Biophysical Journal
Article

Biophysical Society

Disulfide Bridges: Bringing Together Frustrated Structure in a Bioactive Peptide

Yi Zhang,[1] Klaus Schulten,[1,2] Martin Gruebele,[1,2,*] Paramjit S. Bansal,[3] David Wilson,[3] and Norelle L. Daly[3,*]
[1]Beckman Institute for Advanced Science and Technology, Center for Biophysics and Quantitative Biology; [2]Department of Physics and Department of Chemistry, University of Illinois at Urbana-Champaign, Champaign, Illinois; and [3]Centre for Biodiscovery and Molecular Development of Therapeutics, Australian Institute of Tropical Health and Medicine, James Cook University, Cairns, Queensland, Australia

ABSTRACT Disulfide bridges are commonly found covalent bonds that are usually believed to maintain structural stability of proteins. Here, we investigate the influence of disulfide bridges on protein dynamics through molecular dynamics simulations on the cysteine-rich trypsin inhibitor MCoTI-II with three disulfide bridges. Correlation analysis of the reduced cyclic peptide shows that two of the three disulfide distances (Cys^{11}-Cys^{23} and Cys^{17}-Cys^{29}) are anticorrelated within ~1 μs of bridge formation or dissolution: when the peptide is in nativelike structures and one of the distances shortens to allow bond formation, the other tends to lengthen. Simulations over longer timescales, when the denatured state is less structured, do not show the anticorrelation. We propose that the native state contains structural elements that frustrate one another's folding, and that the two bridges are critical for snapping the frustrated native structure into place. In contrast, the Cys^{4}-Cys^{21} bridge is predicted to form together with either of the other two bridges. Indeed, experimental chromatography and nuclear magnetic resonance data show that an engineered peptide with the Cys^{4}-Cys^{21} bridge deleted can still fold into its near-native structure even in its noncyclic form, confirming the lesser role of the Cys^{4}-Cys^{21} bridge. The results highlight the importance of disulfide bridges in a small bioactive peptide to bring together frustrated structure in addition to maintaining protein structural stability.

INTRODUCTION

The concept of minimal frustration is a central idea in the energy landscape theory of protein folding (1): The native arrangement of the polypeptide chain is favored (minimally frustrated) because different elements of the native structure are compatible with one another. Nevertheless, some proteins contain elements that frustrate folding. For example, the WW domain contains a long loop needed for function (2), but the loop slows down folding. As another example, protein Im7 eventually forms native structure, but not before being trapped by some nonnative contacts (3).

Proteins contain a number of interesting structural elements that are not obligatory for the folding of all proteins: salt-bridges, proline kinks, and disulfide bridges, to name a few examples (4). In particular, many small bioactive peptides lack a large hydrophobic core to provide sufficient stability, and a cross link such as a disulfide bridge greatly enhances their stability.

The question then arises to what extent the disulfide bridge merely strengthens structural propensities already encoded in the polypeptide's amino acid sequence, and to what extent it enables frustrated structure to form. For example, two cysteines in a reduced peptide may readily sample short Cys-Cys distances, allowing the disulfide bridge to form quite easily; or Cys-Cys interaction may be rare events high in free energy, in which case the bridge holds together a strained structure.

Here we answer this question by simulation and experiments on MCoTI-II (5), a 34-residue trypsin inhibitor from *Momordica cochinchinensis*, the plant that produces the Vietnamese "gac" fruit. This peptide is a member of the cyclotide family, which is characterized by a cyclic cystine knot (CCK) structural motif. The CCK motif in MCoTI-II involves three disulfide bridges connecting positions 4–21, 11–23, and 17–29, coupled with a cyclic backbone (see Fig. 1). The CCK structure makes the cyclotides sufficiently stable for potential pharmaceutical applications (6), despite their small size. In vivo, oxidation of the cysteine residues to form the cystine knot presumably facilitates subsequent enzymatic cyclization.

