

## Nonequilibrium Dynamics of the Helix-Coil Transition in Polyalanine

Maximilian Conradi,<sup>1</sup> Henrik Christiansen,<sup>1,2</sup> Suman Majumder,<sup>3</sup> Fabio Müller,<sup>1</sup> and Wolfhard Janke<sup>1</sup><sup>1</sup>*Institut für Theoretische Physik, Universität Leipzig, IPF 231101, 04081 Leipzig, Germany*<sup>2</sup>*NEC Laboratories Europe GmbH, Kurfürsten-Anlage 36, 69115 Heidelberg, Germany*<sup>3</sup>*Amity Institute of Applied Sciences, Amity University Uttar Pradesh, Noida 201313, India*

(\*Electronic mail: maximilian.conradi@itp.uni-leipzig.de)

(Dated: 9 February 2025)

In this work, the nonequilibrium pathways of the collapse of the helix-forming biopolymer polyalanine are investigated. To this end, the full time evolution of the helix-coil transition is simulated using molecular dynamics simulations. At the start of the transition short  $3_{10}$ -helices form, seemingly leading to the molecule becoming more aspherical midway through the collapse. After the completed collapse, the formation of  $\alpha$ -helices seems to become the prevalent ordering mechanism leading to helical bundles, a structure representative for the equilibrium behavior of longer chains. The dynamics of this transition is explored in terms of the power-law scaling of two associated relaxation times as a function of the chain length.

## I. INTRODUCTION

The shape of a protein can influence its function or lead to a loss of function altogether, which is the cause of various diseases.<sup>1</sup> While machine learning based approaches are able to accurately predict the structure of a protein based on its sequence of amino acids,<sup>2,3</sup> we still do not understand the underlying mechanisms of secondary and tertiary structure formation. Specifically, the dynamics of the nonequilibrium pathways that enable proteins to fold into their native states in time scales of micro- to milliseconds is not well understood.<sup>4,5</sup> In contrast, nonbiological homopolymers like polystyrene take multiple seconds to fold.<sup>6</sup>

The collapse of the protein backbone plays a major role in this process. The impact on the folding originates from backbone-angle preferences and the chain entropy, which influence the native state of the protein.<sup>5,7–10</sup> Based on this hypothesis numerous studies have investigated this process using homopolymers and polypeptides as model for the backbone.<sup>11–28</sup> Only in recent decades have experimental methods like small-angle X-ray scattering, single-molecule fluorescence or dielectric spectroscopy been able to monitor the collapse of a single molecule.<sup>29–40</sup> It is for this reason that most earlier studies used computational methods like Monte Carlo (MC) simulations or molecular dynamics (MD) simulations to analyze this transition. In this regard, MD simulations have been used more extensively not only because of easily available open source frameworks, but also due to the fact that the physical dynamics can be represented in MD simulations more faithfully than in MC simulations.

In general, a polymer undergoes a transition from an extended, random coil to an ordered, compact state, when the solvent condition is changed from good to poor. The first theoretical description of the kinetics of this process has been proposed by de Gennes.<sup>11</sup> In his seminal “sausage model” the collapse is explained by the formation of sausage-like intermediate structures due to interactions between monomers. This intermediate slowly becomes more spherical as the surface energy decreases until it finally reaches a globular state. Later Halperin and Goldbart provided an alternative phe-

nomenological “pearl-necklace” theory.<sup>12</sup> There, the collapse is described as a stepwise sequence of specific events during the collapse. In the first stage, small clusters start to form along the chain. This leads to a sequence of interconnected pearls on the chain, the pearl-necklace intermediate. The pearls then start to slowly absorb monomers of the connecting segments. Eventually, the clusters merge together and form a single globular structure. This theory is supported by numerous studies of collapse transitions and hence has become the commonly accepted picture for this type of process.<sup>13–16,22,24–28</sup> However, a newer study suggests that depending on the solvent viscosity and temperature a combination of both the pearl-necklace and sausage scenario can be observed.<sup>41</sup> While these theories do describe the collapse of a coarse-grained homopolymer, it is unclear whether this translates to biological homopolymers and proteins.

A quantity of significant interest in such studies of the nonequilibrium dynamics is the behavior of the collapse time  $\tau_c$  as a function of the length of the polymer (typically measured by the number of monomers or residues), which usually follows a power-law scaling of the form

$$\tau_c \sim N^z, \quad (1)$$

with  $z$  being the dynamic exponent. Past studies have found different values for the exponent  $z$ , ranging from  $z \approx 1 - 2$ , depending on the specifics of the simulation. Generally, MC simulations<sup>12–16</sup> provide higher values of  $z \approx 2$  compared to MD simulations<sup>17,18,20</sup> where  $z \approx 1$  is found; for an overview, see Ref. 28. This is explained by the fact, that MD simulations are able to incorporate hydrodynamic effects resulting in a faster dynamics.

All these results have been obtained from simulations of coarse-grained homopolymers, which typically represent non-biopolymers. Similar studies considering biopolymers are rare. Only recently, an all-atom MD simulation study of the collapse dynamics of polyglycine, a bio-homopolymer composed of the amino acid glycine as its residue, in explicit water has been reported.<sup>27</sup> There, although the sequence of events during the collapse is similar to that of non-biopolymers, the dynamics is found to be even faster with  $z = 1/2$ . The au-

thors argued that this is caused by the instantaneous hydrogen-bond formation, and speculated that the faster collapse of protein molecules can be attributed to this ultra-fast dynamics of their polypeptide backbone. However, glycine does not feature any secondary structure formation. Hence, the effect of secondary structure formation on the collapse dynamics is still unexplored. In this context, one of the simplest polypeptides that shows a distinct secondary structure formation, i.e., helix formation, is polyalanine.<sup>42</sup> Alanine contains a methyl group  $\text{CH}_3$  attached to the  $-\text{CONH}_2$  group, in contrast to the H-atom in glycine. The presence of this bulky methyl group increases its helical propensity resulting in the formation of a helical conformation at low temperature, thus exhibiting a helix-coil transition as a function of temperature.<sup>43,44</sup>

Thus, with the motivation to understand the effect of secondary structure formation on the overall collapse of a polypeptide, here, we explore the relaxation dynamics of the collapse of the helix forming biopolymer polyalanine by means of all-atom MD simulations in implicit solvent. Our results reveal that due to the tug-of-war between the collapse and secondary structure formation, the sequence of events observed in the nonequilibrium pathway is significantly different from what is observed for a non-biopolymer, unlike the collapse of polyglycine. We show that in addition to the traditional methods of extracting relaxation times from the time dependence of the radius of gyration, the behavior of various shape factors also allows us to extract related relaxation times. To quantify our observations, we also investigate the scaling of these relaxation times with the length of the polypeptide which in all cases shows a much slower dynamics than what is observed for polyglycine, suggesting a slowing down of the overall collapse dynamics due to the simultaneous formation of secondary structure.

The rest of the paper is structured as follows. In Sec. II, we describe the model and details of our simulation method along with the definition of various observables. In Sec. III, we present our results which contains a qualitative picture of the transition, a quantitative description of the same, and analyses of scaling properties of various relaxation times. Finally, in Sec. IV, we provide a short summary of our conclusions as well as an outlook on future research options.

