Forced Unfolding of Protein-Inspired Single-Chain Random Heteropolymers


As Published: 10.1021/acs.macromol.1c02411

Publisher: American Chemical Society (ACS)

Persistent URL: https://hdl.handle.net/1721.1/142471

Version: Author’s final manuscript: final author’s manuscript post peer review, without publisher’s formatting or copy editing

Terms of use: Attribution-NonCommercial-ShareAlike 4.0 International
Forced unfolding of protein-inspired single-chain random heteropolymers

Zexiang Han†§, Shayna L. Hilburg†§, Alfredo Alexander-Katz†,*

§These authors contributed equally.

†Department of Materials Science and Engineering, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA

*Correspondence: aalexand@mit.edu

KEYWORDS. Random heteropolymer (RHP), bioinspired, compactification, single-molecule mechanics, unfolding, topology, atomistic molecular dynamics, PEGylation

For Table of Contents use only

MMA: hydrophobic
OEGMA: mostly hydrophilic, high-molecular weight
EHMA: hydrophobic
SPMA: charged
ABSTRACT

Synthetic random heteropolymers (RHPs) with high chemical heterogeneity can self-assemble into single-chain nanoparticles that exhibit features reminiscent of natural proteins, such as topological polymorphism. Using all-atom molecular dynamics simulations, this work investigates the structure and single-chain mechanical unfolding of a library of four-component RHPs in water, studying the effects of sequence, composition, configuration, and molecular weight. Results show that compactified RHPs can have highly dynamic unfolding behaviors which are dominated by complex side-chain interactions and prove markedly different from their homopolymer counterparts. For a given sequence and conformation, an RHP’s backbone topology can strongly impact its unfolding response, hinting at the importance of topological design in the nanoscale mechanics of heteropolymers. In addition, we identify enthalpically-driven reconfiguration upon unfolding, observing a solvent-shielding protection mechanism similar to protein stabilization by PEGylation. This work provides the first computational evidence for the force-induced unfolding of protein-inspired multicomponent heteropolymers.

INTRODUCTION

The need to study single molecules stemmed from the desire to develop a more thorough understanding of the biophysics of proteins or nucleic acids when subjected to a form of stress. Compared to chemical or thermal stimuli, mechanical force serves as a distinct and orthogonal strategy to study single-molecule mechanics.\textsuperscript{1} For folded proteins, highly heterogeneous mechanical responses are possible and proteins can undertake a range of mechanical (load-bearing or mechanosensing) and non-mechanical functions. One of the protein systems most commonly studied for its mechanical behavior is titin, a muscle protein exhibiting high elasticity. Force-
induced unfolding of a titin molecule reveals that its modular domains extend sequentially and independently, where each domain unfolding event gives rise to one force peak (or rupture force), attributable to the severing of a specific set of interstrand hydrogen bonds.\textsuperscript{3–4} Titin, therefore, gives rise to sawtooth patterns in force curves, which are also seen for other proteins for which secondary and/or tertiary structure lead to defined domains, such as spectrin and ankyrin.\textsuperscript{3,5,6} More generally, force-extension profiles for protein unfolding have distinguishable features arising from the sequence-defined structural interactions between amino acid side groups, such as the separation of $\beta$-sheets and the uncoiling of $\alpha$-helices. It naturally follows that, through examination of the unfolding behavior, structural information can be gained. Through understanding the relative forces required for dissociation of various portions of the proteins, insights into how the molecules may respond when disrupted by an external force—be it mechanical, chemical, or thermal—are provided.

While unfolding has been studied extensively for proteins and other natural biomacromolecules, there exist fewer studies on the single-chain mechanics of synthetic macromolecules with significant chemical complexity. Unfolding of hydrophobic homopolymer globules is well-studied;\textsuperscript{7–9} though, more complex collapsed synthetic polymers are less well characterized. A subset of these complex macromolecules, which have been examined somewhat more thoroughly, assemble through the incorporation of modularity and orthogonal chemistries.\textsuperscript{10–12} Chung \textit{et al.} designed a modular polymer mimicking titin, exploiting the ability of the constituent monomer 2-ureido-4-[1H]-pyrimidinone (UPy) to self-dimerize through hydrogen bonding.\textsuperscript{10} Hosono \textit{et al.} reported the mechanical unfolding of a similarly designed single-chain nanoparticle (SCNP), where the pendants within the chain can intramolecularly crosslink, leading to supramolecular self-assembly or self-collapse.\textsuperscript{11} As a result of such modularity, these two biomimetic polymers display
characteristic sawtooth patterns in their respective force-extension profiles and stepwise unfolding pathways. Metal-π coordination chemistry has also been exploited for creating self-folding single-chains in which rupturing of transient linkages can dissipate energy. These examples rely on specific, directional interactions. Fewer studies investigate the unfolding of globularly structured SCNPs that assemble due to nonspecific interactions. Geissler and Shakhnovich were the first to lay out an analytical treatment for the mechanical unfolding of general random heteropolymers (RHPs). They proposed that, during heteropolymer unfolding, there will likely have a pearl-necklace-shaped intermediate. The existence of this morphology can be ascribed to solvation effects, where hydrophilic regions are prone to extend upon unfolding whilst hydrophobic regions remain collapsed and compact to minimize solvent exposure, forming “pearls”. We expect that the stability of pearl-necklace morphologies will be dependent on the exact sequence. For instance, a heteropolymer where hydrophobic clusters are periodically spaced within the chain will likely experience this necklace-like intermediate compared to one that is completely random. While insights can be gained from these studies, they lack chemical detail which we have previously demonstrated to be vital to understanding specific random heteropolymer assembly.

Of the existing reports on the forced unfolding of macromolecules, most work is experimentally enabled by a suite of single-molecule force spectroscopy techniques including atomic force microscopy, optical tweezers, and magnetic tweezers. However, these techniques fail to help visualize the purported unfolding events. An area for exploration is in situ imaging of single molecules during mechanical unfolding for multimodal analysis. Both single-molecule fluorescence imaging and in situ liquid-cell electron microscopy may be possible candidates, though their developments are only in their infancy. In silico methods modeling single-molecule nanomechanics are therefore an attractive alternative, capable of providing mechanistic
insights. Common computational methods to study single-molecule mechanical response include all-atom steered molecular dynamics (SMD)\textsuperscript{2,20,21}, coarse-grained Brownian dynamics\textsuperscript{7}, and Monte Carlo simulations\textsuperscript{22}. Among these, all-atom SMD is particularly favored for studying the forced unfolding of chemically heterogenous biomolecules due to the atomistic resolution that proves essential to capturing their conformational flexibility and diversity. Atomistic modeling also explicitly includes polar interactions with water molecules, which is particularly useful in modeling the unique amphiphilic behavior of polyethylene glycol (PEG).\textsuperscript{23,24} One evident limitation of atomistic SMD and similar computational techniques is the difficulty to access micro- to millisecond timescales due to a correspondingly high computational cost. As a result, the typical pulling speed employed in SMD lies in the range of $10 - 100 \text{ Å ns}^{-1}$ (Table S1), which is six to seven magnitudes faster compared to those used in experimental methods.\textsuperscript{20-22,25-31} Due to the orders-of-magnitude difference in pulling velocities, SMD results often overestimate force peak values compared to empirical values. In spite of this, it is still common to correlate experimental findings with simulations in order to gain mechanistic insights into the unfolding events and elucidate unfolding pathway(s), and there is typically a satisfactory qualitative agreement between simulations and experiments.\textsuperscript{21,32}

