The number of hydration waters is defined as the number of water molecules within a 3.1 Å distance of any atom on the peptide's backbone. This distance is taken to be the distance at which the radial distribution function (RDF) between the oxygen of water and an atom on the peptide backbone has its second minimum (Figure S5). Our simulations reveal that as the temperature increases, both the number of hydrating water molecules and the number of peptide–water hydrogen bonds decreases (Figure 2). Moreover, the presence of two distinct clusters one above and one below 330 K (Figure 2b) indicates that there is a strong temperature dependence of peptide–water interactions properties.

The number of intrapeptide hydrogen bonds (Figure 2c) with ( $n_{p-w-p}$ ) and without ( $n_{pp}$ ) mediated water demonstrates that as the temperature increases the intramolecular peptide hydrogen bonding tends to increase and water-mediated hydrogen bonding decreases. Our observation of the temperature-induced increase in intramolecular peptide bonds agrees well with the experimental observation that peptides at high temperature have more  $\beta$ -turns or  $\beta$ -strands.<sup>36,49</sup> The decrease in the number of water-mediated intramolecular peptide hydrogen bonds also correlates well with the possible expulsion of localized waters at higher temperatures. Overall, the changes in the peptide hydration and the strength of intramolecular peptide interactions become accentuated in a critical temperature window (i.e., at around 330~335 K in our model polypeptide).

**Secondary Structure Analysis.** In order to examine the structural propensities of  $(VPGVG)_{18}$  as a function of temperature (Figure 3) we utilized the DSSP method,<sup>50</sup>



**Figure 3.** Secondary structure formation as a function of temperature for (a)  $\beta$ -turn and (b)  $\beta$ -strand structures.

which is a widely used structure recognition algorithm mainly based on H-bonding patterns. In Figure 3,  $\beta$ -turns are defined as having hydrogen bonds between residues  $i$  and  $i+n$ , where  $n = 3, 4, 5$  (Figure 3a) and  $\beta$ -strands are defined to include  $\beta$ -bridges and parallel and anti-parallel  $\beta$ -sheets (Figure 3b). We found that as the temperature increases, more  $\beta$ -strands form, yet the number of turns in a single polypeptide shows a nonmonotonic change. At low temperatures, it is possible that turns are likely to be stabilized by hydrogen bonding due to the low thermal energy. Increasing the temperature and hence the thermal energy may decrease the stability of intrachain hydrogen bonding which may lead to a decrease in the number of turn structures at higher temperatures. We observed that the number of turns increases again at even higher temperatures

when H-bonding with water decreases and intramolecular hydrogen bonding increases, which is possibly due to the loss of hydration at high temperatures.

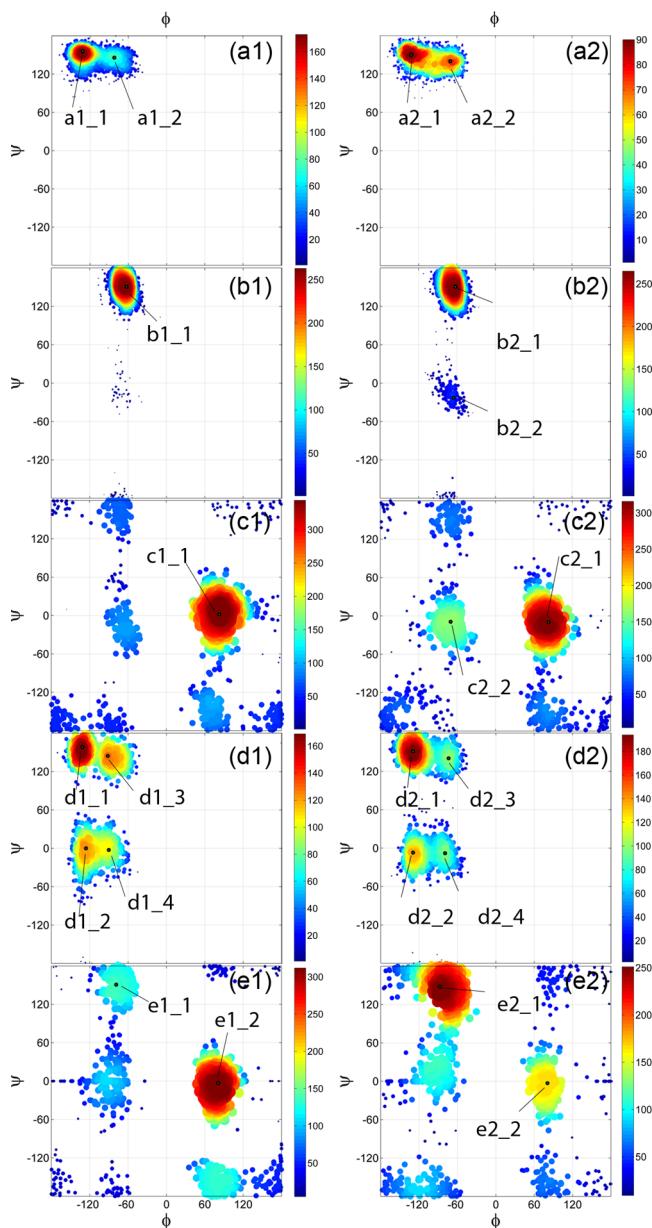
The occurrence frequency of secondary structural motifs for each residue is presented in Figure S8. We observe that Pro and Gly residues involved in turn formation are highly dynamic and undergo rapid interconversion.<sup>51</sup> A small number of helical structures form at the polypeptide termini at all temperatures (Figure S8). Overall, our observations of temperature-dependent secondary structure propensities are in general agreement with the previously discussed experimental observations.<sup>36,37</sup>