## II. MODEL AND METHODS

Molecules of  $(\text{Ala})_N$  are prepared with a hydrogenated N-terminus ( $-\text{NH}_2$ ) and a C-terminus ( $-\text{COOH}$ ). The all-atom MD simulations are run using the OpenMM package.<sup>45</sup> The Amber14 force field<sup>46</sup> is employed for interactions between the atoms. All simulations are performed in a generalized Born implicit solvent model.<sup>47</sup> For the nonbonded interactions a cutoff radius of 1 nm and a switch-off distance of 0.9 nm are used. Initially, the molecules are equilibrated at  $T = 2000$  K for at least 1 ns, where they take random-coil conformations. Following this, the molecules are quenched to a temperature of  $T_q = 300$  K significantly below the transition temperature  $T_c$ , which is at least 415 K for a short chain of length  $N = 20$  residues.<sup>43,44</sup> MD simulations are per-

formed using the Langevin thermostat with a leap-frog integration scheme. We choose a step size of  $\delta t = 1$  fs and a friction coefficient of  $\gamma = 1$  ps<sup>-1</sup> for the integrator. We have simulated polyalanine molecules with twelve different chain lengths  $N = 25, 50, 75, \dots, 300$  residues. For chain lengths up to  $N = 100$  we use 200 independent initial conformations and for longer chains 100 initial conformations each. Each simulation is run for at least 100 ns and measurements are performed in logarithmically spaced intervals. The results are averaged over all realizations for each considered chain length  $N$ . Results for the relaxation times are obtained using the delete-one Jackknife method.<sup>48-50</sup>

The results of the simulations are analyzed using three main types of observables, which we introduce in the following: (1) Rotationally invariant quantities derived from the gyration tensor, (2) energy contributions based on the force field, and (3) hydrogen bonds and secondary structure. Based on the positions of the individual atoms  $\vec{r}_m$  at a given point in time the components of the gyration tensor  $\mathbf{Q}$  can be calculated as follows:

$$\mathbf{Q} = Q_{ij} = \frac{1}{M} \sum_{m=1}^M (r_m^i - r_{\text{CM}}^i)(r_m^j - r_{\text{CM}}^j), \quad i, j = 1, \dots, d. \quad (2)$$

Here,  $M$  is the total number of atoms in the model,  $r_{\text{CM}}^i$  refers to the  $i$ -th component of the center of mass vector, and  $d = 3$  is the spatial dimension. From this we can derive the squared radius of gyration  $R_g^2$  as

$$R_g^2 = \frac{1}{M} \sum_{m=1}^M (\vec{r}_m - \vec{r}_{\text{CM}})^2 = \sum_{i=1}^d Q_{ii} = \text{Tr } \mathbf{Q}. \quad (3)$$

Three more quantities that can be derived from the gyration tensor are the asphericity  $A$ , the prolateness  $S$ , and the nature of asphericity  $\Sigma$ . Using the average eigenvalue of the gyration tensor

$$\bar{\lambda} = \frac{\text{Tr } \mathbf{Q}}{3} = \frac{1}{3} \sum_{i=1}^3 \lambda_i, \quad (4)$$

we can define the asphericity as<sup>51-53</sup>

$$A = \frac{1}{6} \sum_{i=1}^3 \frac{(\lambda_i - \bar{\lambda})^2}{\bar{\lambda}^2} = \frac{3}{2} \frac{\text{Tr } \hat{\mathbf{Q}}^2}{(\text{Tr } \mathbf{Q})^2}, \quad (5)$$

where  $\hat{\mathbf{Q}} = \mathbf{Q} - \bar{\lambda} \mathbf{I}$  and  $\mathbf{I}$  is the unit matrix. This quantity describes the shape of a polymer based on how spherical it is. It takes a value between  $A = 0$  for a perfectly spherelike conformation, where  $\lambda_i = \bar{\lambda}$ , and  $A = 1$  for a completely straight rodlike conformation with all but one eigenvalue equal to zero. The second quantity is the prolateness that is given by<sup>52,53</sup>

$$S = \frac{\prod_{i=1}^3 (\lambda_i - \bar{\lambda})}{\bar{\lambda}^3} = 27 \frac{\det \hat{\mathbf{Q}}}{(\text{Tr } \mathbf{Q})^3}. \quad (6)$$

For a fully prolate, rodlike conformation one obtains  $S = 2$  since  $\lambda_1 \neq 0$  and  $\lambda_2 = \lambda_3 = 0$ . For an absolutely oblate, disk-like conformation, on the other hand, one has  $\lambda_1 = \lambda_2$  and

$\lambda_3 = 0$  which results in  $S = -1/4$ . This implies  $-1/4 \leq S \leq 2$ . Since  $S = 0$  for a spherelike conformation, in general  $S > 0$  corresponds to a prolate, ellipsoidlike conformation and  $S < 0$  to an oblate one. Finally, one can calculate the nature of asphericity<sup>54,55</sup>

$$\Sigma = \frac{4\prod_{i=1}^3(\lambda_i - \bar{\lambda})}{\left[\frac{2}{3}\sum_{i=1}^3(\lambda_i - \bar{\lambda})^2\right]^{3/2}} = \frac{4\det \hat{Q}}{\left(\frac{2}{3}\text{Tr} \hat{Q}^2\right)^{3/2}}. \quad (7)$$

This quantity gives values of  $\Sigma = -1$  for a disk and  $\Sigma = 1$  for a rigid rod. Hence, it has the boundaries  $-1 \leq \Sigma \leq 1$ . Note that for any given conformation, the latter three shape parameters are related by  $\Sigma = S/(2A^{3/2})$ .

Additionally to the aforementioned geometrical quantities, we also monitor the individual components of the energy as obtained from the Amber force field.<sup>46,56</sup> The energies calculated by the force field can be divided into nonbonded energies, which are calculated for atoms that are three or more bonds apart, and bonded energies. The first group consists of the energy based on Coulomb interactions between two charged particles and the energy derived from the Lennard-Jones interactions for pairs of atoms. In regard to the bonded energies we specifically look at the torsion energy, since it is known that the dihedral angles are related to the secondary structure.

Finally, using the DSSP algorithm, one identifies hydrogen bonds and based on their pattern assigns secondary structures to residues.<sup>57</sup> This way we can determine residues that are part of helical segments or other secondary structure elements. One of the two most common secondary structure elements are  $\alpha$ -helices which are present in most proteins.<sup>58</sup> Its basic structure consists of backbone amide bonds between residues  $i$  and  $i + 4$  marking a turn with a length of approximately 3.6 residues and an offset of 0.54 nm.<sup>59</sup> Additionally, there exist two other types of helical structures:  $3_{10}$ - and  $\pi$ -helices.  $3_{10}$ -helices are helix structures that are based on bonds between residues  $i$  and  $i + 3$  making them narrower than  $\alpha$ -helices. Usually, this type is only found in short sequences of 3 to 4 residues at the ends of  $\alpha$ -helices.<sup>58,60,61</sup> Therefore, they are viewed as a possible intermediate in the formation of  $\alpha$ -helices.<sup>59</sup> Finally,  $\pi$ -helices consist of bonds between residues  $i$  and  $i + 5$ . We did not encounter this helix type as it is relatively rare because the backbone angles are energetically unfavorable compared to  $\alpha$ -helices and the entropic cost for the formation is higher than for the other two helix types.<sup>62</sup>

### III. RESULTS

#### A. Qualitative picture of the helix-coil transition in polyalanine

We first qualitatively investigate the collapse kinetics of a single polyalanine molecule of length  $N = 100$  residues simulated up to 1000 ns. In Fig. 1 snapshots from the trajectory of an individual molecule are displayed. Below each snapshot is the corresponding contact map. There, the distances  $\Delta r_{ij}$  between the  $C^\alpha$ -atoms of residues  $i$  and  $j$  are displayed.