Overall, single-molecule mechanics is of paramount importance for understanding the internal structure and response to external forces of polymeric chains. Using a recently reported RHP system as an example\textsuperscript{33}, this work characterizes the single-chain mechanical response of a highly chemically heterogeneous polymer system. Xu and colleagues took inspiration from natural proteins and rationally designed methacrylate-based statistical RHPs, which serve as a novel class of biomimetic materials.\textsuperscript{33} These four-component RHPs incorporate methyl methacrylate (MMA), oligo(ethylene glycol) methacrylate (OEGMA), 2-ethylhexyl methacrylate (EHMA), and 3-
sulfopropyl methacrylate (SPMA) (Figure 1A). The monomer selections are intended to leverage varied amphiphilicity and polarity to recapitulate the heterogeneity of native protein chains and resulting surfaces.\textsuperscript{15} Such a design differs from many past SCNPs, which rely on intramolecular crosslinking strategies, and more closely resembles the self-assembly of natural biomacromolecules.\textsuperscript{10-12} The RHPs can mimic protein functions and interface favorably with proteins, finding applications as synthetic alternatives to molecular chaperones to stabilize proteins in non-native environments\textsuperscript{33,34}, as mimics of transmembrane proteins to facilitate selective proton transport\textsuperscript{35}, and as enzyme protectants to facilitate the degradation of commodity plastics\textsuperscript{36}. Atomistic modeling of these RHPs in water revealed that while the RHP structures are not sequence-defined, some structural motifs emerge in their collapsed form and a variety of assemblies are possible.\textsuperscript{15} We also showed that the RHPs possess minimal backbone mobility in water; however, experiments have demonstrated that the RHPs can interact with other biomacromolecules and small molecules, suggesting that external stimuli could provide the driving force to at least partially unfold portions of the molecules. Herein, we perform all-atom MD simulations to characterize RHPs with degrees of polymerization of 20 and 50 (referred to herein as 20mers and 50mers, respectively) in water. Sequences of various compositions and arrangements of the four methacrylate-based components are investigated (Figure 1A) using SMD to study their unfolding (Figure 1B). To the best of our knowledge, no prior work has been performed to investigate the unfolding behavior of synthetic heteropolymers as chemically complex as the four-component amphiphilic RHPs presented here. Aforementioned theoretical treatments of heteropolymers often neglect atomistic details and cannot accurately capture the conformational complexity of chemically heterogeneous polymers.\textsuperscript{13,14,37} Thus, our investigation of this bioinspired RHP system not only adds a different chemistry to the current portfolio of forced
unfolding of synthetic heteropolymers, but also affords a library of polymer sequences for investigation. Moreover, by understanding how the RHP responds to a tensile force stimulus, we can appreciate what would likely be required for backbone remodeling to take place and gain insights into internal structural dynamics and stability of the RHPs in water.

Figure 1. Protein-inspired RHP studied in this work. (A) RHP chemical structure. Monomers are color-coded as follows: MMA in black, OEGMA in blue, EHMA in red, and SPMA in yellow. (B) Schematic showing the forced unfolding of an RHP from its collapsed state.

RESULTS & DISCUSSION

RHP chain compactification

The self-assembly of explicitly solvated single-chain heteropolymers is studied, with sequence schematics of all 20mer and 50mer RHPs investigated in this work provided in Figure S1. Similarly to 100mer RHPs of the same chemistry, individual 20mer and 50mer RHP sequences can self-assemble into multiple conformational states. For a given chain, 10 annealing cycles lead to the sampling of 10 distinct conformers, each with unique topological organization in the backbone (Figures S2 and S3). Such topological heterogeneity has also been observed in certain intrinsically disordered proteins as well as synthetic SCNPs of similar sizes, alluding to the rich
conformational energy landscape of these comparable systems. Standard deviation of the absolute value of backbone dihedral angles is used as a measure for backbone mobility over the relevant timescale. By this measure, the ten conformational states sampled are believed to be metastable in nature as the backbones minimally change, as illustrated by the small standard deviation of dihedral angles over 40 ns for the equilibrated system in water (Figure 2A). For both 20mers and 50mers, the ends of the RHPs are, as expected by the configurational entropy of linear polymer chains, more mobile than the middle.

Mobility comparisons between RHPs of different molecular weights show that both the middle and end segments of the 50mers reconfigure less than their counterparts in the shorter polymers in unbiased MD simulations. From this trend, we can imply that the longer polymers are more compact, impeding backbone rearrangement. This is confirmed by Figure 2B, whereby an analog for density (polymer mass divided by the radius of gyration, $R_g$, cubed) generally increases with molecular weight. One notable exception is sequence 19, which shows a decrease in density when comparing its 20 and 50mer lengths. This result stems from the anion-anion repulsion of the SPMA monomers, of which sequence 19 has the highest proportion investigated in this work, leading to polyelectrolyte-like behavior. We would not expect an RHP with a high negative charge to compactify into a globular morphology; instead, an amphiphilic polymer with a high net charge would adopt a more extended conformation, giving rise to a lower density. The typical RHP trend, however, shows compactification as well as a narrowing of the range of densities between sequence conformations as molecular weight increases. The narrower range of density values stems from the greater similarity in $R_g$ values for 50mers, as despite different backbone topologies, configurations all led to compact globules. For 20mers, some sequence conformations formed denser assemblies while others remained extended, indicating a stochastic compactification with
close energetic competition between the entropic cost of limiting mobility in a compact globule and the enthalpic surface energy penalty of exposing hydrophobic monomers to the aqueous environment. As smaller oligomers, the RHPs will be soluble, because even MMA, one of the more hydrophobic monomers in our polymer, is soluble at extremely low degrees of polymerization.\textsuperscript{42} Therefore, based on the simulation results, most compositions of RHPs with degree of polymerization of 20 appear to be near the energetic cliff for compact globule formation, while 50mers in the same windows are nearly all compact and more uniform in density.

\textbf{Figure 2. RHP chain mobility and its compactification.} (A) (Top) Box-and-whiskers plot of the raw standard deviation $\sigma$ of dihedral angle for each conformation from the 40 ns of equilibration, for all 20mer and 50mer RHP sequences and conformations, respectively. Dashed lines represent the mean. The “ends” refer to the 8 dihedral angles from the 5 monomers on each end (thus in total 16 dihedrals) regardless of the RHP length; the remaining dihedrals are in the “middle” of the RHP. (Bottom) Percentage of $\sigma$ of dihedral angle values greater than 20° for RHP ends vs. middle over the 40 ns of equilibration. Error bars represent standard error around the mean for the 200
conformations (20 sequences with 10 conformations each). (B) Density analog for all 20mer and 50mer sequence conformations versus molecular weight (MW). The density analog in g cm\(^{-3}\) is calculated as MW \(\times\) Avogadro’s number \(N_A / R_g^3\), with appropriate unit conversions. RHPs generally show compactification as molecular weight increases, with several exceptions explained in text. Corresponding sequence schematics are given in Figure S1.