In order to compare our observations with available NMR data, we calculated the torsion angles for each type of amino acid (Val1, Pro2, Gly3, Val4, and Gly5) in  $(VPGVG)_{18}$  at low (290 K) and high (350 K) temperatures (Figure 4). We compare our results to the SS-NMR data for the VPGVG segment in  $(VPGVG)_3^{37}$  and  $[(VPGVG)_4(VVPKG)_3]^{31}$  obtained by Hong et al., and the SS-NMR data for the middle segment in  $(VPGVG)_6^{34}$  obtained by Ohgo et al. These studies indicate a significant structural heterogeneity and neither study supported the torsion angles of the  $\beta$ -spiral.<sup>33</sup> Although there is a reasonable agreement between these studies on the overall structural propensities of the VPGVG subunit, the observed differences are primarily in the torsion angles for the valine residues in the VPGVG repeat unit.

Our study indicates that the distribution of dihedral angles is a function of temperature and that this temperature sensitivity varies from residue to residue in the pentapeptide motif. The Pro2 and Gly3 are the central two residues,  $i+1$  and  $i+2$ , of a  $\beta$ -turn structure. In Figure 4, the regions represented by labels b1\_1 and b2\_1 include the torsion angle pairs for the  $i+1$  residue of a type-II  $\beta$ -turn ( $\varphi = -60^\circ, \psi = 120^\circ$ ). Our results indicate that, in general, proline in  $(VPGVG)_{18}$  tends to adopt the torsion angles for the type II  $\beta$ -turn, a polyproline helix ( $\varphi = -75^\circ, \psi = 150^\circ$ ) or a collagen helix ( $\varphi \sim -75^\circ, \psi \sim 160^\circ$ ). At high temperature (350 K), proline can also adopt torsion angles around the point b2\_2 ( $-65.4^\circ, -23.02^\circ$ ), which is consistent with the torsion angles for the  $i+1$  residue of a type I  $\beta$ -turn. Ohgo et al. also observed Pro2 angles near ( $-60^\circ, -30^\circ$ ) by NMR on the middle segment of  $(VPGVG)_6^{34}$ .

As expected, Gly3 can adopt many different conformations. However, within  $(VPGVG)_{18}$ , it preferentially adopts conformations in the regions marked as c1\_1 ( $\varphi = 83^\circ, \psi = 2^\circ$ ) at 290 K and c2\_1 ( $\varphi = 83^\circ, \psi = -10^\circ$ ) at 350 K. Since the torsion angles of  $i+2$  residues in type-II  $\beta$ -turns are  $\varphi = 80^\circ$  and  $\psi = 0^\circ$ , the Pro2-Gly3 pair favors the type II  $\beta$ -turn structures at low and high temperatures. At high temperature Gly3 also adopts torsion angles labeled as c2\_2 ( $\varphi = -70^\circ, \psi = -9^\circ$ ), which are consistent with  $i+2$  residue in type-I  $\beta$ -turn ( $\varphi = -80^\circ, \psi = 0^\circ$ ) structures. The data for Pro2 and Gly3 indicate that  $(VPGVG)_{18}$  adopts a number of type I  $\beta$ -turn structures at high temperatures in addition to the expected type II  $\beta$ -turn structures.

The highest intensity regions for valines indicate that both Val1 and Val4 tend to adopt  $\beta$ -strand conformation, which corresponds to torsion angles of about  $\varphi = -130^\circ$  and  $\psi = 150^\circ$ . Val1 conformations are more restricted than Val4, due to the steric hindrance from the pyrrolidine ring of proline. In a solid state NMR study by Yao and Hong,<sup>37</sup> the torsion angles of the Val1 residue were reported to be ( $-96^\circ, 145^\circ$ ) or ( $-144^\circ, 145^\circ$ ), which agrees well with our results. For Val1, the a2\_2 ( $-70^\circ, 140^\circ$ ) region has higher intensity at 350 K than that at 290 K which indicate that Val1 has propensity to form