The colors encode the distance, where yellow corresponds to small separations and blue to larger values of  $\Delta r_{ij}$ . We apply a cutoff for distances  $\Delta r_{ij} > 1$  nm, to increase the visibility of contacts. The diagonal represents the contacts  $\Delta r_{ii} = 0$  nm and is hence colored in yellow. In the first image one can see that before the quench the molecule is in an extended, random-coil state without helical residues. This is also visible in the corresponding contact map where almost no yellow regions aside from the diagonal exist. Following the quench the molecule still remains in an extended conformation for some time which can be seen in the second frame at  $t = 0.003$  ns. Nonetheless, first helical segments have formed along the chain as visible in the snapshot and in the contact map where they appear as yellow signals close to the diagonal. In general, the molecule has slightly contracted in some areas. At  $t = 0.15$  ns a short straight  $\alpha$ -helix with a length of two turns is present in the molecule. In the contact map this structure is indicated by the broader yellowish region of close contacts around the diagonal approximately between residues 50 and 60. On both sides of this area, small yellow patterns of close contacts appear on the diagonal, signaling the short  $3_{10}$ -helices that have formed. Off the diagonal there are now also some contacts between residues farther apart. Overall, the molecule has straightened in comparison to the previous frames as can be seen from the snapshot. At  $t = 1$  ns the molecule takes a much more compact conformation where the segments on both parts of the helix are visibly stronger contracted. This is also reflected in the contact map by yellow regions off the diagonal that signal contacts with residues that are significantly farther away. In addition, the  $\alpha$ -helix has extended to more than three full turns and is visible as a broad yellow region on the diagonal.  $3_{10}$ -helices are also still present in the parts of the molecule on both sides of the  $\alpha$ -helix. In the following, at  $t = 10$  ns, the residues in the molecule have rearranged to form a more compact globular structure in order to decrease the surface exposed to the solvent. This is reflected in the contact map by close contacts even between residues that are very far apart such as  $i = 10 - 20$  and  $j = 90 - 100$ . At this point, we also see in the snapshot that a number of short  $3_{10}$ -helices are present throughout the entire molecule. The longer  $\alpha$ -helix from the previous frame can still be seen in the bottom of the frame, although it is now classified as a  $3_{10}$ -helix. Finally, in the last frame at  $t = 1000$  ns we see that the molecule has changed from a globular structure to a bundle of helices. The basis for this structure are the four  $\alpha$ -helices that are now aligned in the middle, each consisting of three to four turns. In the contact map this shows as broad yellow stripes of close contacts that have formed alongside the diagonal. Aside from these regions next to the diagonal, we also observe yellowish patterns of close contacts between residues further apart. In contrast to previous frames, these brighter regions now appear in regular patterns of lines parallel and perpendicular to the diagonal. They belong to the contacts between two helices with perpendicular lines corresponding to antiparallel helices and parallel lines to parallel helices. This regular layout shows that the molecule is highly organized with the helices aligned into a bundle. The observed pattern already resembles contact maps for equilibrium conformations found in previous studies.<sup>63,64</sup>

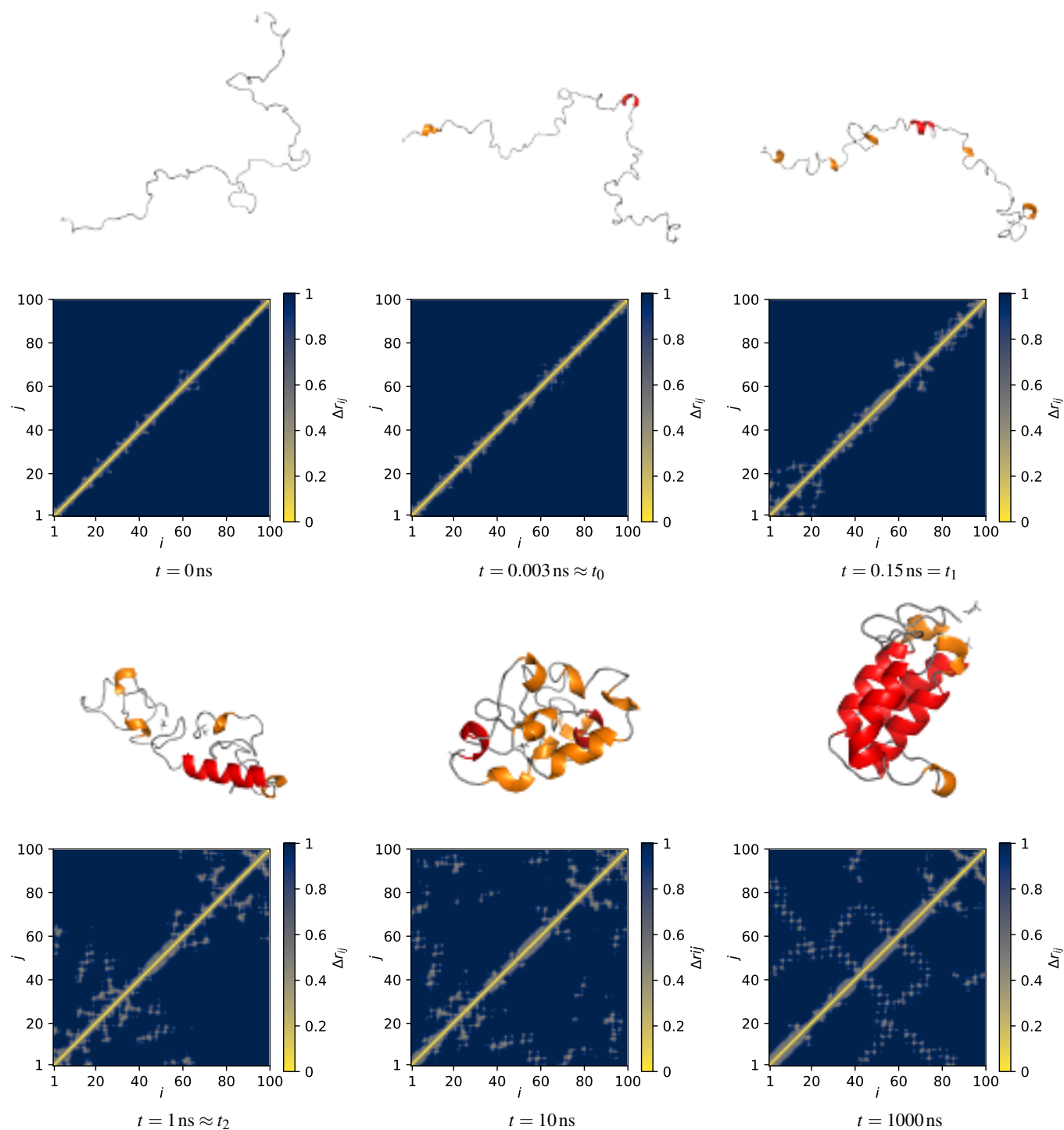


FIG. 1. Snapshots during the transition of a single polyaniline molecule with length  $N = 100$ .  $\alpha$ -helical residues are colored red,  $3_{10}$ -helical residues are colored orange. Below each snapshot is the corresponding contact map. The distances  $\Delta r_{ij}$  between residues  $i$  and  $j$  are given in nm. Distances above  $\Delta r_{ij} > 1$  nm are set to 1 nm. Colors are assigned for each distance  $\Delta r_{ij}$  based on the color bar on the right side of the map. The times  $t_0, t_1, t_2$  mark specific events in the trajectory discussed in the following section.

The evolution of the helices is even better displayed in Fig. 2. Here, each residue is matched with the type of helix it is part of at any point during the helix-coil transition. Trajectory 1 matches the conformations shown in Fig. 1. In this frame multiple vertical orange lines are visible. These represent  $3_{10}$ -

helices that start to form early in the transition as shown in the snapshots. These structures stay in place once they are formed resulting in line patterns. It is notable that all  $3_{10}$ -helices remain short, as visible in the frames of Fig. 1. The  $\alpha$ -helix in the middle of the conformation forms early in the

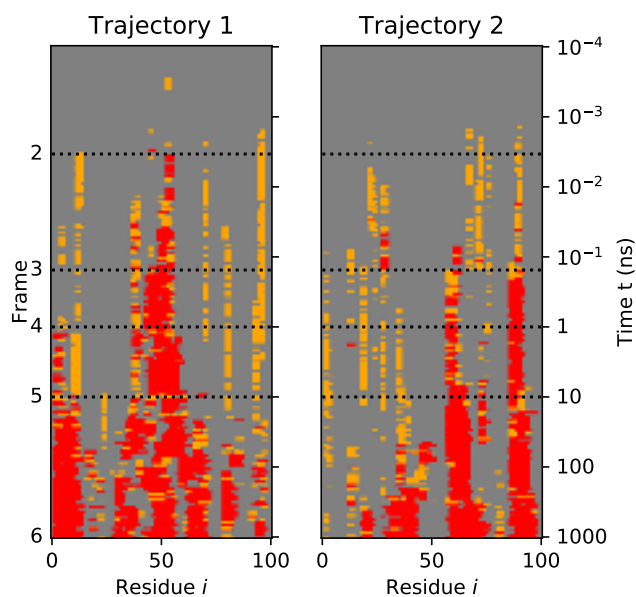


FIG. 2. Time evolution of helices in two selected conformations. The left side shows the time evolution displayed in Fig. 1, the right side another randomly chosen time evolution.  $\alpha$ -helical residues are colored red,  $3_{10}$ -helical residues are colored orange. The dotted lines mark the times of the snapshots in Fig. 1. Frame 1 is not shown.

transition in a spot where previously a  $3_{10}$ -helix existed. In the following this  $\alpha$ -helix expands in both directions by merging with adjacent  $3_{10}$ -helices and eventually bends into multiple helices as can be seen close to  $t = 10$  and  $t = 100$  ns. In the final part at  $t = 1000$  ns we can see four  $\alpha$ -helices, which are all of similar length. Notably, three of them are only separated by a few residues and have seemingly emerged from the same  $\alpha$ -helix that formed early in this trajectory. Aside from these a significant part of the molecule has no  $\alpha$ -helix structures, although these regions exhibited helices at earlier times. For comparison the right frame in Fig. 2 displays another randomly chosen trajectory. Here, it also seems like the helices persist in the spot in which they were formed and that almost all the  $\alpha$ -helices originate from  $3_{10}$ -helices. The trajectories in Fig. 2, however, also show some limitations of this analysis with the DSSP algorithm where the classification of individual residues can oscillate between  $3_{10}$ -helix and  $\alpha$ -helix.