We studied cyclic MCoTI-II by all-atom molecular dynamics (MD) simulation, and looked at the chromatography and nuclear magnetic resonance (NMR) spectra of a linear mutant to test a prediction from the simulation. The answer

Submitted October 5, 2015, and accepted for publication March 21, 2016.
*Correspondence: mgruebel@illinois.edu or norelle.daly@jcu.edu.au
Editor: Elizabeth Komives.
http://dx.doi.org/10.1016/j.bpj.2016.03.027

to our question about frustration, as is often the case for complex biological systems, is "all of the above". We performed 22 μs of all-atom MD simulations of the solvated peptide, starting with either folded or unfolded structures, and with various combinations of disulfide bridges intact. The picture that emerges is that two of the disulfide bridges (11–23, 17–29) are critical for snapping the native structure into place. The 11–23 and 17–29 distances are anticorrelated in the reduced peptide when it reaches more nativelike conformations, so the native state is somewhat frustrated. The anticorrelation disappears when the peptide is simulated over longer periods and the denatured state loses more structure. The third 4–21 distance can correlate with one or the other of the two key bridges. It neither frustrates folding, nor is it the key for maintaining the frustrated native fold.

To test this idea further, we performed an experiment where we synthesized a linear peptide with the presumably noncritical 4–21 bridge deleted. Chromatography and NMR data show that although the peptide is somewhat more flexible than the full linear native fold, deleting the 4–21 bridge indeed produces a peptide with near-native structure, even without cyclization. This result highlights that disulfide bridges do not just help stabilize the native structure of a small bioactive peptide, but also help shoehorn it into a somewhat frustrated native fold.

MATERIALS AND METHODS

MD simulations

The initial structures in all MD simulations were based on the solution NMR structure of MCoTI-II (PDB: 1IB9) (5). The *psfgen* plugin in VMD (7) was used to prepare modified structures, each containing a different combination of disulfide bridge. The prepared structures were solvated using the TIP3P water model (8) and 0.15 M NaCl. Each prepared system contained ~32,000 atoms.

The MD simulations presented in this study used the CHARMM36 force field with CMAP corrections for proteins and ions (9–11). Eight systems, listed in Table 1 were simulated with a total time of 22 μs. The minimization step and the first 1-μs production run for all systems except for SS0(c) were carried out with NAMD2 (12), explained in detail in the rest of the paragraph. Each system was minimized for 5000 steps and equilibrated for 6 ns while applying harmonic restraints on the heavy atoms of the protein (k =1 kcal/(mol Å^2)). In the next step, systems were simulated for several nanoseconds in the NPT ensemble conditions at 1.0 atm and 310 K, until the volume of each system stopped changing. Then, the simulations were continued as an NVT ensemble. As indicated in Table 1, some simulations required an additional heating step to denature the native protein structure, which was realized by raising the temperature to 400 K in the NPT ensemble conditions, while holding pressure constant. In all simulations, periodic boundary conditions were applied. Temperature was controlled by Langevin dynamics with a damping coefficient of 1 ps^{-1}; pressure was controlled by the Nosé-Hoover Langevin piston method (13,14). The cutoff for short-range nonbonded interactions was 12 Å, and the particle-mesh Ewald method was employed to calculate long-range electrostatic forces (15). The SHAKE algorithm was used to constrain bond distances involving hydrogen atoms (16), so a 2-femtosecond integrations step could be chosen. Figure rendering and data analysis were performed in VMD (7).

To obtain more statistics, SS0u(c) and additional production runs beyond 1 μs on other systems were carried out on the Anton platform (17). The *multigrator* integration method is applied with a time step of 2 fs (18). Short-range forces were evaluated every time step with a cutoff of 13.88 Å; long-range electrostatics was calculated every three time steps using the Gaussian split Ewald method with a 32 × 32 × 32 grid (19).