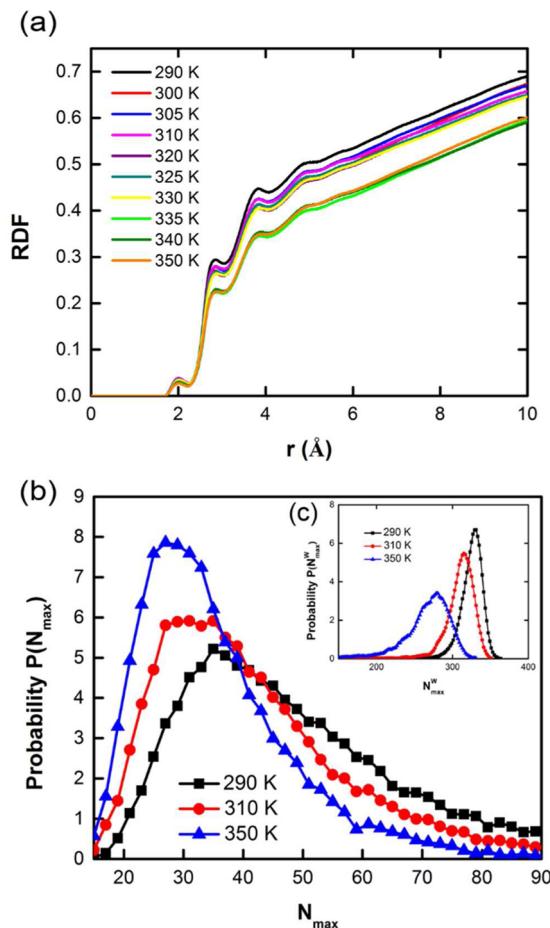
**Figure 4.** Ramachandran plot ( $\phi$ ,  $\psi$  distributions) for each residue in a pentamer, (a1,2) Val1, (b1,2) Pro2, (c1,2) Gly3, (d1,2) Val4, (e1,2) Gly5, at two temperatures, 290 K (left column) and 350 K (right column), colored by intensities. The most populated regions on the Ramachandran plots are labeled.

more  $\beta$ -turn structures at high temperature. In contrast with Val1, Val4 torsion angles are distributed between four regions where  $\varphi$  can be  $-130^\circ$  or  $-90^\circ$  and  $\psi$  can be  $150^\circ$  or  $-5^\circ$ . The experimental studies by Yao and Hong<sup>37</sup> and Ohgo et al., which disagree with each other, in combination show four different angle propensities for Val4, which are  $(-148, 145)$ ,<sup>37</sup>  $(-92, 145)$ ,<sup>37</sup>  $(-110, 130)$ ,<sup>34</sup> and  $(-75, -15)$ .<sup>34</sup> These four high intensity regions roughly agree with our observation. Moreover, we show that at high temperatures, Val4 preferentially adopts  $\beta$ -strand structures which is indicated by the higher intensity of  $\varphi$  angles around  $-130^\circ$  (d2\_1) in combination with the reduced intensity of d2\_3 peak at 350 K compared to 290 K.

The Ramachandran plots for Gly5 and Gly3 are very similar at low temperatures and indicate their propensity to form  $\beta$ -

turns. Both glycine residues at 290 K have similar torsion angles with the highest occupied region at around  $(80^\circ, 0^\circ)$ , which corresponds to the torsion angle of  $i + 2$  residues in type II  $\beta$ -turns. However, their angle distributions show different structural preferences at high temperatures. For Gly5, the most populated region at low temperature is around e1\_2 of  $(81^\circ, -3^\circ)$  and at high temperature is around e2\_1 of  $(-88^\circ, 148^\circ)$ . Thus, at high temperature Gly5 tend to preferentially adopt an extended conformation ( $\varphi = 150^\circ$ ). Overall, the observed increase in  $\beta$ -strands at high temperatures (Figure 3b) may be attributed to the conformational changes of Val1, Val4, and Gly5. Based on the dihedral angles obtained from the highest intensity points in Ramachandran plots (Figure 4), the representative structures of G(VPGVG)V segment at 290 and 350 K are built and shown in Figure S9.

**Structure of Hydration Water.** The structural properties of water close to the surface of the polypeptide can be quantified by the RDF between the polypeptide and water (Figure 5a, Figure S5, and Table S2). We observed that the locations of the first ( $r_{\min} = 2.23 \text{ \AA}$ ) and second ( $r_{\text{sec}} = 3.05 \text{ \AA}$ ) local RDF minima are essentially independent of temperature and are primarily due to water–backbone interactions (Figure S5). There is a slight gap in the RDF peaks intensity between the temperatures of 330 and 335 K, which corresponds to the



**Figure 5.** (a) Radial distribution function (RDF) of oxygen atoms in water molecules around polypeptide atoms as a function of temperature. (b) Probability distribution of the size  $N_{\max}$  of the largest water network around polypeptide. (c) Probability distribution of the size  $N_{\max}^w$  of the largest water network in pure water systems.