**RHP responses to force-induced unfolding**

For a selection of polymer sequences, five independent replicates were studied under an applied tensile force to induce unfolding at a constant rate of 1 Å ns\(^{-1}\). To ensure a sufficiently slow pulling rate for our RHP system, we perform *in silico* stress-relaxation experiments on one 50mer RHP, sequence 12. Snapshots during the unfolding simulation were extracted every 10 ns (i.e., 0 ns, 10 ns, 20 ns, ... 100 ns), and each is then stress-relaxed by maintaining the end-to-end distance restraint and allowing the chain to equilibrate for 20 ns while monitoring for relaxation behaviors (Figure 3A). The backbone dihedrals show only a few changes upon stress relaxation for essentially all unfolding intermediates (Figure 3B and S4), indicating relatively insignificant backbone reconfiguration. In addition, the magnitudes of changes in the dihedral angles for the partially unfolded structures are not far from those for the initial equilibrated structure (Figure 3C, top), and the mobility of the backbone of the structural intermediates remains low (Figure 3C, bottom). This suggests that the initial structural snapshots do not deviate much from their stress-relaxed states. We also see minimal reconfiguration within the side-chains, which are generally more mobile than the backbone, during the 20-ns stress relaxation. In fact, RHP solvation, which is dominated by side-chain/water interactions, remains nearly constant over the course of the stress-relaxation (Figures 3D and S5), indicating extremely rapid water solvation. Provided this
minimal RHP reconfiguration, the pulling rate employed is sufficiently slow for gaining mechanistic insights into the RHP behavior upon unfolding. Our results show that the RHPs are in a pseudo-equilibrium regime during force-induced unfolding, and the unfolding observed through our procedures is likely a low energy pathway.

Figure 3. Insignificant backbone and side-chain remodeling upon RHP stress-relaxation. Results shown here are for 50mer RHP sequence 12 conformation 8. (A) Illustration of the stress-relaxation protocol on an RHP snapshot from unfolding time $t = 60$ ns. End-to-end distance is maintained constant during the 20-ns stress-relaxation. Note that the stress-relaxation time is represented by $\tau$ in order to differentiate from the unfolding time $t$. (B) Stress-relaxation responses
in the backbone of the snapshot RHP structure from \( t = 60 \) ns. Time evolution of dihedral plots, each averaged over 2 ns. Representative time-averaged \( C_\alpha \)-based contact map evolution, confirming insignificant remodeling in the RHP backbone. (C) (Top) Average standard deviation \( \sigma \) of the absolute value of all 97 dihedral angles throughout the 20-ns stress-relaxation against different unfolding intermediates. The end-to-end distances correspond to the extracted RHP snapshots from time \( t = 0 \) ns to \( t = 100 \) ns during its unfolding process. Each data point is from one independent stress-relaxation simulation. (Bottom) Average number of dihedrals with a \( \sigma \) greater than 20°. Error bars represent standard errors around the mean. (D) Normalized first solvation shell (NFSS) along the sequence for different unfolding intermediates for (top) stress-relaxation and (bottom) mechanical unfolding, showing minimal change in solvation after RHP stress-relaxation. The corresponding time points for each of the end-to-end distances are time-averaged: for stress-relaxation, they refer to the average NFSS for the last 2 ns of the stress-relaxation, and data from three replicates are shown. For unfolding, they refer to the average NFSS at the corresponding unfolding times \( t \) (for example, 9.7 Å means averaging across \( t = 0–2 \) ns, and 19.7 Å means averaging across \( t = 9–11 \) ns).

Force curves are recorded from each constant-velocity unfolding replicate using SMD. There thus exist three possible unfolding events which can have varying extent of overlap: (i) concerted breakage of a set of noncovalent interactions, which require a high force/energy; (ii) breakage of noncovalent interactions one by one, or “unzipping”, which require a lower unfolding force that spans over a longer range of extension, as well as (iii) breakage of a set of dynamically evolving intramolecular interactions, whose force curve features will likely be diffuse or stochastic and more difficult to interpret. For our RHPs, visual inspection of the unfolding trajectories does not
establish an unambiguous correlation between force curve features – peaks or plateaus – with molecular snapshots. Examination of individual unfolding replicates of a given RHP conformer shows that side-chains have highly dynamic interactions that vary across the five replicates. For example, the solvent-accessible surface area (SASA) evolution over time exhibits different behaviors across the five replicates, even though some force curves exhibit very similar features for a given 20mer RHP (Figure S8). Variable pathways for backbone restructuring are also observed through the evolution of dihedral angles. We can hence suspect that force curve features are a result of a combination of topological variations in the polymer backbone as well as side-chain interactions rather than a one-to-one correlation of a particular force peak to the disassociation of two moieties. Though each RHP sequence conformation replicate has a unique unfolding trajectory, they share commonalities reflected in the averaged force curve from which we can gain insight to the polymer assembly.

Since the coil-to-globule transition for hydrophobic homopolymers is well studied, the force responses for unfolding collapsed homopolymers of chemistries relevant to our RHPs – namely, PMMA and PEHMA – are studied as controls (Figure 4). The unfolding force is rate-dependent in polymeric systems; therefore, the magnitude of the unfolding forces recorded in SMD simulations will be higher than measured by experimental single-molecule techniques which operate at a pulling velocity orders of magnitudes lower. For PMMA, an initial force peak is present in several of the simulations, and is attributable to the disruption of intrachain dipole-dipole interactions due to PMMA’s directional and polar side group. Once the dipole-dipole self-packing interactions are disrupted upon initial unfolding, polar side groups of PMMA can become solvated with water, and it only requires minimal (near-zero) force to unfold the chain further, resulting in a force curve traditionally consistent with PMMA’s glassy nature. For PEHMA homopolymers, their force
curves exhibit more features and show a greater variability between the unfolding replicates
compared to that of PMMA, demonstrating highly dynamic side-chain interactions, as seen by
visual inspection in the simulation trajectories. The averaged force curve for PEHMA_{50} also
displays a characteristic plateau-like behavior at a non-zero force which is ascribed to hydrophobic
hydration.\textsuperscript{9}

\textbf{Figure 4. Unfolding force curves of hydrophobic homopolymers.} Force curves for (A) 20mer
and (B) 50mer homopolymers PMMA (black) and PEHMA (red) are based on moving averages
over 2-Å intervals. Shown are the data from five independent unfolding simulation replicates in
different shades of the same color, and the average is given in a thicker dotted line. Example
snapshots of initial, intermediate (midpoint), and final conformations of the polymer during
unfolding are provided, where side-chains have been rendered semi-translucent to highlight
polymer C–C backbone topologies.