The correlations of sulfur-sulfur bond distances x and y corresponding to a pair of disulfide bonds seen in the native structure were quantified through the Pearson correlation coefficient, ρ_{xy}, defined as

$$\rho_{xy} = \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\overline{y})^2}} \quad (1)$$

where n is the number of frames in simulation analyzed; $\overline{x}$, $\overline{y}$ are the average bond distances in that simulation; and x_i and y_i are the actual sulfur-sulfur distances in frame i of the simulation. The value $r_{xy} = 1$ indicates a total positive correlation and $r_{xy} = -1$ indicates a total negative correlation between the distances in the two sulfur-sulfur pairs compared.

Peptide synthesis

MCoTI-II(C4A) was synthesized by manual solid-phase peptide synthesis using Fmoc chemistry on a 2-chlorotritylchloride resin, with HBTU as the amino-acid coupling agent. The peptide was synthesized on a 0.5-mmol scale and cleaved using a mixture of 95% TFA (trifluoroacetic acid)/2.5% triisopropylsilane/2.5% H_2O. After removal of TFA, the peptide was precipitated with diethyl ether and extracted in 50% acetonitrile/50% H_2O. The aqueous layer was lyophilized and the resulting crude peptide purified using reversed-phase high-performance liquid chromatography (RP-HPLC) on a model No. 1260 Infinity system (Agilent Technologies, Santa Clara, CA) and characterized using mass spectrometry on a matrix-assisted laser desorption/ionization time-of-flight (TOF)/TOF 5800 (SCIEX, Framingham, MA). The purified, reduced peptide was oxidized in 0.1 M ammonium bicarbonate, pH 8.2. Aliquots were removed at time points between 1 and 30 min and quenched with 10% TFA. The aliquots were analyzed using analytical scale RP-HPLC and products characterized with mass spectrometry. MCoTI-II(C17A,C29A) was synthesized using similar procedures to MCoTI-II(C4A) but with selective protection of cysteine residues. Cys^4 and Cys^{21} were protected with *S*-acetamidomethyl groups with deprotection and oxidation carried out using the procedure given in Hunter and Komives (20).

NMR spectroscopy

NMR spectra were recorded on Avance III 600-MHz or 700-MHz NMR spectrometers (Bruker, Billerica, MA). Peptides were dissolved in 90% H_2O and 10% D_2O, and total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY) spectra were recorded at 298 K. For TOCSY spectra, an 80-ms mixing time, and for NOESY spectra a 200-ms mixing time were used. Two-dimensional spectra were collected over 4096 data points in the f2 dimension and 512 increments in the f1 dimension over a spectral width of 12 ppm.

RESULTS

Dynamics of initially folded MCoTI-II with and without disulfide bridges

In this study, our goal was to understand if the MCoTI-II peptide has an inherent dynamical structural propensity to bring the correct cysteines together for linkage, or if the disulfides stabilize frustrated structure that the reduced

polypeptide chain does not prefer. In other words: which of the three cysteine pairs find each other efficiently when disulfide bonds are not linked, and which do not?

Our first set of simulations started in the folded state. Equilibrium simulations of cyclic MCoTI-II with all three disulfides intact (SS123f) and with none intact (SS0f(a,b)) were initialized in the conformation of the folded NMR structure shown in Fig. 1 *A*, and were run for microseconds as listed in Table 1. The root-mean-square deviation (RMSD) of the MCoTI-II peptide starting from the folded structure is shown in Fig. 2. SS123f maintains nativelike structure throughout (RMSD $\approx$ 3 Å). The RMSDs of MCoTI-II without any disulfide bonds are initially slightly higher than the one with intact disulfides (~3–6 Å). After ~0.6 μs, SS0f(a) unfolds significantly to ~9 Å RMSD. This result shows that unfolding can occur in less than a microsecond when none of the disulfide bonds is formed, whereas the fully oxidized peptide remains folded over the same time period.