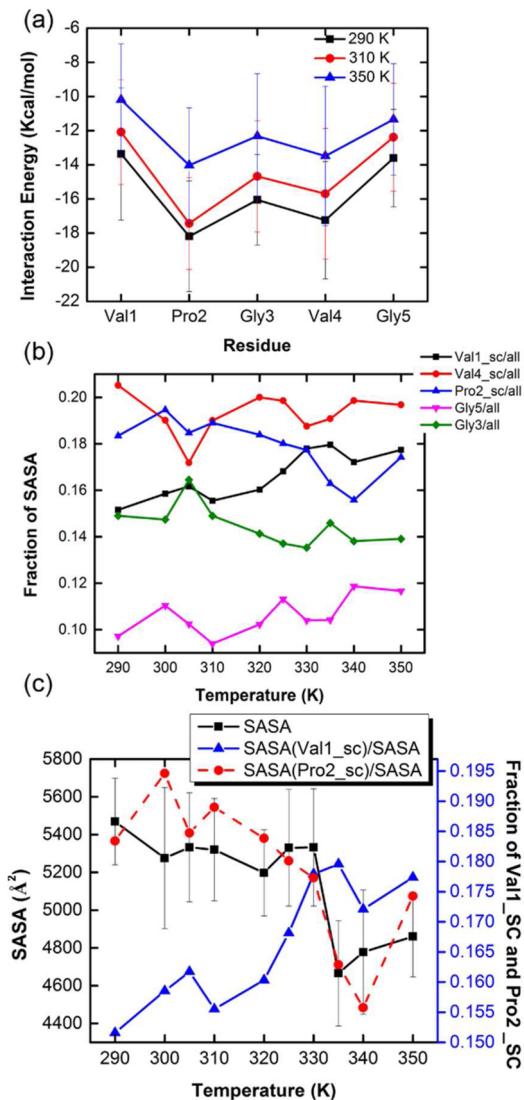
same temperature window where change in the ELPs conformational properties has occurred. An increase in temperature leads to a gradual loss of local hydration layers, as indicated by the decrease in the heights of the hydration peaks at higher temperature (Figures 5a and S5).

The existence of a hydration shell around the polypeptide is considered to be essential for the LCST behavior of  $(\text{VPGVG})_n$  ELPs.<sup>52,53</sup> The state of the hydrogen-bonded water network in the hydration shell of the short peptide  $\text{GVG}(\text{VPGVG})_3$  at various temperatures was studied by percolation analysis, where the thermal breaking of the spanning network of the hydration water upon heating was observed.<sup>39</sup> It was proposed that the thermal breaking of the H-bonded network can cause rapid changes of the thermodynamic properties of hydration water which make the surface of the peptide effectively more hydrophobic.<sup>52,54</sup>

In our simulations, the 400 water molecules that are closest to the polypeptide, which includes the first and second hydration layers, are located within 3.23 Å from the peptide atoms at 290 K and expand to a distance of 3.5 Å at 350 K. Two water molecules were considered as hydrogen bonded when  $R_{\text{OO}} \leq 3.5$  Å and  $\varphi_{\text{O}_1\cdots\text{O}_2\cdots\text{H}_2} \leq 30^\circ$ . We chose this criterion from the simulations of pure water at various temperatures, where each water molecule experiences an average of about 2.24 H-bonds at 300 K and 2.05 H-bonds at 350 K.<sup>55</sup> The size of the largest H-bonded water network in the closest water shell was used to estimate the connectivity between the water molecules around the ELP. A total of  $1.6 \times 10^4$  configurations after equilibrium were examined to identify the largest network of water molecules at each temperature. Probability distribution of the largest water network around the polypeptide at  $T = 290$  K, 310 and 350 K are displayed in Figure 5b. The profile at lower temperature not only has a larger average network size, but also is broader than the one at higher temperature. Thus, at lower temperature, hydration water molecules appear to be more ordered than at higher temperatures, which corresponds to a decrease in both enthalpy and entropy of hydration water molecules.<sup>52,54</sup> In order to determine the effect of the peptide presence on the network formation of water molecules we compared the water network distributions with and without the peptide. Generally, the temperature increase in bulk water will also reduces the size of the water network (Figures 5c and S6). However, the presence of the peptide clearly changes the size distribution of these clusters. Specifically, at high temperatures the presence of a peptide induces the formation of many small water clusters, which is indicated by the narrow distribution as compared to the one for bulk water.

**Solvent Accessible Surface Area and Energy.** To learn what role each amino acid plays in the VPGVG repeat, the interaction energies of each amino acid residue in the repeat unit of the ELP with water molecules were calculated (Figure 6a). The interactions between each amino acid and water molecules weakens as the temperature is increased. Interestingly, proline interacts with water more strongly than other amino acids, regardless of the temperature (Figure 6a). Also, Val1 is generally more hydrophobic than Val4 possibly due to the steric restriction by Pro2.