For the multicomponent RHPs, their equilibrated structures are often stabilized by the
hydrophobic attraction between EHMA-EHMA side-chains. Some sequences also have
conformations stabilized by OEGMA-OEGMA interactions or by a mixture of EHMA and
OEGMA interactions, depending on the monomer availability within the chain. Compared to their
homopolymer counterparts, 20mer and 50mer RHPs display more varied mechanical responses
(Figures 5, S6, and S7). In terms of force curve features, some 20mer RHPs exhibit clear and
pronounced force peaks across five replicates, some exhibit a plateau behavior, and some can have
more frequent force peaks than others. This demonstrates that the RHP system has an extremely
rich energy landscape which can result in vastly different single-chain mechanical responses, even
for identical sequences as short as 20 monomers in length. Similarly, force curves of 50mer RHPs
also display varying extents of force peaks and plateaus depending on the exact sequence and
conformation of the chain. The diverse force curves for unfolding RHPs are distinct from those
observed for biological heteropolymers such as proteins or nucleic acids of similar sizes. For
example, the unfolding force curve of single-stranded DNA hairpin structures of 55 bases in size
reveals a characteristic “rip” feature, indicative of disruption of the hydrogen bonding. This
disparity arises from three major design differences between RHPs and biopolymers (e.g. proteins
or nucleic acids). First, RHPs do not have intramolecular hydrogen bonding that proves
fundamental for the secondary structural formation in proteins. Secondly, RHPs possess several
bulkier and longer side-chains compared to native amino acids that make up proteins (or nucleotides that make up nucleic acids), giving rise to unfolding behavior dominated by side-chain interactions and reconfiguration. Thirdly, RHPs have a racemic mixture of monomers with randomly assigned chiralities – that is, our RHPs are heterochiral – whereas proteins and nucleic acids are intrinsically homochiral. As a result, no characteristic rupture forces from the breakage of specific ordered interactions are observed in RHPs.

Figure 5. RHPs show varied responses to force-induced unfolding. Shown here are the force curves for the unfolding of selected 20mer RHPs with their unfolding trajectory snapshots (initial, intermediate, and final conformations). An exhaustive overview of the unfolding force curves of all 20mer and 50mer RHPs studied are provided in Figures S6 and S7, respectively.
Monitoring the unfolding pathway and behavior of individual RHP sequences

The unfolding behavior of an RHP is investigated in detail by monitoring the time evolution of dihedral angles and intramolecular contacts to understand backbone restructuring and elucidate the unfolding events. Here, we detail the unfolding pathway of a molten globular 50mer RHP, sequence 12 conformation 8. As with many 50mer RHPs, the mechanical unfolding of this RHP is mostly driven by the dissociation of the hydrophobic core formed by EHMA-EHMA interactions (Figures 6 and S9). During the first 30–40 ns of unfolding, there are few apparent dihedral angle transitions except for the two ends. In fact, the internal structure of the RHP is well preserved during the initial mechanical perturbation, confirmed by the preservation of contacts from pulling time $t = 1$ ns to 30 ns (Figure 6C). Therefore, conformational changes are possibly mediated by the extension of the two chain ends which possess higher mobility as demonstrated by unbiased MD simulations and by diffuse, small changes throughout the entire chain in order to enable the initial increase in end-to-end distance. On a morphological level, the globule becomes somewhat distorted under the applied force. From $t = 50$ to 60 ns, the chain undergoes segmental separation where the EHMA hydrophobic core dissociates into two separate, smaller EHMA-stabilized cores. Examining the dihedral evolution, this is believed to be mediated by the extension of the MMA block (highlighted by red boxes in Figure 6C), as indicated by the distinct gauche to trans angle transition. This pearl-like intermediate is similar in shape to that predicted by theory\textsuperscript{13,14}; however, in the case of this specific RHP, a hydrophobic MMA block instead of a hydrophilic cluster unfolds first. Upon segmental separation, the extensions of the two smaller units unfold sequentially, where the longer EHMA segment near the beginning of the sequence has persistent hydrophobic interactions until $t = 90$ ns and unfolds last (highlighted by green boxes in Figure 6C). Overall,
during the latter stages of the unfolding pathway (after $t = 50$ ns), EHMA side-chains, which previously had longer range hydrophobic interactions, have an increased number of local interactions due to their physical proximity within the sequence, giving rise to the emergence of some gauche angles in the more extended conformation. Analyses on the four other replicates of this RHP sequence and conformation also show relatively similar unfolding pathways; however, the segmental separation can take place at a different MMA block via its extension (Figure S10). This suggests that the MMA blocks can behave as a hinge, such as the $\alpha$-helical linker in spectrin$^4$, to mediate or propagate unfolding. This is supported by our finding that PMMA homopolymers require minimal force to unfold, leading MMA blocks within an RHP to be more amenable to reconfigure during unfolding. Nevertheless, it should be noted that MMA blocks do not always behave in this manner in the sequences investigated, and there are a myriad of complex interactions that would allow alternative unfolding pathways to exist, making this phenomenon highly sequence- and conformation-specific.

Figure.
Figure 6. Monitoring the unfolding pathway of a 50mer RHP. This particular RHP corresponds to sequence 12 conformation 8, and results from a representative unfolding replicate is shown here. (A) Unfolding force curve. Vertical dotted lines indicate specific time points of interest. (B) Backbone dihedral angle evolution for the entire unfolding process of this RHP. (C) Selected 2-ns-time-averaged $C_\alpha$-based contact maps. Red/green boxes highlight regions of interest. Corresponding VMD snapshots of the RHP are also shown to visualize the unfolding events. The unfolding trajectory showing the unfolding intermediates per every 10 ns is provided in Figure S9. The unfolding trajectory of this RHP visualized in VMD is provided in Supporting Information Movie S1.