Distances between native Cys pairs in reduced (unlinked) SS0f(a) are shown in Fig. 3 *A*, with two notable features. First, at the beginning of the simulation, the distance between the Cys^{4} and Cys^{21} pair increases most readily, while the distance of the $Cys^{11\text{-}23}$ and $Cys^{17\text{-}29}$ pairs remains more nativelike. Second, a clear anticorrelation of sulfur-sulfur distances is observed between the 11–23 and 17–29 pairs starting at 0.6 μs in the SS0f(a) simulation, whereas the 4–21 and 17–29 pairs remain clearly correlated during the same time period. We will see this pattern of anticorrelated and correlated sulfur-sulfur distances in further simulations discussed below, indicating that bridges 11–23 and 17–29 frustrate one another when the peptide structure is not too far from the native state, while bridge 4–21 can go along with either of the two other bridges.

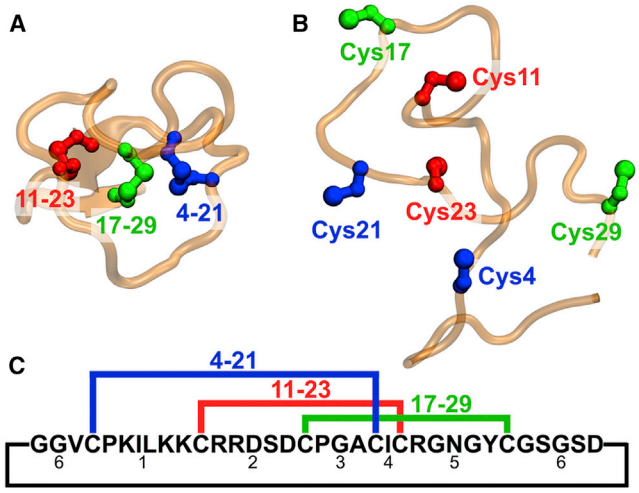

FIGURE 1 Structure and sequence of MCoTI-II. (*A*) Native structure solved by solution NMR (PDB: 1IB9) (5). Three disulfide bonds are labeled based on their Cys numbering. (*B*) MCoTI-II without disulfide bonds formed. Shown is just one of the denatured peptide structures after 0.05 μs of heating at 400 K. Colors as in (*A*). (*C*) Sequence of the cyclic peptide MCoTI-II with native cysteine matches labeled above, and intercysteine loops labeled below. Loop 1 contains the binding site for trypsin. To see this figure in color, go online.

TABLE 1 Summary of Simulations Performed

Simulation	Description	Time per Run (μs)	Number of Runs
SS0f(a,b)	equilibration of MCoTI-II without any disulfide bridges	3.5 (3)	2
SS0u(a,b,c)	$\approx$0.5 μs of heating was applied before equilibration of MCoTI-II without any disulfide bridges	3.5	3
SS123f	equilibration of MCoTI-II with all three disulfide bridges linked	1	1
SS2f	equilibration of MCoTI-II with 11–23 linked	0.8	1
SS2u	intermediate state selected from SS0u, followed by equilibration with 11–23 linked	0.8	1
SS3f	equilibration of MCoTI-II with 17–29 linked	0.8	1
SS3u	intermediate state selected from SS0u, followed by equilibration with 17–29 linked	0.8	1
SS23f	equilibration of MCoTI-II with 11–23 and 17–29 linked	0.8	1

Dynamics of MCoTI-II after high temperature denaturation

To explore further if the unfolded peptide has structural propensity to bring native cysteine matches together, we examined MCoTI-II in simulations that start with unfolded peptide. Three simulations of this type for a peptide without any disulfide linkages were carried out, and abbreviated as SS0u(a), SS0u(b), and SS0u(c). In each simulation, the peptide was heated to 400 K for ~0.05 μs to yield three independent starting structures, then simulated at 310 K for the rest of the simulation. One representative denatured structure of MCoTI-II after heating is shown in Fig. 1 *B*. All the secondary structure has disappeared, and the overall structure is much less compact and rigid than its folded counterpart in Fig. 1 *A*.