The solvent accessible surface area (SASA) is another important element in the analysis of protein folding and protein–protein interactions. The concept of accessible surface area also provides a convenient way to define the peptide surface and interior. The SASA is not just a geometric measure, but has physical significance as a gain in hydrophobic



**Figure 6.** Temperature dependence of properties of a single ELP chain: (a) Interaction energy of each individual residue in an ELP pentamer with water. (b) Fraction of SASA of Val1\_side chain (black), Pro2\_side chain (blue), Val4\_side chain (red), Gly3 (green), and Gly5 (pink). (c) SASA of peptides (black; square) and SASA fraction of Val1\_side chain (blue; triangle) and Pro2\_side chain (red; circle).

interaction free energy is accompanied by a reduction in the SASA.<sup>56</sup> The SASAs of the polypeptide backbone, side chain, and various chemical groups are listed in Table S3. The percentage of buried surface area for  $(\text{VPGVG})_{18}$  is between 43 and 51% of the SASA of extended polypeptide over the studied temperature range (Table S3). This fraction of buried SASA is slightly lower than that of a globular 100-residue protein, where ~55% of the polypeptide accessible surface becomes buried upon folding.<sup>57</sup> This can be explained by the lack of a hydrophobic core in VPGVG-ELPs due to the regular distribution of nonpolar side chain of each pentamer.

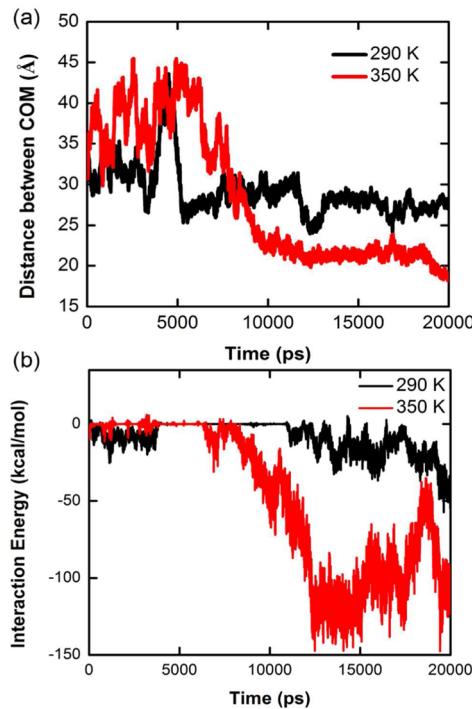
In order to provide a quantitative relationship between the SASA of side chains and the temperature, the SASA fraction associated with the valine and proline side chains and the SASA fraction for glycines are plotted in Figure 6b. We observed that as the temperature increases the SASA fraction for the Val1 side chain and for Gly5 increases, which also correlates with the changes in their torsional angles. This can be explained by an

increase in the local turn structures of Val1 and Gly5 at high temperature, which leads to a reduction in structural heterogeneity of the chain.

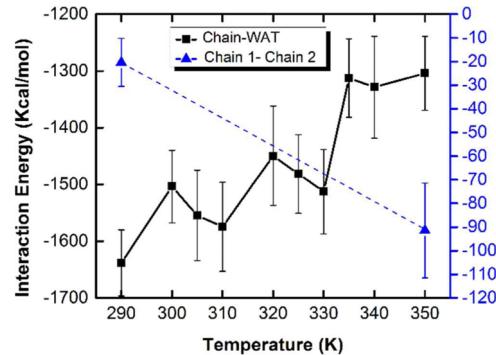
Our simulations suggest that the LCST behavior of  $(\text{VPGVG})_{18}$  may be associated with an abrupt decrease in SASA (Figure 6c) which is correlated with the behavior of Val1 and Pro2 side chains. Specifically, as the temperature increases more side chains of Val1 residues become accessible for water and Pro2 residues are hidden from water. Generally the side-chain mobility increases as the temperature increases, however in our case the observed reduction in ELP radius of gyration reduces the side-chain motion. The observed thermoresponsive changes in SASA of Val1 and Pro2 side chains in the polypeptide are due to the competition between these two effects. Our simulation also suggests that increasing temperature leads to less heterogeneity of the peptide structure.