In general, we observe that already early on helical segments are forming throughout the molecule, with the majority being  $3_{10}$ -helices. The molecule then seems to collapse in some regions while simultaneously straightening around the helical segments. Subsequently, the molecule collapses into a single globular state, which already contains short  $\alpha$ - and  $3_{10}$ -helices. These  $\alpha$ -helices appear to form in spots that previously showed  $3_{10}$ -helices supporting observations that  $3_{10}$ -helices can act as an intermediate in the formation of  $\alpha$ -helices.<sup>59</sup> Following this, the  $\alpha$ -helices begin to extend which leads to a transition into a bundle of helices. This seems to deviate from the traditional pearl-necklace picture which is likely due to the formation of short helices early in the col-

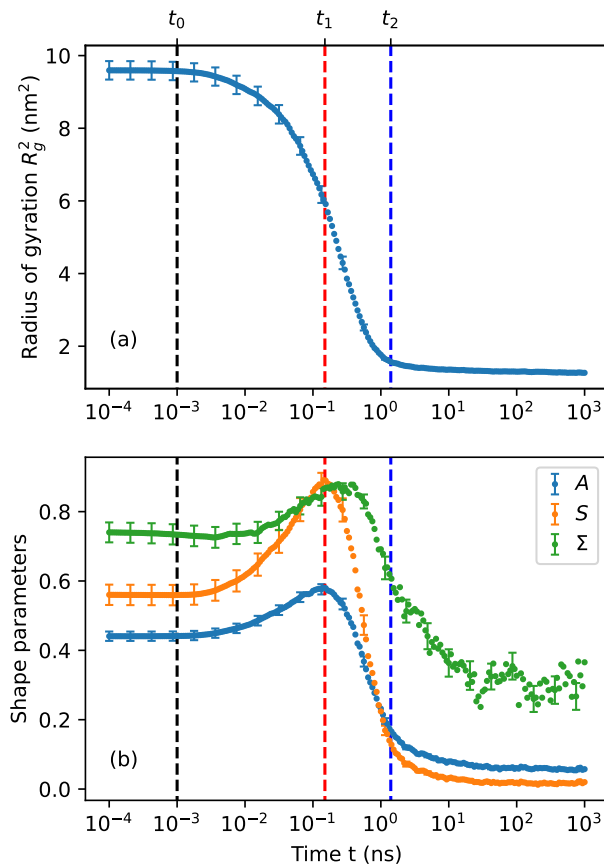


FIG. 3. Time evolution of (a) the squared radius of gyration  $R_g^2$  and (b) asphericity  $A$ , prolateness  $S$ , and nature of asphericity  $\Sigma$  for molecules of chain length  $N = 100$ . The dashed vertical lines at  $t_0 = 0.001$  ns,  $t_1 = 0.15$  ns, and  $t_2 = 1.4$  ns mark notable events in the collapse. All results are averaged over 200 simulations with independent time trajectories.

lapse and their specific bonding pattern. While these exemplary trajectories provide a good illustration of the path that the molecule follows towards equilibrium, they are not necessarily statistically representative and are thus backed up by a more quantitative analysis in the next section.

## B. Quantitative picture of the helix-coil transition

Next, we consider the four shape parameters defined in Sec. II. For this we performed additional simulations with independent initial conformations for a total of 200 independent time trajectories. From the average of their time evolutions displayed in Fig. 3 we identify three characteristic times  $t_0$ ,  $t_1$ , and  $t_2$ . The times  $t_0 = 0.001$  ns and  $t_2 = 1.4$  ns are derived from the time dependence of the squared radius of gyration  $R_g^2(t)$  in (a), marking the beginning and end of the collapse,

respectively. The time  $t_1 = 0.15$  ns marks the maximum in asphericity  $A$  and prolateness  $S$  which is visible in (b). Looking at the time evolution of the squared radius of gyration we see that after the quench, the system needs a short amount of time to become unstable to fluctuations, i.e., to break the already-developed equilibrium correlations in the extended state at high  $T$ . This gets reflected in the behavior of  $R_g^2$  which initially is almost flat until  $t_0$ . From there it starts to decrease until the time  $t_2$  is reached, where it stabilizes at a low value corresponding to the compact, globular state observed in Fig. 1.

Comparing this to the other three shape parameters, we can once again see that initially after the quench all three quantities remain at a constant level taking on values that correspond to a random coil.<sup>52</sup> Approximately at time  $t_0$ , when the collapse begins, both asphericity and prolateness start to increase until they reach a maximum at  $t_1$ . From this point on, both observables show a rapid decrease until they reach stable values at approximately time  $t_2$  once the molecule has fully collapsed. At this point, both quantities have decreased to values close to zero indicating that the shape of the molecule has become almost spherical. In contrast, the nature of asphericity  $\Sigma$  shows a slightly different time evolution. Similar to asphericity and prolateness, it initially remains at a stable value indicating no changes. This region of constant value continues past the beginning of the collapse at  $t_0$  until approximately  $t \approx 10^{-2}$  ns. Following this,  $\Sigma$  starts to increase, albeit slower than the other two quantities. Consequently, it also reaches its maximum, which indicates a rodlike shape, slightly later. Similar to the asphericity and prolateness it decreases afterwards, before fluctuating around a somewhat stable value of  $\Sigma \approx 0.3$  that indicates a slightly rodlike conformation. Strikingly, the maximum in the shape parameters happens in the middle of the collapse. This shows that despite the molecule collapsing, a more rodlike structure is able to form. A similar pattern of the time evolution for these shape parameters has also been observed for a non-biopolymer in a recent study.<sup>41</sup> There it was ascribed to a collapse separated into an initial pearl-necklace stage that results in a single sausage-like globule at the maximum of both asphericity and prolateness. This is then followed by a rearrangement of the residues into a spherical shape resulting in a decrease in  $A$  and  $S$  and, eventually, reaching a stable value. While we observe the same pattern in the time evolution of asphericity and prolateness, the structural properties reflected in the shape parameters are quite different. In fact, it is likely that in our case helical structures are related to this maximum in the three shape parameters. When comparing with the evolution of helices in Fig. 2 we can see that already at  $t_0$  helices start to form in the molecule and at  $t_1$  there are multiple short  $3_{10}$ -helices and one long  $\alpha$ -helix present in the molecule.