In Fig. 2, the overall RMSD of simulations SS0u(a,b,c) increased well above the native value during the first 0.05 μs (T = 400 K). When the temperature was brought back to 310 K, the RMSD fluctuated around 9 Å, the same value that SS0f(a) reached after 0.64 μs. In microsecond-long simulations at 310 K, the denatured peptides were not able to recover nativelike structure.

Correlations among the three disulfide distances are revealed for SS0f(a) and SS0u(a) in Fig. 3 (see Fig. S2 in the Supporting Material for SS0f(b), SS0u(b), and SS0u(c)). In both simulations, a clear anticorrelation between distances 11–23 and 7–29 was observed a short time after denaturation as monitored by RMSD. To quantify these observations, the Pearson correlation coefficient in Eq. 1 was computed in a

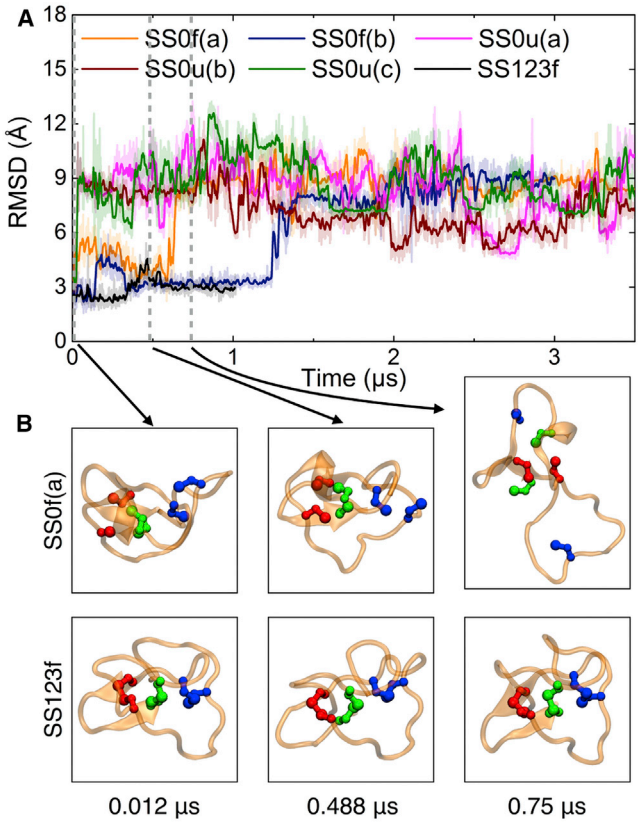

FIGURE 2 Structural dynamics of MCoTI-II. (*A*) RMSDs measured for Cα atoms for structures of MCoTI-II as a function of time under different conditions; RMSD values were determined all relative to the native structure (PDB: 1IB9) (5). Simulations involving initial simulated heating (SS0u(a), SS0u(b), and SS0u(c)) show a large increase in RMSD values immediately, while simulations without heating (SS0f(a) and SS0f(b)) can reach a high final RMSD value of ~9 Å as well. (*B*) Comparison of MCoTI-II structures from simulations SS0f(a) and SS123f, at times pointed to by arrows in (*A*). Colors as in Fig. 1. To see this figure in color, go online.

sliding time window for distances 11–23 and 17–29 and plotted in Fig. 3, *B* and *D*. In the case of SS0f(a) shown in Fig. 3, *A* and *B*, the structure unfolded naturally ~0.6 μs, where distance 11–23 started to anticorrelate with distance 17–29. The strong anticorrelation remained for 0.3 μs and came back and forth throughout the simulation. A similar scenario is observed for SS0u(a) shown in Fig. 3, *C*–*E*, where a heating step was applied in the simulation to accelerate unfolding. In that simulation, no more anticorrelation was observed after 1.5 μs when the peptide got trapped in less nativelike structures. A sampling of both (correlated and anticorrelated) structures is shown in Fig. S4. Meanwhile, distance 4–21 was occasionally correlated with distance 11–23 or distance 17–29 throughout both simulations. Additional correlational coefficients among the three disulfide distances right after the initial unfolding transitions are listed in Table 2, where 11–23 and 17–29 are highly anticorrelated ($\rho \sim -0.92$), whereas 4–21 can go along with either ($\rho \sim 0$ on average, with range ~−0.8 to 0.9 in different simulations).