**Polypeptide Aggregation.** The temperature-triggered coacervation of ELPs has been described as a concentration-dependent phase separation,<sup>7</sup> where hydrophobic intermolecular association between peptides becomes the dominant process. The formation of microscale aggregates by systems containing multiple polypeptide chains occurs on a time scale that is not readily accessible in atomistic MD simulations. Therefore, to examine whether the temperature-dependent changes at a single peptide level will lead to aggregation, we conducted simulations of two temperature-equilibrated polypeptides at low temperature (290 K) and at high temperature (350 K). Initially, the polypeptides, despite being spatially close, lack the necessary molecular contacts that would bias their aggregation into a compact assemble. This simulation was performed to reveal the early stage of polypeptide aggregation. Figure 1 shows snapshots of the two polypeptides at 290 and 350 K after 15 ns of simulation time. At high temperature (350 K), the two polypeptides form a strongly compacted aggregate (Figure 1e), with water expelled from the interface between the two molecules. At low temperature (290 K), the polypeptides are in a much more extended conformation and can intermittently interact with each other due to the periodic boundary conditions (Figure 1d). Thus, upon transition from low to high temperature, the system undergoes a transition from a disordered state to an aggregate with a higher degree of compactness. The distance between the center-of-mass of the two chains is displayed in Figure 7a to show the dynamics of aggregation. The interaction energy profiles between the polypeptides for both cases (Figure 7b) correlates with the distance between their centers of mass and shows that the chains interact strongly at high temperatures. During the simulations, no significant changes in the radius of gyration and secondary structure propensities were observed (Figure S10).

The increase in hydrophobicity with temperature of a single peptide may play a key role in the process of aggregation. The correlation between the interaction energies between water and the peptide and between the two polypeptides is plotted in Figure 8. The temperature increase leads to a reduction in water-peptide interactions (i.e., increase in hydrophobicity of a single polypeptide) and, consequently, to a strong association between the polypeptides. A single polypeptide molecule in water is more stable at low temperature than high temperature, while a polypeptide aggregate is less stable at low temperature than at high temperature.



**Figure 7.** Interaction between two ELPs. (a) Distance between center-of-mass of two polypeptides and (b) peptide–peptide interaction energy at 290 and 350 K.



**Figure 8.** Peptide–water interaction energy in single-peptide system (black; square). Peptide–peptide interaction energy in double-peptide system (blue; triangle). The data are averaged over the last 40 ns for each MD run in a single chain system and over the last 5 ns run for double chain system.

## CONCLUSIONS

Using long atomistic molecular dynamics simulations of  $(\text{VPGVG})_{18}$  in aqueous solutions, we investigated the temperature-dependent structure and dynamics of the polypeptide chain and its hydration water, in order to understand the mechanisms underlying the LCST behavior of VPGVG polypeptides. Based on our simulations, we conclude that the transition behavior includes thermal disruption of the water network, loss of hydration of the polypeptide, and an increase in its hydrophobicity, leading to physical association of chains.

Our simulations show that at high temperatures, an ELP molecule tends to form a more ordered secondary structure in the form of  $\beta$ -strands. Detailed analysis of residue-based torsion angles in the VPGVG repeat unit shows that  $\beta$ -turn structures exist at all temperatures. Specifically, the Pro2-Gly3 pair can

adopt type II  $\beta$ -turn structures at all temperatures. At high temperatures there is also a number of type I  $\beta$ -turns formed for Pro2-Gly3. In contrast, Val1 adopts  $\beta$ -strand conformations at all temperatures and more turn conformations at high temperatures. However, the formation of  $\beta$ -strand structures becomes more prominent for Val4 and Gly5 at high temperatures. The presence of prolines and glycines generally limits the formation of extended secondary structures due to the steric constraints imposed by prolines and the high entropic penalty of glycine confinement.<sup>49,58</sup> However, this limitation can be compromised when the temperature is raised,<sup>2</sup> which allow for more  $\beta$ -strand and distorted  $\beta$ -strand structures to be adopted at high temperatures.

At high temperatures we observed thermal disruption of the hydrogen-bonded water network around the polypeptide and an abrupt decrease in the polypeptide's SASA, which are accompanied by increases in secondary structure formation and exposure of hydrophobic side chains as well as displacement of local hydrating water molecules into the bulk. From the traditional peptide backbone centric view, where the peptide backbone is considered responsible for the resultant conformation, the role of hydrophobicity is nonspecific. However, from the peptide side chain centric view, hydrophobicity is the dominant force for protein folding.<sup>59</sup> Moreover, proline in elastomeric proteins is considered to be a "gatekeeper" residue, which maintains the disordered structure and prevents collapse into a hydrophobic core.<sup>2,49,60</sup> Interestingly, we observed a distinct change in the exposure of Val1 and Pro2 hydrophobic side chains at high temperatures. Thus, as the temperature increases, ELPs undergo conformational changes of the backbone along with the hydrophobicity changes of the side chains, exposing hydrophobic valine side chain to the solvent and hiding the "gatekeeper" proline residue.