Another finding observed in the study of sequence 12 conformation 8, as well as several other compositions and configurations, is that upon RHP unfolding, OEGMA side-chains wrap around the surface of the main RHP chain. Throughout the unfolding trajectory, different OEGMA side-chains preferentially interact with hydrophobic regions, as highlighted in Figure 7, where darker regions in the contact maps calculated from side-chain centers of mass denote the physical proximity between the side groups of each monomer. Since the contact maps are time-averaged over 2 ns, the residues have significant residence times rather than being coincidental instantaneous occurrences. OEGMA, shown to be amphiphilic in nature and whose side-chains possess high conformational flexibility, can establish a favorable interface between the hydrophobic monomers and the water molecules, serving as a possible protection mechanism via solvent shielding. To obtain direct, quantitative evidence of OEGMA wrapping, separation distances between the tail atoms of OEGMA and EHMA residues are computed and plotted against time for all possible OEGMA/EHMA pairs in the RHP sequence (Figure 7C). Small separation
distances between the two tail atoms can arise either from sequence effects due to mere geometric proximity between covalently bound monomers within the chain or from the wrapping phenomena we intend to capture. To decouple the two effects, EHMA monomers have been grouped into four groups based on their positions in this particular sequence. The fourth OEGMA monomer, highlighted in magenta, is actively involved in the wrapping process, as observed from the nonmonotonic time-dependence in the separation-distance curves with multiple regions of EHMA (Figure 7). OEGMA-EHMA distance curves for the other OEGMA residues do not manifest the same extent of wrapping (Figure S11). Moreover, this wrapping phenomenon has been observed for numerous 20mer and 50mer RHPs studied, establishing its relative generality. Whilst this observation is novel for our particular chemistry, the solvent-shielding phenomenon is reminiscent of previously studied interactions between PEG — which makes up the long OEGMA side-chain — and biomolecules. Conjugation of proteins with PEG chains (PEGylation) and similar molecular brushes is a method to enhance the thermal and mechanical stability of certain proteins using covalent modifications. While comprehensive studies providing mechanistic insights are limited, DeBenedictis et al. report that PEG chains can wrap around an α-helical protein and shield water molecules from attacking the hydrogen bonds, delaying the unfolding process. Moreover, they show that the PEG chains disproportionately favor hydrophobic and charged residues, phenomenologically similar to the OEGMA wrapping in our system.
Figure 7. OEGMA side-chain wrapping is observed during RHP unfolding. (A) RHP sequence schematic (sequence 12) and a structural snapshot at pulling time \( t = 50 \) ns. EHMA monomers within this sequence can be approximately grouped into four sets, color-coded in red, lime green, blue, and purple. OEGMA monomers are numbered and color-coded as shown in the sequence. The same color-coding is used to indicate OEGMA monomers in the snapshot, showing OEGMA wrapping around hydrophobic EHMA and charged SPMA residues. The OEGMA side-chains have been rendered semi-translucent and water has been excluded for ease of visualization. (B) A time-averaged side-chain-based contact map showing preferential interactions of some OEGMA side-chains with hydrophobic residues. (C) Time evolution of tail-atom separation distances for the second OEGMA monomer to the other EHMA monomers based on 2-ns-time-averaged tail-atom contact map data. Therefore, each line in the plot corresponds to the tail-atom distance of one particular OEGMA/EHMA pair. Similar plots for the other OEGMA monomers can be found in Figure S11.
Differently positioned OEGMA monomers have side-chains wrap around the main chain to varying extents at different time points of unfolding. Consequently, the wrapping behavior is not a one-size-fits-all process, and sequence effects play an important role to enable the protection mechanism at a given time point. This is analogous to protein PEGylation, where the site of conjugation in a protein also affects the extent of mechanical reinforcement imparted by PEG chains.\textsuperscript{50,51} For our RHPs, per the side-chain-based contact maps provided in Figure S11, it may be hypothesized that OEGMA monomers surrounded by MMA monomers are less prone to partake in the proposed protection mechanism \textit{via} side-chain wrapping. Previous analysis has shown that OEGMA monomers in equilibrated RHPs are better solvated when they are surrounded by MMA residues compared to when surrounded by EHMA residues\textsuperscript{15}, corroborating our observations. Since EHMA is more hydrophobic than MMA, side-chain wrapping around the EHMA residues to minimize water contact is energetically beneficial for the system, suggesting a protection mechanism at play.

\textit{Physicochemical factors affecting RHP unfolding}

A variety of physicochemical parameters, including chain length, chemical composition, sequence characteristics, and backbone topology, can influence the RHP unfolding pathway, force curve features, mechanostability, and nonequilibrium unfolding work. Specifically, RHP mechanostability, or the mechanical resistance of RHPs to the applied tensile force, is characterized using peak forces and their distributions. First, on average, RHPs of 50 monomers in length are found to be more mechanically stable than the 20mers (Figure S12). This can be reasoned from the previous conclusion that 50mers typically begin from a state of greater chain compactification due to lower solubility and higher propensity for hydrophobic collapse. As a result of such compactification, the unfolding pathways of 20mers and 50mers also differ. Many
20mer RHPs and homopolymers can unfold via an Ω-shaped topological intermediate (Figure 5). Since this is common to sequences having different chemical compositions and sequence traits, the Ω-shaped intermediate may have a topological origin given that most compact 20mers investigated initially assume a U- or O-shape, where the 2D projection of an RHP backbone onto any surface gives a U-shape if there is no intersection and an O-shape if there is one intersection. Notably, an unfolding pathway phenomenologically similar to the pathway proposed here has been observed during the mechanical unfolding of biological β-hairpin structures52,53, despite the fact that specific hydrogen bonding is responsible for stabilizing β-hairpins whereas nonspecific hydrophobic interactions are responsible for stabilizing our RHPs. 50mer RHPs, on the other hand, have more diverse topologies and, as a result, more diverse unfolding behaviors. Previous theoretical analysis on amphiphilic heteropolymer unfolding suggests the existence of a pearl-necklace unfolding intermediate due to favorable solvation in hydrophilic residues and unfavorable solvation in hydrophobic ones.13,14 However, this is rarely seen in the multicomponent RHPs. When the pearl-necklace unfolded intermediate is observed, we find that the pearl formation is not necessarily due to the unfolding of hydrophilic cluster(s) within the polymer sequence; rather, reduced steric hindrances in MMA segments facilitate unfolding through a linker-like mechanism as discussed in the case study. In addition, we suspect that the solvent-shielding protection mechanism discussed previously can alter the unfolding pathway by modulating the barriers pertinent to the water solvation energetics, thus eliminating the need for a necklace structure during the force-induced globule-coil transition.54 Therefore, we find that, in addition to hydrophobicity and hydrophilicity of the monomers, their exact chemistry and steric impact the RHP unfolding response, which would not have been captured by coarse-grained simulations and theories.
Our RHPs possess high chemical heterogeneity with lengthy and/or bulky side-chains that prove impactful to their unfolding responses. Since chemical composition, sequence characteristics, and backbone topologies cannot be easily decoupled for our statistically random polymers, these are examined holistically to provide insights into their effects on unfolding and mechanostability. The effect of chemical composition is investigated by examining the monomer content in each chain using pooled data across sequences (Figures S13 and S14). No strong correlations are observed for 20 or 50mer sequences, indicating that chain composition alone cannot dictate the mechanical response of a chemically heterogeneous RHP. One might expect that, as the most prevalent stabilizing interactions in RHPs are EHMA-EHMA hydrophobic attraction, its content would have a high correlation with mechanostability. However, even though PEHMA homopolymer is relatively mechanically stable, having a high fractional content of EHMA in the RHP does not guarantee the same. A weak negative correlation can be observed between SPMA content and RHP mechanostability due to electrostatic repulsions, though RHPs with the same chemical composition (20mer RHP sequences 9 and 13) can produce disparate mechanical responses. These results hint that sequence and/or topological effects may be of more relevance than chemical composition for mechanostability.