The value $\rho \approx 1$ means that two processes are locked together ($\Delta\Delta G$, the effective free energy coupling the two processes, is <0). The value $\rho \approx 0$ means they are unrelated ($\Delta\Delta G \approx 0$), and $\rho \approx -1$ means that they are mutually exclusive ($\Delta\Delta G >> 0$). The observation that 11–23 and 17–29 are anticorrelated means that the formation of one disulfide bridge makes the other one harder to form. These two bridges stabilize native substructures that frustrate each other. The remaining 4–21 distance is flexible in its structural propensity, and was also by far the easiest to unravel starting with the native peptide (Figs. 2 *A* and S2 *A*). It is worth mentioning that the frustration is not severe: deviations from −1 of correlation coefficients in Table 2 by 2–17% indicate that the effective free energy cost $\Delta\Delta G \equiv -kT\ \ln[(1+\rho)/(1-\rho)]$ of the frustration is <2.5–5 *kT*. The native state of MCoTI-II is still rather stable, and can fold in minutes in the presence of the 11–23/17–29 frustration: assuming a speed-limit prefactor $k_m \approx (1\ \mu s)^{-1}$ in the Arrhenius expression $k = k_m \times \exp[-\Delta G^{\dagger}/kT]$ for the unimolecular folding rate, $\Delta G^{\dagger} \approx kT \ln[60\ s/10^{-6}\ s] \approx 17\ kT$, 3–7 times larger than the frustration between 11–23 and 17–29 that we observe.

Dynamics of MCoTI-II with partially linked disulfide bond matches

To further confirm frustration between disulfide bond pairs, and to predict the likelihood of disulfide bond formation within the MCoTI-II structure, we simulated the MCoTI-II structure with partially linked bridges. SS23f is a simulation under native conditions with disulfide bridges 11–23 and 17–29 intact, SS2u is a simulation under heat-denatured conditions with only bridge 11–23 intact, and so forth, for other cases listed in Table 1. (The "u" simulations started with the structure at ~0.36 μs in Fig. 3 *C*, where all three sulfur-sulfur distances were approximately equal.)

RMSD analysis (Fig. S1) and distance plots (Fig. S3) support that the overall structure is less flexible than the fully reduced peptide when one of the disulfide bonds is formed. Simulations that started from the native state conformation also confirmed that 4–21 is the least rigid match: these simulations showed that the 4–21 sulfur-sulfur distance always fluctuates a lot, regardless of whether 11–23 only, 17–29 only, or both, are bridged.

A prediction based on simulation

The picture that emerges is then the following: the 4–21 sulfur-sulfur distance fluctuates most easily, fluctuating significantly even in the native state, whereas sulfur-sulfur distances 11–23 and 17–29 tend to remain locked-in to their native state values. However, once the peptide has unfolded, the distances of pairs 11–23 and 17–29 becomes anticorrelated, and thus a frustrated native structure is formed when these two form disulfide bridges. Therefore, formation of 11–23 and 17–29 is critical