Simulations of the initial stages of the aggregation process between two polypeptides demonstrated that the observed gradual changes in the properties of a single polypeptide with temperature are responsible for the experimentally observed aggregation of ELPs at or above LCST. Based on analysis of the interaction energy, we suggest that the competition between peptide-peptide and peptide-water interactions may determine the LCST of the system. Overall, we propose that the LCST in VPGVG polypeptides is a collective phenomenon that originates from gradual changes in the structure of single polypeptide chains and the abrupt change in properties of hydration water around the peptide as temperature increases. Because we only observed gradual structural changes with temperature at the single peptide level, this new molecular level understanding of the structure, dynamics, and thermodynamics of ELP (VPGVG)<sub>18</sub> as a function of temperature reinforces the concept of cooperative nature of the pronounced phase transition exhibited by VPGVG polypeptides.

## ASSOCIATED CONTENT

### Supporting Information

Additional information about the simulation systems, convergence of simulations, and analysis. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: yara\_yingling@ncsu.edu.

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by the NSF's Research Triangle MRSEC (DMR-1121107). The computer support was provided by the High Performance Computing Center at North Carolina State University.

## REFERENCES

- (1) Bressan, G. M.; Argos, P.; Stanley, K. K. *Biochemistry* **1987**, *26*, 1497–1503.
- (2) Muiznieks, L. D.; Keeley, F. W. *Biophys. J.* **2011**, *100*, 199–199.
- (3) Debelle, L.; Tamburro, A. M. *Int. J. Biochem. Cell Biol.* **1999**, *31*, 261–272.
- (4) He, D.; Chung, M.; Chan, E.; Alleyne, T.; Ha, K. C. H.; Miao, M.; Stahl, R. J.; Keeley, F. W.; Parkinson, J. *Matrix Biol.* **2007**, *26*, 524–540.
- (5) Urry, D. W. *Sci. Am.* **1995**, *272*, 64–69.
- (6) Martino, M.; Perri, T.; Tamburro, A. M. *Macromol. Biosci.* **2002**, *2*, 319–328.
- (7) Urry, D. W. *Faraday Discuss.* **1976**, *61*, 205–212.
- (8) Urry, D. W. *Angew. Chem., Int. Ed.* **1993**, *32*, 819–841.
- (9) Reguera, J.; Urry, D. W.; Parker, T. M.; McPherson, D. T.; Rodriguez-Cabello, J. C. *Biomacromolecules* **2007**, *8*, 354–358.
- (10) Urry, D. W. *J. Phys. Chem. B* **1997**, *101*, 11007–11028.
- (11) Urry, D. W.; Luan, C. H.; Parker, T. M.; Gowda, D. C.; Prasad, K. U.; Reid, M. C.; Safavy, A. *J. Am. Chem. Soc.* **1991**, *113*, 4346–4348.
- (12) Meyer, D. E.; Chilkoti, A. *Biomacromolecules* **2004**, *5*, 846–851.
- (13) Meyer, D. E.; Shin, B. C.; Kong, G. A.; Dewhirst, M. W.; Chilkoti, A. *J. Controlled Release* **2001**, *74*, 213–224.
- (14) Meyer, D. E.; Chilkoti, A. *Nat. Biotechnol.* **1999**, *17*, 1112–1115.
- (15) Ge, X.; Yang, D. S. C.; Trabucco-Carlson, K.; Kim, B.; Chilkoti, A.; Filipe, C. D. *M. J. Am. Chem. Soc.* **2005**, *127*, 11228–11229.
- (16) McDaniel, J. R.; Bhattacharyya, J.; Vargo, K. B.; Hassouneh, W.; Hammer, D. A.; Chilkoti, A. *Angew. Chem., Int. Ed.* **2013**, *52*, 1683–1687.
- (17) Dreher, M. R.; Raucher, D.; Balu, N.; Colvin, O. M.; Ludeman, S. M.; Chilkoti, A. *J. Controlled Release* **2003**, *91*, 31–43.
- (18) Lim, D. W.; Nettles, D. L.; Setton, L. A.; Chilkoti, A. *Biomacromolecules* **2007**, *8*, 1463–1470.
- (19) McHale, M. K.; Setton, L. A.; Chilkoti, A. *Tissue Eng.* **2005**, *11*, 1768–1779.
- (20) Tamburro, A. M. *Nanomedicine* **2009**, *4*, 469–487.
- (21) Chen, Y. L.; Guan, Z. B. *J. Am. Chem. Soc.* **2010**, *132*, 4577–4579.
- (22) Amiram, M.; Quiroz, F. G.; Callahan, D. J.; Chilkoti, A. *Nat. Mater.* **2011**, *10*, 141–148.
- (23) Urry, D. W.; Long, M. M.; Cox, B. A.; Ohnishi, T.; Mitchell, L. W.; Jacobs, M. *Biochim. Biophys. Acta* **1974**, *371*, 597–602.
- (24) Luan, C. H.; Urry, D. W. *J. Phys. Chem.* **1991**, *95*, 7896–7900.
- (25) Rodriguez-Cabello, J. C.; Alonso, M.; Perez, T.; Herguedas, M. M. *Biopolymers* **2000**, *54*, 282–288.
- (26) Henze, R.; Urry, D. W. *J. Am. Chem. Soc.* **1985**, *107*, 2991–2993.
- (27) Pepe, A.; Armenante, M. R.; Bochicchio, B.; Tamburro, A. M. *Soft Matter* **2009**, *5*, 104–113.
- (28) Lessing, J.; Roy, S.; Reppert, M.; Baer, M.; Marx, D.; Jansen, T. L. C.; Knoester, J.; Tokmakoff, A. *J. Am. Chem. Soc.* **2012**, *134*, 5032–5035.
- (29) Urry, D. W.; Khaled, M. A.; Rapaka, R. S.; Okamoto, K. *Biochem. Biophys. Res. Commun.* **1977**, *79*, 700–706.
- (30) Asakura, T.; Ashida, J.; Ohgo, K. *Polym. J.* **2003**, *35*, 293–296.
- (31) Hong, M.; Isailovic, D.; McMillan, R. A.; Conticello, V. P. *Biopolymers* **2003**, *70*, 158–168.
- (32) Urry, D. W. *J. Protein Chem.* **1988**, *7*, 1–34.
- (33) Urry, D. W.; Trapane, T. L.; Sugano, H.; Prasad, K. U. *J. Am. Chem. Soc.* **1981**, *103*, 2080–2089.