The work required for unfolding an RHP is calculated as the area under the force curve, with the integral evaluated from initial end-to-end distance to 45 Å for 20mers and 110 Å for 50mers. Unfolding work is rate-dependent for polymeric systems, but can nevertheless provide information on the internal friction in polymeric globules, which in turn reflects the roughness of their conformational energy landscape. Figure 8 maps the specific unfolding work against the change in specific total SASA upon polymer unfolding for homopolymers and RHPs, where “specific” properties are normalized by the polymer molecular weights. The change in SASA
should encompass effects of both molecular weight and overall hydrophilicity, allowing each sequence to be compared. We notice that there are several cases where $\Delta \text{SASA}/\text{MW}$ is negative (Figure 8), all of which come from different RHP unfolding replicates. This can be explained by the amphiphilic nature of OEGMA side-chains as well as the fact that OEGMA is a main determinant of the magnitude of the overall SASA as it has lengthy side-groups. The PEG chains in OEGMA residues can either wrap around the main chain as previously discussed, giving rise to relatively low values of total SASA values, or become fully solvated, giving rise to high SASA. The balance between the two at the start and end of the unfolding trajectory then partly dictates the magnitude of $\Delta \text{SASA}$ – where some of which can be negative in value. First, comparing 20mers and 50mers, the longer polymers which compactify more require a greater specific unfolding work. In addition, relative to the limits established by the homopolymers, namely PMMA and PEHMA specific unfolding work, the 50mer data also has a tighter distribution. Our chemically heterogeneous 50mer RHPs experience relatively uniform compactification and are more dense compared to 20mers, suggesting that the specific unfolding work correlates well with polymer compactness and molecular weight. This has also been demonstrated on unfolded or disordered proteins that the magnitude of internal friction correlates with protein compactness. Closer examination of the nonequilibrium unfolding work shows that there are minimal trends relating unfolding work to the chemical composition of the polymer chains and corroborate the observation that all of the 50mer RHPs studied have similar extents of internal friction (Figure S15). A weak positive correlation exists for unfolding work and the number of OEGMA monomers in a 20mer RHP. Since OEGMA has long side-chains, giving a brush-like architecture, it will have the greatest contribution to internal friction which in our case would predominantly be due to intramolecular side-chain interactions.
Figure 8. Specific non-equilibrium unfolding work (work / molecular weight MW) for homopolymers and RHPs versus the change in specific total SASA (ΔSASA / MW). ΔSASA is calculated by subtracting the average total SASA of the polymer in the first 2 ns of the unfolding simulation from that in the last 2 ns of the simulation. Individual data points are from each unfolding replicate, including five replicates for each sequence of thirty 20mer RHP sequence/conformation, nineteen 50mer RHP sequence/conformation, two PMMA_{20}, one PMMA_{50}, two PEHMA_{20}, and four PEHMA_{50}. A similar plot where calculated unfolding work is based on a defined chain extension interval is shown in Figure S16.

Examining the effect of sequence characteristics is less obvious for our data, yet postulates can be put forward. Chains with alternating hydrophilic (OEGMA or SPMA) and hydrophobic (MMA or EHMA) monomers were likely to display force curves that are generally increasing in a monotonic fashion without pronounced features (Figure 4). Similarly, sequence effects on the mechanical response of 50mer RHPs can be rationalized on a case-by-case basis, albeit not
deterministically. RHPs with high SPMA content and SPMA spaced out within a chain result in very mechanically labile unfolding because electrostatic repulsions between anionic groups facilitate the hydration and unfolding of the chain. Yet, there are exceptions, and this is not the case for sequence 5. The destabilizing effect of having a relatively high SPMA content is counteracted by high hydrophobic content, leading to more compact chain morphology upon hydrophobic collapse and giving rise to pronounced peak features. To enable a predictable power for rational design of the multicomponent RHPs, further and more systematic analysis on sequence characteristics is needed. This points to future work on developing statistical models that can be used to assess and quantify sequence traits. Similar analysis has been carried out in protein homology considering the vast number of possible permutations of the amino acid constituents, which can be informative when investigating the RHP systems.57

Topological design in heteropolymer systems is an emerging topic in macromolecular engineering.58,59 Our RHPs provide a library of sequences with distinct metastable conformational states stabilized by reversible intramolecular interactions, offering opportunities to decouple sequence and topological effects. We first select two compact 20mer conformers and compare their peak force distributions (Figure S17), noting that different initial equilibrated structures of identical 20mer sequences can have dissimilar mechanical responses to the tensile stimulus. While 20mers have backbones that are topologically simple and similar, 50mers have more diverse conformations and topological organizations and can further shed light on how polymer topology affects the single-chain mechanics of RHPs. Four conformations each of the homopolymer PEHMA and of two RHP sequences (sequence 15 and sequence 19) have been investigated. Visual inspection of the different initial conformations reveals that the homopolymers are more spherical in shape – though the exact chain topologies differ for the four conformations, whereas the two
heteropolymers in question adopt more varied morphologies (Figure 9). For both homo- and heteropolymers, unique chain topologies result in distinct mechanical responses to forced unfolding (Figure 9). This aligns with previous results on proteins, whose native topology can greatly impact its unfolding behavior, as the initial chain topology restrains the progression of unfolding events that can take place.\textsuperscript{50-62} For the heteropolymers, the effect of chain topology is more evident. Structure 2 of sequence 15 has a slightly planar morphology, and the two chain ends point in opposite directions. As a result, mechanical unfolding of this conformation leads to an initial resistance, followed by a drop to near-zero force, indicating nearly spontaneous unfolding behavior at relatively high chain extension. Closer scrutiny of the initial resistance reveals two contributing factors. During initial mechanical perturbation of the polymer, the planar-like morphology deforms laterally and becomes more prolate/ellipsoidal (Figure S18). The internal friction from OEGMA side-chain interactions likely contribute to the initial peak observed in the averaged force curve. A tadpole-like unfolding intermediate, where a globule-shaped head is in coexistence with a stretched chain end\textsuperscript{63}, is observed at around $t = 30$ ns for this structure but not the other topologies of the same RHP sequence. At around $t = 40–45$ ns, EHMA-EHMA hydrophobic interactions become disrupted to allow further chain extension, after which force drops to a near-zero value as essentially all stabilizing intramolecular interactions have been disrupted, and no further pronounced force peaks are observed (Figures 9B and S18). On the other hand, prolate structures of the heteropolymer (structures 2 and 3) can display a relatively sustained mechanical resistance over a large range of end-to-end distances (Figure 9). Sequence 19, which is high in negatively-charged SPMA content, is chosen as another RHP for investigating the effect of topology. While two of the native conformations (structures 1 and 3) unfold easily, mostly attributable to electrostatic repulsion between anions which facilitates unfolding, the remaining
two conformations display pronounced force peak features in their unfolding force curves. Therefore, even with high SPMA content, there may exist topologies for which high mechanostability may be achieved. Interestingly, structure 4 of sequence 19 also requires the highest force (285 pN on average, with the highest being 410 pN in one of the replicates) to unfold compared to the different conformers of PEHMA or sequence 15. We hypothesize that, for this topological organization of the backbone, the arrangement of negatively-charged sulfate ions within the SPMA side-chains amplifies the strength of the hydrophobic attraction, similar to previous experimental and simulation results proving the modulation of hydrophobic effect by proximal (i.e., within 1 nm) covalently-attached charged moieties in an ion-specific fashion.\textsuperscript{54,65} For this RHP conformation (sequence 19 structure 4), its unfolding requires the cooperative breaking of multiple types of interactions at approximately the same time, leading to one pronounced force peak at low extension (Figure 9C).