The early formation of helices becomes even more apparent when looking at the time evolution of the average number and average length of helices in Fig. 4. Here, we can see in (a) that already with the beginning of the collapse at  $t_0$ ,  $3_{10}$ -helices start to form as the number of helices increases. Shortly after the beginning of the collapse also  $\alpha$ -helices start to form. While the average number of  $\alpha$ -helices is growing through-

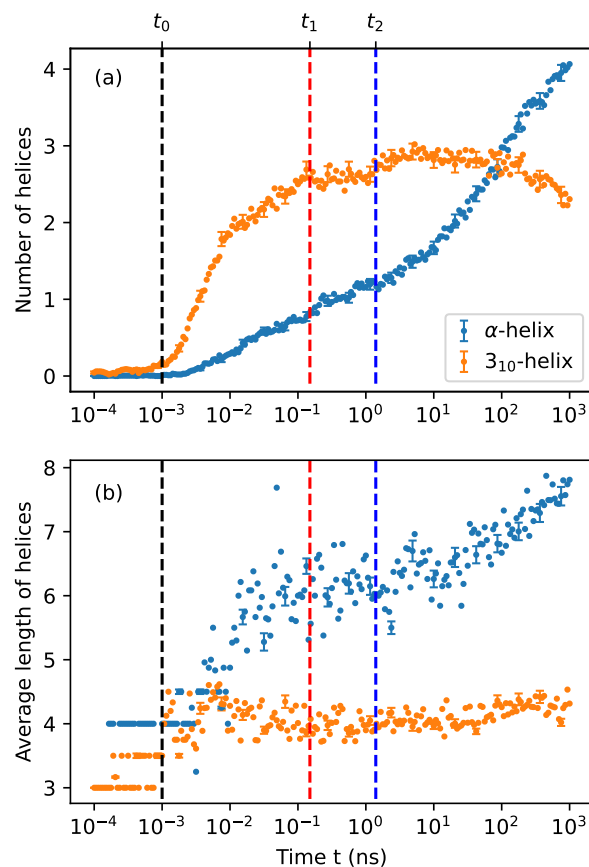


FIG. 4. Time evolution of (a) the average number of helices and (b) the average length of  $\alpha$ - and  $3_{10}$ -helices for molecules of chain length  $N = 100$ . All results are averaged over 200 simulations with independent initial conformations. Legend for both frames is displayed in (a).

out the entire transition, the number of  $3_{10}$ -helices starts to plateau at  $t_1$ , the maximum in asphericity and prolateness. This suggests that the formation of  $3_{10}$ -helices is related to the increase in asphericity and prolateness. After this point the average number of  $3_{10}$ -helices remains at about three helices per molecule. Only on longer time scales past the end of the collapse the number of  $3_{10}$ -helices is starting to slightly decrease again.

The average length of the helices in Fig. 4 (b) shows a slightly different picture. Initially, the few existing  $3_{10}$ - and  $\alpha$ -helices show average lengths of approximately 3 and 4 residues, respectively, which is in both cases the minimum number of residues needed to form one turn. Close to  $t_0$  the average length of the  $3_{10}$ -helices increases to approximately 4 residues. However, already before this point some  $3_{10}$ -helices were longer than 3 residues as indicated by the data points at an average length of 3.5 residues. The length of these helices then does not change throughout the remaining parts of

the transition in accordance with previous observations that this helix type is only forming short helices.<sup>58,60,61</sup> The average length of  $\alpha$ -helices remains at the initial value longer and only increases when the number of  $\alpha$ -helices in Fig. 4 (a) starts to increase. It then appears to increase until approximately  $t_1$  where the average length fluctuates around a value of approximately 6 residues per helix. The large fluctuations in the average length could be caused by the inaccuracies in the classification of DSSP that was already observed in Fig. 2. After the end of the collapse the average length of  $\alpha$ -helices starts to increase again around the time at which also the number of  $\alpha$ -helices is showing a signal. This shows that, while able to form throughout the entire collapse,  $\alpha$ -helices are not able to extend between  $t_1$  and  $t_2$  where the molecule collapses to a more compact, globular state. Only at later times do the helices continue to extend while the molecule is transitioning into equilibrium states with longer straight  $\alpha$ -helices such as the helical bundle shown in the last frame of Fig. 1. At the end of our simulations  $\alpha$ -helices have reached an average length of 8 residues which is equivalent to two full turns. The average length at this time is close to what we observe in our own equilibrium simulations using precisely the same model setup, but significantly lower than what has been observed in vacuum where such helices are more than double in size.<sup>42,63,64</sup> The authors of these studies, however, also find a lower number of helices for the respective chain length. The reason for these longer helices, is that in the absence of a solvent extended conformations with long helices are stabilized.

This pattern of helix formation is also displayed in the number of intramolecular hydrogen bonds in Fig. 5 (a). We can see that the number of hydrogen bonds is strongly correlated to the formation of helices, as one would expect. It starts to increase approximately at  $t_0$  when  $3_{10}$ -helices start to form and continues even after the end of the collapse, when only  $\alpha$ -helix-formation is happening. Also in (b) the different contributions of the energy provided by the Amber14 force field<sup>46</sup> display signals of helix formation. The nonbonded energies start to decrease already early on, significantly before  $t_0$ . A first signal can be seen around  $t \approx 10^{-2}$  ns, the time at which  $\alpha$ -helix formation is setting in. Further signals are observed at  $t_1$  and around  $t_2$  aligning with the peak in asphericity and prolateness and the end of the collapse. These signals in the time evolution could be related to the formation of helices and the difference in nonbonded energy between  $3_{10}$ - and  $\alpha$ -helices.<sup>65</sup> The torsion energy on the other hand shows a simpler profile. Again, we observe that it decreases early on continuing past  $t_0$  and reaches a constant level around the time when  $\alpha$ -helices begin to form, suggesting that the decrease in torsion energy could be related to the delayed formation of  $\alpha$ -helices. At late times no significant changes in the torsion energy are visible. While harmonic bond energy and angle energy were also determined, they are not displayed as their behavior shows no specific signals for the collapse.

In general, we are able to distinguish two stages of the transition: The collapse and the subsequent folding to a helical bundle. During the first stage the molecule transitions from an extended random coil state to a compact state. The initial part of the collapse is characterized by the formation of short  $3_{10}$ -

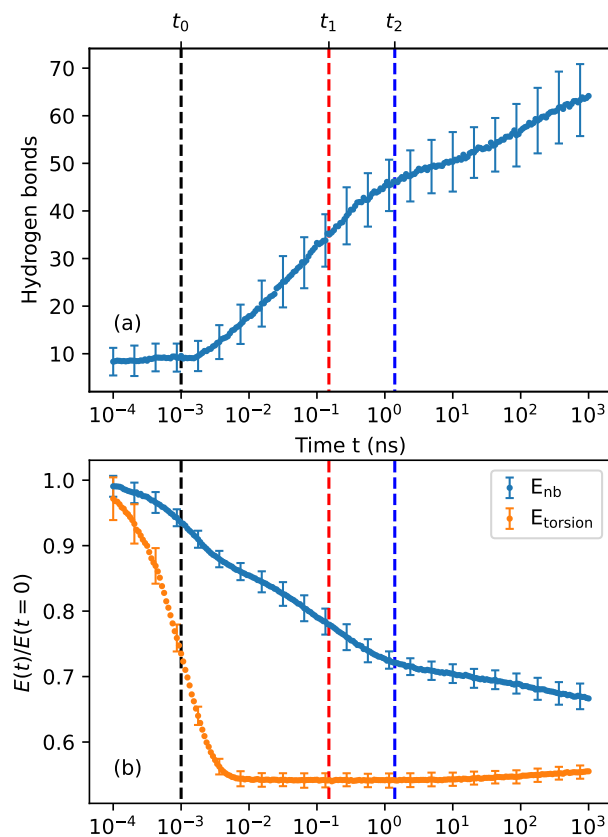


FIG. 5. (a) Time evolution of the number of intramolecular hydrogen bonds. (b) Time evolution of the nonbonded energies  $E_{nb}$  and torsion energy  $E_{torsion}$  normalized with their respective values at  $t = 0$  ns. All results are for  $N = 100$  and averaged over 200 simulations with independent initial conformations.

helices and a corresponding increase in asphericity and prolateness indicating a more rodlike structure. This is followed by a decrease in  $R_g^2$  as well as the other shape factors to values corresponding roughly to a spherical state. Already during this process we observe formation of  $3_{10}$ - and  $\alpha$ -helices. The second stage of the process is characterized by a slight decrease in the number of  $3_{10}$ -helices and the ongoing formation of new  $\alpha$ -helices and extension of existing  $\alpha$ -helices which eventually organize in equilibrium states such as helical bundles. Even though the shape parameters do not change during this process, the torsion energy appears to slightly increase. While a previous study<sup>66</sup> found a separation between the formation of  $\alpha$ -helices and the collapse, we do not observe such a strict separation. A possible explanation could be that we use chains that are significantly longer where equilibrium states would likely contain more than just one  $\alpha$ -helix.<sup>64</sup>