- (34) Ohgo, K.; Ashida, J.; Kumashiro, K. K.; Asakura, T. *Macromolecules* **2005**, *38*, 6038–6047.
- (35) Reiersen, H.; Clarke, A. R.; Rees, A. R. *J. Mol. Biol.* **1998**, *283*, 255–264.
- (36) Gross, P. C.; Possart, W.; Zeppezauer, M. *Z. Naturforsch., C: J. Biosci.* **2003**, *58*, 873–878.
- (37) Yao, X. L.; Hong, M. *J. Am. Chem. Soc.* **2004**, *126*, 4199–4210.
- (38) Rousseau, R.; Schreiner, E.; Kohlmeyer, A.; Marx, D. *Biophys. J.* **2004**, *86*, 1393–1407.
- (39) Kruckau, A.; Brovchenko, I.; Geiger, A. *Biomacromolecules* **2007**, *8*, 2196–2202.
- (40) Li, B.; Alonso, D. O. V.; Daggett, V. *J. Mol. Biol.* **2001**, *305*, 581–592.
- (41) Case, D. A.; Darden, T. A.; T.E. Cheatham, I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 11*; University of California: San Francisco, 2011.
- (42) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (43) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (44) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (45) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (46) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (47) Shao, J. Y.; Tanner, S. W.; Thompson, N.; Cheatham, T. E. *J. Chem. Theory. Comput.* **2007**, *3*, 2312–2334.
- (48) Deshmukh, S. A.; Sankaranarayanan, S. K. R. S.; Suthar, K.; Mancini, D. C. *J. Phys. Chem. B* **2012**, *116*, 2651–2663.
- (49) Rauscher, S.; Baud, S.; Miao, M.; Keeley, F. W.; Pomes, R. *Structure* **2006**, *14*, 1667–1676.
- (50) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.
- (51) Tamburro, A. M.; Bochicchio, B.; Pepe, A. *Biochemistry* **2003**, *42*, 13347–13362.
- (52) Oleinikova, A.; Brovchenko, I. *Phys. Chem. Chem. Phys.* **2012**, *14*, 5686.
- (53) Vogel, M. *J. Phys. Chem. B* **2009**, *113*, 9386–9392.
- (54) Oleinikova, A.; Brovchenko, I. *J. Phys. Chem. Lett.* **2011**, *2*, 765–769.
- (55) Wernet, P.; Nordlund, D.; Bergmann, U.; Cavalleri, M.; Odelius, M.; Ogasawara, H.; Naslund, L. A.; Hirsch, T. K.; Ojamae, L.; Glatzel, P.; Pettersson, L. G. M.; Nilsson, A. *Science* **2004**, *304*, 995–999.
- (56) Chothia, C. *Nature* **1974**, *248*, 338–339.
- (57) Miller, S.; Janin, J.; Lesk, A. M.; Chothia, C. *J. Mol. Biol.* **1987**, *196*, 641–656.
- (58) Miao, M.; Bellingham, C. M.; Stahl, R. J.; Sitarz, E. E.; Lane, C. J.; Keeley, F. W. *J. Biol. Chem.* **2003**, *278*, 48553–48562.
- (59) Ma, B. Y.; Nussinov, R. *J. Mol. Biol.* **2000**, *296*, 1091–1104.
- (60) Muiznieks, L. D.; Weiss, A. S.; Keeley, F. W. *Biochem. Cell Biol.* **2010**, *88*, 239–250.