Comparing the intramolecular contact evolutions of the unique topological conformations (Figures 9 and S19) offers additional insight. First, both the time-evolution of contact maps and the contact reduction trends are similar between replicates of a specific topology for a given polymer sequence. For the RHPs, this would suggest that, once we have a metastable structure, there exists a generally consistent energetic pathway for unfolding, though small deviations can exist. Secondly, we find that homopolymer unfolding leads to both dissolution and reformation of contacts whereas heteropolymer unfolding is predominantly mediated by the dissolution of contact features. Thus, backbone restructuring is more common in the homopolymer compared to in the heteropolymer upon unfolding, following the mobility trends obtained in previous unbiased simulations.\textsuperscript{15} From an enthalpic point of view, monomer-monomer interactions are all identical within homopolymers, whereas heteropolymers have a multitude of possible interactions between
the different monomer types. These inter-monomeric potentials lead to a much more textured energy landscape of the evolving RHP conformations since there are now ten unique monomer pairings rather than the single homopolymer self-interaction. Polarity, electrostatics, and hydrophilicity lead EHMA-EHMA and OEGMA-OEGMA associations to be favored while SPMA-SPMA and SPMA-EHMA interactions are avoided. From a topological perspective, all four conformations of PEHMA investigated possess an initial antiparallel topology (contacts forming a line perpendicular to the main diagonal), which then induces an unzipping-type behavior by reconfiguring to a helical-like topology (contacts residing next and parallel to the main diagonal). This unfolding response via unzipping is enabled by the high entropic conformational flexibility of the backbone as well as that of the side-chains evidenced by the gradual reduction in intramolecular contacts for homopolymers, while the heteropolymers show more varied contact reduction rates throughout the trajectories (Figure 9). Moreover, force peak occurrences appear to correlate with a rapid reduction in intramolecular contacts, with the magnitude of the force peak being dependent on the nature of monomer-monomer interactions disrupted upon unfolding. There is also greater variability in the contact reduction curve behavior between each topological conformation as well as between replicates of the same conformation for the RHPs compared to the PEHMA homopolymer.
**Figure 9.** Backbone topology affects RHP unfolding responses. Unfolding force curves of four different starting conformations, numbered 1–4, of (A) homopolymer PEHMA and (B, C) RHP sequence 15 and 19. Superimposed are contact data (in dark green) showing the reduction in the number of intramolecular contacts during RHP unfolding (one curve for each replicate; see Methods for details). Corresponding contact map evolutions and contact reduction curves in a different presentation style can be found in Figure S19.

For further comparison of chemical and topological contributions to unfolding behavior, we compare the dihedral angle dynamics between the different polymer sequences and between...
conformations. The standard deviation of dihedral angles across the entire unfolding trajectory can be used as a proxy for the extent of backbone reconfiguration, allowing identification of mechanically stable regions (i.e., chain segments that undergo minimal transition in all five unfolding replicates) within a polymer chain (Figure S20A–C). Both PEHMA and RHP sequences showed regions that remained mechanically stable from each initial topology. However, in the RHPs, the segments closest to the chain ends offered a disproportionately consistent opportunity for reconfiguration. Further analysis of the chiral nature of the monomers along the sequence revealed that the emergence of mechanically stable segments correlates with alternating chirality in the residues (Figure S20D). Additionally, both intra- and inter-monomeric dihedrals can mediate conformational changes upon mechanical unfolding with no preference of one over the other (Figure S21). This highlights the capability of atomistic MD simulations to capture the high conformational flexibility of the polymer backbone. Overall, our results demonstrate that, for a multicomponent polymeric system with topological heterogeneity, both the backbone conformation and sequence are important factors in affecting the single-chain unfolding response, with topological effects capable of outweighing chemistry or sequence effects.

The above analysis highlights the importance of topology in affecting the unfolding response, pathway, and dynamics of single-chain heteropolymer systems. Additionally, topological design can be a viable option to tune the mechanical response of a heteropolymer. In practice, both internal and external confinement strategies can be employed to force single-chain heteropolymers into certain topologies. Internal confinement utilizes intramolecular crosslinking and/or orthogonal chemistry. Perez-Baena et al. have used long bifunctional crosslinkers via thiol-yne coupling reaction to promote the formation of long-range loops in compactified SCNPs. External confinement refers to physical confinement of the polymers at the nanoscale, consequently altering
the configurational sampling space of the system. Therefore, synthesis of high concentrations of RHPs within a nanofluidic device with optimized channel geometries may be of interest in future experimental work for directed topological RHP synthesis. Processing conditions can also be coupled to the aforementioned strategies to assist formation of desired topological structures. To gain further understanding of how RHP topological motifs might correlate with force curve features, emerging concepts such as circuit topology and graph theory could be applied. Additionally, a greater variety of RHP sequences with various chemistries and characteristics could provide enough data for a rigorous quantitative model, though computational costs of the pulling simulations remain a hurdle. As demonstrated in one RHP (sequence 19), the realization that spatially-defined anions within a fold can modulate the strength of hydrophobic attraction should also promote new research into this area by combining topological control with sequence-defined polymer synthesis. This would allow us to design in structural (in)stabilities for relevant applications for synthetic heteropolymers or SCNPs self-assembled through hydrophobic collapse.

CONCLUSIONS

In summary, this work investigates the structure and single-molecule mechanics of protein-inspired RHPs in water using all-atom simulations. The four-component heterochiral RHPs sample from a broad statistical distribution of metastable conformations, and so do their properties. As a result, their structure-property landscape proves highly complex, and there is no singly defined response to forced unfolding for the RHPs. Nevertheless, our data suggest that the physicochemical parameters of the RHPs, particularly the backbone topology, can be tuned to enable specific unfolding responses, which may be leveraged to mediate specific interactions with biomacromolecules. As the RHP system presented here also belongs to the active field of SCNPs, our findings on the chemical heterogeneity in polymer design, topological organization of the
polymer backbone, and the importance of side chain length and bulkiness in modulating polymer behavior may open doors for further research in this area. More generally, this work highlights the necessity of atomistic details in elucidating single-molecule mechanics of multicomponent heteropolymers, revealing phenomena (such as the existence of a dynamic multitude of unfolding pathways as well as the OEGMA wrapping as a protection mechanism) that cannot be easily captured by previously proposed theories. Overall, heteropolymer systems with high chemical and conformational heterogeneity, such as the one presented here, necessitate further exploration.

**METHODS**

**Unbiased molecular dynamics (MD) simulation details**

RHP sequences with degrees of polymerizations of 20 and 50 (referred as 20mers and 50mers, respectively, in this work) were simulated by selecting the first 20 and the first 50 residues of 100mer RHPs with target compositions of MMA:OEGMA:EHMA:SPMA in ratios of either 50:25:20:5 or 50:5:30:15 by number generated and parameterized per methods in Hilburg et al.\(^\text{15}\)

Annealing protocol initially minimized and equilibrated at 500 K for 40 ns and then ramped down to 300 K over 40 ns in implicit solvent. This was repeated ten times and each resulting structure at 300 K was extracted for explicit solvation in a periodic octahedral geometry with approximately 40,000 molecules of SPC/E water and potassium counterions (to offset SPMA charges). Each structure was then annealed to 650 K, held for 20 ns, and cooled down to 300 K over 40 ns. The structures were then held at 300 K for 60ns, the latter 40 ns of which were used for analysis. The final frames of these trajectories were then used for unfolding simulations.

**Steered molecular dynamics (SMD) simulations**

SMD simulations were performed on the obtained equilibrated structures using a constant-velocity protocol in order to mechanically unfold the RHPs in explicit water at 300 K. For all RHP
sequences and conformations studied, the two ends are defined to be the backbone $C_\beta$ atom of the first monomeric unit and the backbone $C_\alpha$ atom of the terminal unit (Figure S2). A biasing potential is applied to the two ends with a force constant of 7.0 kcal mol$^{-1}$ Å$^{-2}$, pulling the two ends apart at a constant speed of 1.0 Å ns$^{-1}$. Other parameterizations, including Langevin thermostat and Berendsen barostat, remain identical to those reported in Hilburg et al.$^{15}$ All 20mer RHPs are stretched until the end-to-end distance reaches 50 Å, and all 50mers are stretched until 120 Å, where there are no apparent interactions between non-adjacent monomers. Five independent replicates initiated with new random velocities were performed for all RHP sequences studied. The choices of simulation parameters have been informed by previous works involving the use of SMD (Table S1).