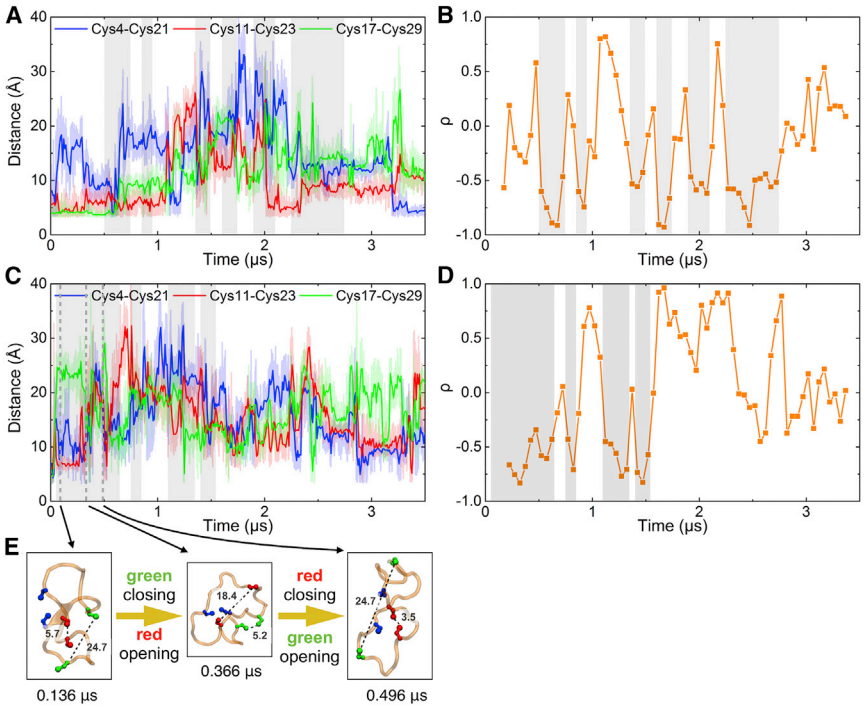

FIGURE 3 Sulfur-sulfur distances and correlations monitored in simulations of different conditions. (*A*) Sulfur-sulfur distances monitored in simulation SS0f(a). Average distances taken over 0.01 μs windows are shown as bold traces; light traces show raw distance data. Gray shading highlights the time when Cys11-Cys23 and Cys17-Cys29 distances are anticorrelated. (*B*) Pearson correlation coefficient between Cys11-Cys23 and Cys17-Cys29 distances in simulation SS0f(a) from (*A*). A gliding window of size 0.3 μs and step 0.05 μs is applied to monitor correlation as a function of time. The same gray shading as in (*A*) cover the time when large negative correlation coefficients are present. (*C* and *D*) Same as (*A*) and (*B*) for simulation SS0u(a). Other cases are shown in the Supporting Material. At longer times in the simulation, when the peptide is trapped in less nativelike microstates, the anticorrelation disappears (see also Fig. S4). (*E*) Comparison of structures from simulation SS0u(a) in (*C*), where an anticorrelation between sulfur-sulfur distances 11–23 (red) and 17–29 (green) is observed. To see this figure in color, go online.

for snapping MCoTI-II into its native structure, whereas the labile 4–21 goes along for the ride. We thus predict that deleting the 4–21 bridge would still allow the peptide to form the native structure, albeit a native structure that is more flexible at the C-terminus. On the other hand, deleting one of the two critical disulfide bridges would yield a peptide with impaired folding, even if 4–21 is able to form.

Experimental confirmation of a simulated propensity

To test the picture that formation of disulfide bonds between 11–23 and 17–29 is critical for native structure and that these two disulfide linkages frustrate one another, while the third is less important for maintaining native structure, we synthesized a linear version of the cyclotide MCoTI-II (21) using Fmoc chemistry and analyzed its oxidative folding. The N- and C-termini were chosen to be equivalent to the gene sequence (22), which corresponded to an N-terminal glycine and a C-terminal aspartic acid; the residue numbering is also based on the gene sequence.