### C. Scaling laws governing the collapse of polyaniline

Finally, by extracting relaxation times, we dive into a more quantitative analysis. To this end, we have added further simulations for chains of length  $N = 25, 75 \dots 300$  with 200 realizations for  $N \leq 100$  and 100 for  $N > 100$ . All simulations are run up to at least  $t = 100$  ns. In a first attempt to model the decay of  $R_g^2$  from its random coil value at  $t = 0$  ( $R_g^2 \propto N^{1.2}$  where  $1.2 = 2v_{SAW}$  with the Flory approximation  $v_{SAW} = 3/5$ ) to the close-to-asymptotic compact state ( $R_g^2 \propto N^{2/3}$ ), we employed the template from studies of generic homopolymers<sup>25,26</sup> and polyglycine<sup>27</sup>, namely a stretched exponential (or Kohlrausch) ansatz

$$R_g^2(t) = b_0 + be^{-(t/\tau_c)^\beta}. \quad (8)$$

This works fairly well also here and since the exponent  $\beta$  turned out to be close to unity ( $\beta \approx 0.7$ ) even a simple exponential ansatz does qualitatively represent the decay, at least in its initial part. For the decay of the asphericity  $A$  for times beyond its peak value, however, this is no longer true. Here, a much more satisfactory fit of the data for  $t > t(A_{\max})$  is achieved with the 2-exponential form

$$A(t) = ae^{-t/\tau_1} + be^{-t/\tau_2} + c, \quad (9)$$

where  $\tau_1$  describes the initial decay (the overall ‘‘collapse’’) and  $\tau_2 \gg \tau_1$  the final, much slower relaxation process within the already formed compact state. Eventually, we used the same ansatz (9) for  $R_g^2$  as well since also here it describes the data better than the stretched exponential (8) even though the difference is not very pronounced (we obtain  $\tau_c \approx \tau_1$  within a few percent). This good representation can be seen from the fits in Fig. 6 (a) for different chain lengths where the fits align with the data within error bars. Even better results were obtained with this ansatz for the asphericity with  $t > t(A_{\max})$  across different chain lengths shown in (b).

Since the raw data of  $R_g^2(t)$  and  $A(t)$  are correlated in time we repeated all fits for delete-one Jackknife subsets<sup>48–50</sup> and estimated the statistical errors on the fit parameters from the Jackknife variance. The resulting  $\tau_1$  and  $\tau_2$  are displayed in Fig. 7. One can see in (a) that for both  $R_g^2(t)$  and  $A(t)$  the smaller relaxation times  $\tau_1$  of the fast decay nicely follow the power law in  $\tau_1 \propto N^z$  represented by the straight lines. From the scaling of the relaxation time  $\tau_1$  with  $N$  of  $R_g^2$  we obtain  $z = 1.32(3)$ , using the fit over all data points (with  $\chi^2/\text{d.o.f.} = 2.3$ ), and for  $A$  we find the compatible estimate  $z = 1.38(4)$  (with  $\chi^2/\text{d.o.f.} = 1.2$ ). These are the displayed fits. When changing the fit range, the estimates for  $z$  vary only slightly as shown in (b). We hence conclude that the initial relaxation (the overall ‘‘collapse’’) is characterized by  $z \approx 1.35$ . The estimates for  $\tau_2$  of the slower decay are less reliable since this final relaxation within the already formed compact structure probes the rightmost regime of  $R_g^2(t)$  and  $A(t)$  where the signal-to-noise ratio becomes rather poor. The estimated  $\tau_2$  seem to also be compatible with a power law  $\tau_2 \propto N^{1.35}$ . The result for the scaling exponent  $z$  aligns with values typically found for homopolymers in MD simulations. It is, however,

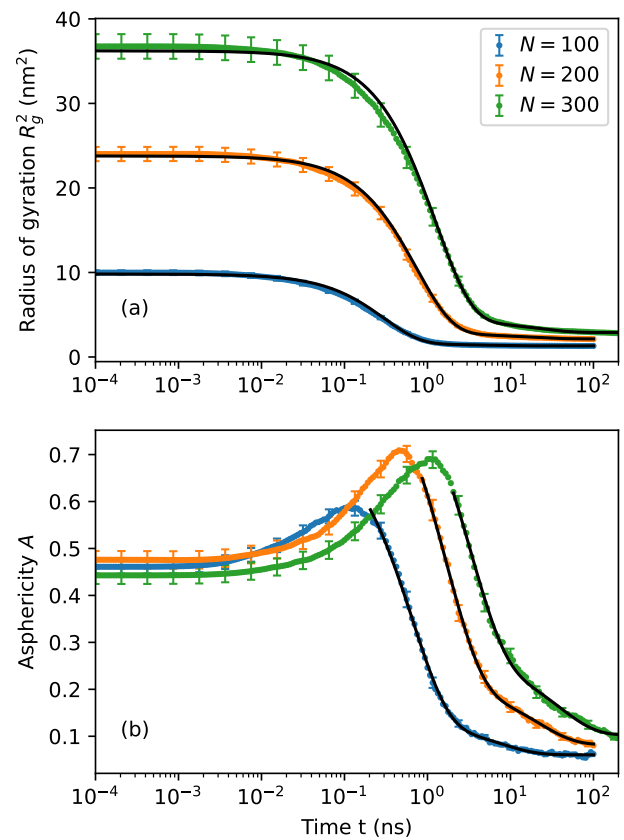


FIG. 6. Time evolution of (a) the squared radius of gyration and (b) the asphericity for different values of  $N$  as indicated. All results are averaged over the 200 (or 100 for  $N \geq 125$ ) simulations with independent initial conformations. Legend for both frames is displayed in (a). The black solid lines represent fits with the 2-exponential ansatz (9).

at odds with an earlier study on the collapse of polyglycine<sup>27</sup> that found a very fast decay and conjectured that intrachain hydrogen bonds might be the reason why biological polymers are able to fold so fast. A possible explanation for this difference is the formation of short helices early in the collapse in our study. By locally stabilizing the polymer they seem to be related to the formation of a more rodlike structure during the collapse, which could be responsible for the slowing down. However, it would be interesting to investigate whether the use of explicit water could influence this behavior which is our plan for future work.

## IV. CONCLUSION

We have investigated the helix-coil transition in polyaniline using secondary structure analysis and shape factors derived from the gyration tensor. Based on the study of the col-



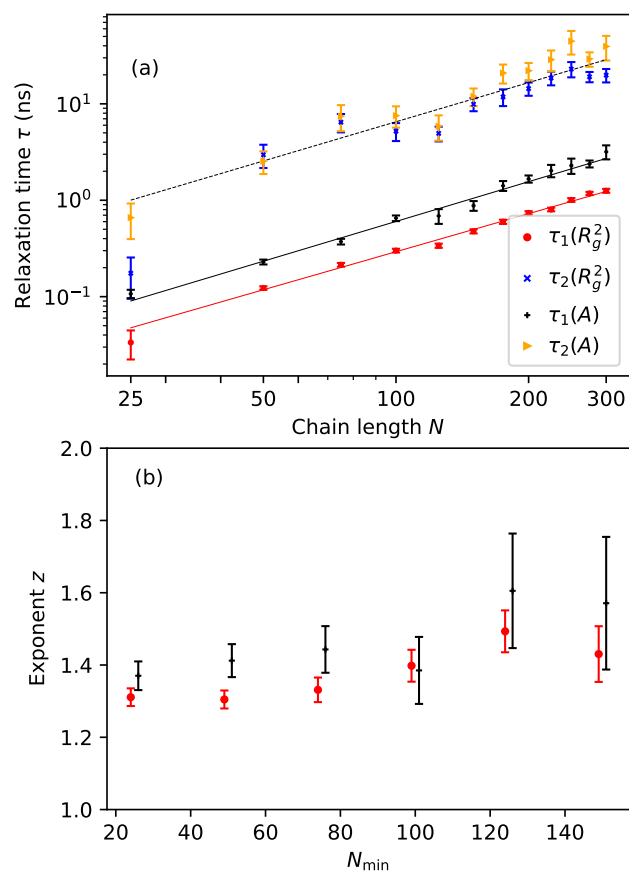


FIG. 7. (a) Relaxation times  $\tau_1$  and  $\tau_2$  obtained from ansatz (9) for different chain lengths  $N$ . The fitted lines correspond to  $\tau \propto N^z$ . The straight dotted line  $\propto N^{1.35}$  through the data for  $\tau_2$  is only a guide to the eye. All results are averaged over the 200 (or 100 for  $N \geq 125$ ) simulations with independent initial conformations. (b) Dependence of the scaling exponents  $z$  on the minimum of the fitting range  $N_{\min}$  (for better visibility, the data for  $A$  are displaced to  $N_{\min} + 1$ ). Legend for both frames is displayed in (a).

lapse in molecules with a chain length of  $N = 100$ , we observed that the transition in general can be split into two major stages: collapse and the subsequent transition into an equilibrium state with straight  $\alpha$ -helices. Already during the first stage helix formation is encountered, which may be responsible for the observed straightening of the molecule halfway through the collapse. The second stage of the transition is then characterized by the formation of straight  $\alpha$ -helices leading eventually to structures like helical bundles.