_in silico_ stress-relaxation experiments were performed to gain insights into the dissipation of induced tensile stress from the unfolding process of single-chain RHPs. Snapshots of interest were selected at particular time points from throughout the one-stage pulling trajectories and used to obtain the atomic coordinates. Randomized velocities are used to initiate the stress-relaxation simulations. The restraint on the end-to-end distance in SMD is set to be a constant, equivalent to that at which the snapshot was extracted, and RHPs are allowed to relax at that restrained end-to-end distance for 20 ns. For each snapshot, three independent stress-relaxation experiments are reported.

_analysis_

Cpptraj and Pytraj$^{68}$ are used to analyze all simulation trajectories as per AMBER19 manual$^{69}$, and VMD$^{70}$ (Visual Molecular Dynamics) is used for visualization.

**Force curves and analysis.** For each unfolding experiment using SMD, the force applied is recorded as a function of the polymer’s end-to-end distance. A moving average over 2-Å intervals
is calculated to improve signal-to-noise ratios in the force curves. Peak force analysis is done using the SciPy library in Python to characterize the mechanical stability of RHPs and of their homopolymer counterparts. In particular, a feature with peak value greater than 42 pN and a prominence value greater than 28 pN is considered as a peak. For any given sequence, all peaks pooled from the five independent replicates are included in the peak force distribution analysis. Nonequilibrium unfolding work is calculated as the integral of the unfolding force curve evaluated from the initial end-to-end distance to 45 Å for 20mers and 110 Å for 50mers, respectively.

**Dihedral angles.** A dihedral angle characterizes bond rotations in a polymer and thus its conformational state. As every four neighboring atoms define a dihedral angle, plotting the dihedral angles along the carbon-carbon backbone of a RHP chain gives a 1D topological fingerprint for that RHP. Dihedral angles are averaged over 2-ns of simulation with standard errors computed, and the absolute value is reported to produce dihedral plots.

**Contact analysis.** Contact maps plotting the intramolecular distances of all possible pairs of monomers in a given chain as a two-dimensional matrix have been extensively used in the literature for protein structural analysis. The internal structure of our RHPs and its evolution upon unfolding is used in an analogous fashion. Three types of contact map analyses are performed in this work based on: $C_\alpha$ atoms in the backbone, the center-of-mass of the side-chains, and tail atoms in the side-chains (Figure S2). All contact map data are averaged over 2-ns of simulation trajectories. Contact map data is directly used to determine the number of contacts formed between monomeric residues at a given time (averaged over 2 ns). For 50mers, a contact is considered to be established if the monomer-monomer distance, whether it is $C_\alpha$-based or side-chain-based, is less than 20 Å. Excluding non-contact entries in a contact matrix and avoiding double-counting, the number of contacts at a given time is thus given by $n_{\text{contacts}} = \frac{n_{\text{values below threshold}} - n_{\text{diagonal}}}{2}$, where
$n_{\text{values below threshold}}$ is the number of entries in the contact matrix below the threshold and $n_{\text{diagonal}}$ is the number of diagonal entries, \textit{i.e.}, 50 for 50mers.

**Solvent-accessible surface area.** Solvent accessible surface area (SASA) of an RHP is the surface area that is exposed to water molecules as calculated using the linear combination of pairwise overlaps algorithm (LCPO) as implemented in AMBER19 (ref. 69), and all atom contributions from a given RHP molecule are considered for SASA evaluation.

**Water shell solvation.** Solvation data provides information on the solvent-accessible regions of RHPs upon unfolding. The normalized first solvation shell (NFSS) is computed as the number of water molecules in the first solvation shell relative to that in a well-solvated monomer of the same type per Hilburg \textit{et al.} 15

**Statistical analysis**

ANOVA tests were conducted using Python to ascertain statistical significance (*$p < 0.05$) between sampling distributions.

**ASSOCIATED CONTENT**

**Supporting Information.**

The Supporting Information is available free of charge at xxx.

RHP sequence schematics. Atom designations. Heterogeneous conformational sampling in RHPs. Additional stress-relaxation results. Unfolding force curves for all RHPs studied. Characterization of independent unfolding replicates. Snapshots showing the unfolding pathway of a 50mer RHP. OEGMA wrapping: snapshots, contact maps, and monomer-monomer separational distances. Peak force distributions. Additional analysis on non-equilibrium unfolding
work. Effects of backbone topology: unfolding trajectory snapshots, contact map evolutions, and contact reduction curves. Dihedral dynamics. (PDF)

Movie of a 50mer RHP unfolding trajectory. (MOV)

AUTHOR INFORMATION

Corresponding Author

* Alfredo Alexander-Katz – Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States; Email: aalexand@mit.edu

Authors

Zexiang Han – orcid.org/0000-0002-8968-0860

Shayna L. Hilburg – orcid.org/0000-0003-0825-7698

Funding Sources

This work was supported by the Defense Threat Reduction Agency contract HDTRA11910011.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Z.H. acknowledges the support from the Imperial-MIT International Research Opportunities Program (IROP) Exchange Program. We are thankful to Ting Xu and Zhiyuan Ruan for insightful discussions.
REFERENCES


(22) Brockwell, D. J.; Paci, E.; Zinober, R. C.; Beddard, G. S.; Olmsted, P. D.; Smith, D. A.; Perham, R. N.; Radford, S. E. Pulling Geometry Defines the Mechanical Resistance of a β-Sheet Protein. *Nat. Struct. Biol.* 2003, 10 (9), 731–737. DOI: 10.1038/nsb968


(33) Panganiban, B.; Qiao, B.; Jiang, T.; DelRe, C.; Obadia, M. M.; Nguyen, T. D.; Smith, A. A. A.; Hall, A.; Sit, I.; Crosby, M. G.; Dennis, P. B.; Drockenmuller, E.; De La Cruz, M.


(44) Hyeon, C.; Thirumalai, D. Mechanical Unfolding of RNA: From Hairpins to Structures with


(49) Debenedictis, E. P.; Hamed, E.; Keten, S. Mechanical Reinforcement of Proteins with Polymer Conjugation. *ACS Nano* 2016, 10 (2), 2259–2267. DOI: 10.1021/acsnano.5b06917


Heidari, M.; Heidari, M.; Schiessel, H.; Mashaghi, A. Circuit Topology Analysis of Polymer Folding Reactions. ACS Cent. Sci. 2020, 6 (6), 839–847. DOI: 10.1021/acscentsci.0c00308


Polotsky, A. A.; Charlaganov, M. I.; Leermakers, F. A. M.; Daoud, M.; Borisov, O. V.; Birshtein, T. M. Mechanical Unfolding of a Homopolymer Globule Studied by Self-Consistent Field Modeling. Macromolecules 2009, 42 (14), 5360–5371. DOI: 10.1021/ma9004742