In the third part of our analysis, we applied a 2-exponential ansatz to the time evolutions of radius of gyration and the asphericity. For both observables we find that the smaller relaxation time  $\tau_1$  scales with an exponent of  $z \approx 1.35$ , similar to exponents obtained in previous studies for homopolymers. The larger relaxation times  $\tau_2$  also appear to follow a power law with exponent  $z \approx 1.35$ . It would be interesting to investigate the scaling of this relaxation time with longer simulations to better capture this relaxation on larger time scales.

Our results suggest that the collapse of polyaniline is influenced by the formation of  $3_{10}$ -helices and  $\alpha$ -helices. The formation of these structures during the collapse could slow down this process compared to other polymers. An interesting extension of this study would be inclusion of explicit solvent which may allow to better disentangle the different factors influencing the two relaxation times.

## ACKNOWLEDGMENTS

This project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 469 830 597 under project ID JA 483/35–1. S.M. thanks the Science and Engineering Research Board (SERB), Govt. of India for a Ramanujan Fellowship (file no. RJF/2021/000044). All simulations were performed on the GPU cluster of the Universitätsrechenzentrum (URZ) at Universität Leipzig.

## AUTHOR DECLARATIONS

### Conflict of interest

The authors have no conflicts to disclose.

## DATA AVAILABILITY STATEMENT

Data available on request from the authors.

## REFERENCES

- <sup>1</sup>C. M. Dobson, “Protein folding and misfolding,” *Nature* **426**, 884–890 (2003).
- <sup>2</sup>J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, “Highly accurate protein structure prediction with alphafold,” *Nature* **596**, 583–589 (2021).
- <sup>3</sup>M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch, R. D. Schaeffer, C. Millán, H. Park, C. Adams, C. R. Glassman, A. DeGiovanni, J. H. Pereira, A. V. Rodrigues, A. A. van Dijk, A. C. Ebrecht, D. J. Opperman, T. Sagmeister, C. Buhlinger, T. Pavkov-Keller, M. K. Rathinaswamy, U. Dalwadi, C. K. Yip, J. E. Burke, K. C. Garcia, N. V. Grishin, P. D. Adams, R. J. Read, and D. Baker, “Accurate prediction of protein structures and interactions using a three-track neural network,” *Science* **373**, 871–876 (2021).
- <sup>4</sup>K. A. Dill, S. B. Ozkan, M. S. Shell, and T. R. Weikl, “The protein folding problem,” *Annu. Rev. Biophys.* **37**, 289–316 (2008).
- <sup>5</sup>K. A. Dill and J. L. MacCallum, “The protein-folding problem, 50 years on,” *Science* **338**, 1042–1046 (2012).
- <sup>6</sup>B. Chu, Q. Ying, and A. Y. Grosberg, “Two-stage kinetics of single-chain collapse. Polystyrene in cyclohexane,” *Macromolecules* **28**, 180–189 (1995).
- <sup>7</sup>G. Reddy and D. Thirumalai, “Collapse precedes folding in denaturant-dependent assembly of ubiquitin,” *J. Phys. Chem. B* **121**, 995–1009 (2017).

- <sup>8</sup>G. Haran, "How, when and why proteins collapse: The relation to folding," *Curr. Opin. Struct. Biol.* **22**, 14–20 (2012).
- <sup>9</sup>M. Sadqi, L. J. Lapidus, and V. Muñoz, "How fast is protein hydrophobic collapse?" *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12117–12122 (2003).
- <sup>10</sup>C. J. Camacho and D. Thirumalai, "Kinetics and thermodynamics of folding in model proteins." *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6369–6372 (1993).
- <sup>11</sup>P. De Gennes, "Kinetics of collapse for a flexible coil," *J. Physique Lett.* **46**, 639–642 (1985).
- <sup>12</sup>A. Halperin and P. M. Goldbart, "Early stages of homopolymer collapse," *Phys. Rev. E* **61**, 565–573 (2000).
- <sup>13</sup>A. Byrne, P. Kiernan, D. Green, and K. A. Dawson, "Kinetics of homopolymer collapse," *J. Chem. Phys.* **102**, 573–577 (1995).
- <sup>14</sup>E. G. Timoshenko, Y. A. Kuznetsov, and K. A. Dawson, "Kinetics at the collapse transition Gaussian self-consistent approach," *J. Chem. Phys.* **102**, 1816–1823 (1995).
- <sup>15</sup>Y. A. Kuznetsov, E. G. Timoshenko, and K. A. Dawson, "Kinetics at the collapse transition of homopolymers and random copolymers," *J. Chem. Phys.* **103**, 4807–4818 (1995).
- <sup>16</sup>Y. A. Kuznetsov, E. G. Timoshenko, and K. A. Dawson, "Kinetic laws at the collapse transition of a homopolymer," *J. Chem. Phys.* **104**, 3338–3347 (1996).
- <sup>17</sup>C. F. Abrams, N.-K. Lee, and S. P. Obukhov, "Collapse dynamics of a polymer chain: Theory and simulation," *Europhys. Lett.* **59**, 391–397 (2002).
- <sup>18</sup>N. Kikuchi, J. F. Ryder, C. M. Pooley, and J. M. Yeomans, "Kinetics of the polymer collapse transition: The role of hydrodynamics," *Phys. Rev. E* **71**, 061804 (2005).
- <sup>19</sup>J. Xu, Z. Zhu, S. Luo, C. Wu, and S. Liu, "First observation of two-stage collapsing kinetics of a single synthetic polymer chain," *Phys. Rev. Lett.* **96**, 027802 (2006).
- <sup>20</sup>J. Guo, H. Liang, and Z.-G. Wang, "Coil-to-globule transition by dissipative particle dynamics simulation," *J. Chem. Phys.* **134**, 244904 (2011).
- <sup>21</sup>J. B. Udgaonkar, "Polypeptide chain collapse and protein folding," *Arch. Biochem. Biophys.* **531**, 24–33 (2013).
- <sup>22</sup>S. Majumder and W. Janke, "Cluster coarsening during polymer collapse: Finite-size scaling analysis," *Europhys. Lett.* **110**, 58001 (2015).
- <sup>23</sup>S. Majumder and W. Janke, "Evidence of aging and dynamic scaling in the collapse of a polymer," *Phys. Rev. E* **93**, 032506 (2016).
- <sup>24</sup>S. Majumder and W. Janke, "Aging and related scaling during the collapse of a polymer," *J. Phys.: Conf. Ser.* **750**, 012020 (2016).
- <sup>25</sup>S. Majumder, J. Zierenberg, and W. Janke, "Kinetics of polymer collapse: Effect of temperature on cluster growth and aging," *Soft Matter* **13**, 1276–1290 (2017).
- <sup>26</sup>H. Christiansen, S. Majumder, and W. Janke, "Coarsening and aging of lattice polymers: Influence of bond fluctuations," *J. Chem. Phys.* **147**, 094902 (2017).
- <sup>27</sup>S. Majumder, U. H. E. Hansmann, and W. Janke, "Pearl-necklace-like local ordering drives polypeptide collapse," *Macromolecules* **52**, 5491–5498 (2019).
- <sup>28</sup>S. Majumder, H. Christiansen, and W. Janke, "Understanding nonequilibrium scaling laws governing collapse of a polymer," *Eur. Phys. J. B* **93**, 142 (2020).
- <sup>29</sup>W. A. Eaton, V. Muñoz, S. J. Hagen, G. S. Jas, L. J. Lapidus, E. R. Henry, and J. Hofrichter, "Fast kinetics and mechanisms in protein folding," *Annu. Rev. Biophys. Biomol. Struct.* **29**, 327–359 (2000).
- <sup>30</sup>S. J. Hagen and W. A. Eaton, "Two-state expansion and collapse of a polypeptide," *J. Mol. Biol.* **301**, 1019–1027 (2000).
- <sup>31</sup>B. Chu and B. S. Hsiao, "Small-angle X-ray scattering of polymers," *Chem. Rev.* **101**, 1727–1762 (2001).
- <sup>32</sup>S. Doniach, "Changes in biomolecular conformation seen by small angle X-ray scattering," *Chem. Rev.* **101**, 1763–1778 (2001).
- <sup>33</sup>J. Kubelka, J. Hofrichter, and W. A. Eaton, "The protein folding 'speed limit'," *Curr. Opin. Struct. Biol.* **14**, 76–88 (2004).
- <sup>34</sup>E. Haas, "The study of protein folding and dynamics by determination of intramolecular distance distributions and their fluctuations using ensemble and single-molecule FRET measurements," *ChemPhysChem* **6**, 858–870 (2005).
- <sup>35</sup>H. Roder, K. Maki, and H. Cheng, "Early events in protein folding explored by rapid mixing methods," *Chem. Rev.* **106**, 1836–1861 (2006).
- <sup>36</sup>M. Arai, E. Kondrashkina, C. Kayatekin, C. R. Matthews, M. Iwakura, and O. Bilsel, "Microsecond hydrophobic collapse in the folding of escherichia coli dihydrofolate reductase, an  $\alpha/\beta$ -type protein," *J. Mol. Biol.* **368**, 219–229 (2007).
- <sup>37</sup>T. Konuma, T. Kimura, S. Matsumoto, Y. Goto, T. Fujisawa, A. R. Fersht, and S. Takahashi, "Time-resolved small-angle X-ray scattering study of the folding dynamics of barnase," *J. Mol. Biol.* **405**, 1284–1294 (2011).
- <sup>38</sup>S. V. Kathuria, L. Guo, R. Graceffa, R. Barrea, R. P. Nobrega, C. R. Matthews, T. C. Irving, and O. Bilsel, "Minireview: Structural insights into early folding events using continuous-flow time-resolved small-angle X-ray scattering," *Biopolymers* **95**, 550–558 (2011).
- <sup>39</sup>G. Haran, "How, when and why proteins collapse: the relation to folding," *Curr. Opin. Struct. Biol.* **22**, 14–20 (2012).
- <sup>40</sup>P. R. Banerjee and A. A. Deniz, "Shedding light on protein folding landscapes by single-molecule fluorescence," *Chem. Soc. Rev.* **43**, 1172–1188 (2014).
- <sup>41</sup>S. Majumder, H. Christiansen, and W. Janke, "Temperature and Viscosity Tune the Intermediates during the Collapse of a Polymer," *Macromolecules* **57**, 10586–10599 (2024).
- <sup>42</sup>P. Palenčár and T. Bleha, "Molecular dynamics simulations of the folding of poly(alanine) peptides," *J. Mol. Model.* **17**, 2367–2374 (2011).
- <sup>43</sup>E. Arashiro, J. R. Drugowich De Felício, and U. H. E. Hansmann, "Short-time dynamics of the helix-coil transition in polypeptides," *Phys. Rev. E* **73**, 040902(R) (2006).
- <sup>44</sup>E. Arashiro, J. R. Drugowich De Felício, and U. H. E. Hansmann, "Short-time dynamics of polypeptides," *The Journal of Chemical Physics* **126**, 045107 (2007).
- <sup>45</sup>P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L.-P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, and V. S. Pande, "OpenMM 7: Rapid development of high performance algorithms for molecular dynamics," *PLOS Comput. Biol.* **13**, e1005659 (2017).
- <sup>46</sup>J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, and C. Simmerling, "ff14sb: improving the accuracy of protein side chain and backbone parameters from ff99sb," *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
- <sup>47</sup>H. Nguyen, D. R. Roe, and C. Simmerling, "Improved generalized Born solvent model parameters for protein simulations," *J. Chem. Theory Comput.* **9**, 2020–2034 (2013).
- <sup>48</sup>B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans* (Society for Industrial and Applied Mathematics, Philadelphia, 1982).
- <sup>49</sup>B. Efron and C. Stein, "The jackknife estimate of variance," *Ann. Statist.* **9**, 586–596 (1981).
- <sup>50</sup>R. G. Miller, "The jackknife—a review," *Biometrika* **61**, 1–15 (1974).
- <sup>51</sup>D. N. Theodorou and U. W. Suter, "Shape of unperturbed linear polymers: polypropylene," *Macromolecules* **18**, 1206–1214 (1985).
- <sup>52</sup>V. Blavatska and W. Janke, "Shape anisotropy of polymers in disordered environment," *J. Chem. Phys.* **133**, 184903 (2010).
- <sup>53</sup>H. Arkin and W. Janke, "Gyration tensor based analysis of the shapes of polymer chains in an attractive spherical cage," *J. Chem. Phys.* **138**, 054904 (2013).
- <sup>54</sup>K. Alim and E. Frey, "Shapes of semiflexible polymer rings," *Phys. Rev. Lett.* **99**, 198102 (2007).
- <sup>55</sup>K. Ostermeier, K. Alim, and E. Frey, "Buckling of stiff polymer rings in weak spherical confinement," *Phys. Rev. E* **81**, 061802 (2010).
- <sup>56</sup>W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules," *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).
- <sup>57</sup>W. Kabsch and C. Sander, "Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features," *Biopolymers* **22**, 2577–2637 (1983).
- <sup>58</sup>D. Barlow and J. Thornton, "Helix geometry in proteins," *J. Mol. Biol.* **201**, 601–619 (1988).
- <sup>59</sup>A. J. Doig, "The  $\alpha$ -helix as the simplest protein model: Helix-coil theory, stability, and design," in *Protein Folding, Misfolding and Aggregation: Classical Themes and Novel Approaches*, edited by V. Muñoz (Royal Society of Chemistry, Cambridge, 2008) pp. 1–27.
- <sup>60</sup>G. Némethy, D. C. Phillips, S. J. Leach, and H. A. Scheraga, "A second right-handed helical structure with the parameters of the Pauling-Corey  $\alpha$ -helix," *Nature* **214**, 363–365 (1967).

- <sup>61</sup>E. Baker and R. Hubbard, "Hydrogen bonding in globular proteins," *Prog. Biophys. Mol. Biol.* **44**, 97–179 (1984).
- <sup>62</sup>C. A. Rohl and A. J. Doig, "Models for the  $3_{10}$ -helix/coil,  $\pi$ -helix/coil, and  $\alpha$ -helix/ $3_{10}$ -helix/coil transitions in isolated peptides," *Protein Sci.* **5**, 1687–1696 (1996).
- <sup>63</sup>P. Palenčár and T. Bleha, "Folding of polyalanine into helical hairpins," *Macromol. Theory Simul.* **19**, 488–495 (2010).
- <sup>64</sup>P. Palenčár and T. Bleha, "Folding of  $\alpha$ -helices into bundles in long polyalanines," *Comput. Theor. Chem.* **1006**, 62–69 (2013).
- <sup>65</sup>J. Tirado-Rives, D. S. Maxwell, and W. L. Jorgensen, "Molecular dynamics and Monte Carlo simulations favor the  $\alpha$ -helical form for alanine-based peptides in water," *J. Am. Chem. Soc.* **115**, 11590–11593 (1993).
- <sup>66</sup>S. Huo and J. E. Straub, "Direct computation of long time processes in peptides and proteins: Reaction path study of the coil-to-helix transition in polyalanine," *Proteins* **36**, 249–261 (1999).

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0245056