Further analysis of the role of the 4–21 disulfide bond was achieved by synthesizing a peptide with Cys^4 mutated to an alanine to prevent formation of the Cys^4-Cys^{21} bond, and allow analysis of favored disulfide isomers before formation of the fully folded peptide. The reduced, purified form of the mutant MCoTI-II(C4A) was incubated in ammonium bicarbonate at pH 8.2 to initiate formation of the disulfide bonds. These conditions were used to determine the most stable disulfide connectivity in the absence of shuffling reagents. The oxidation reaction was analyzed using RP-HPLC and mass spectrometry. An oxidized product with two disulfide bonds was formed within 1 min of dissolving the reduced peptide in buffer, as shown in Fig. 4. This peptide (designated IIa) eluted earlier than the reduced form, consistent with the folding of the native peptide (23). Longer incubations resulted in a decrease in the reduced peptide and an increase in the IIa intermediate, but no other disulfide isomers accumulated to a significant extent.

A larger-scale oxidation reaction was carried out to allow purification of sufficient quantities of the IIa peptide for analysis with NMR spectroscopy. Chemical shift assignments were made using TOCSY and NOESY spectra. Comparison of the αH chemical shifts is a powerful method for analyzing structural changes between related peptides. The similarity between the αH shifts of the wild-type MCoTI-II and IIa indicates that the overall fold is maintained and consequently the native disulfide connectivities are formed (i.e., Cys^{11}–Cys^{23} and Cys^{17}–Cys^{29}) as shown in Fig. 5. The NMR spectra did not change over several days, highlighting the stability of the peptide. Specifically, it appears that the free thiol from Cys^{21} does not induce any disulfide

TABLE 2 Correlation Coefficients ρ among Three Disulfide Distances Measured over 0.3 μs Immediately after the Initial Unfolding Transitions

Simulation	4–21 and 11–23	11–23 and 17–29	4–21 and 17–29
SS0f(a)	−0.83	−0.91	0.93
SS0f(b)	0.58	−0.98	−0.49
SS0u(a)	0.82	−0.83	−0.77
SS0u(b)	0.16	−0.98	−0.11
SS0u(c)	−0.80	−0.89	0.74

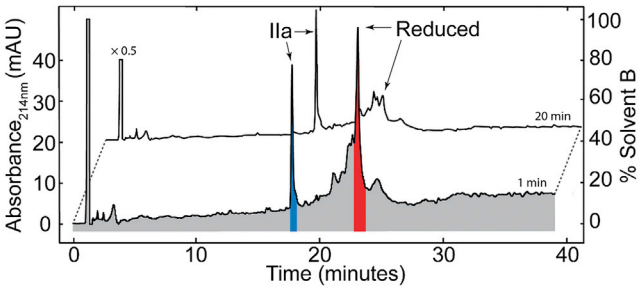

FIGURE 4 RP-HPLC of oxidative folding of linear MCoTI-II(C4A). Reduced peptide was dissolved in 0.1 M ammonium bicarbonate (pH 8.2) and quenched after 1 min with 10% TFA. A well-resolved peak, corresponding to a peptide with two disulfide bonds formed (IIa), eluted earlier than the reduced peptide. Time points taken at 1 and 20 min are shown. The axes are for the 1-min data. The 20-min data has been shifted and scaled by 0.5 for clarity. To see this figure in color, go online.

bond shuffling or formation of an intermolecular disulfide bond. Thus, Cys^4 is not critical for native structure formation.

As an additional control, a peptide with Cys^{17} and Cys^{29} mutated to alanine residues was synthesized with selective protection of the cysteine residues used to form disulfide bonds between 11–23 and 4–21. Analysis of the NMR spectra of this peptide indicated significant overlap in the amide region, which prevented assignment of the individual resonances. The NMR spectra of native MCoTI-II is characterized by significant dispersion in the amide region as a result of the β-sheet structure (5). The lack of dispersion in MCoTI-II(C17A, C29A) suggests that the native fold is not present. These results clearly show that disulfide bonds between 11–23 and 4–21 alone are insufficient to allow the native structure to form, and are consistent with the hypothesis that 11–23 and 17–29 are both critical for native structure. As expected, we have confirmed that the peptide without disulfide bonds also has no nativelike structure based on NMR analysis.

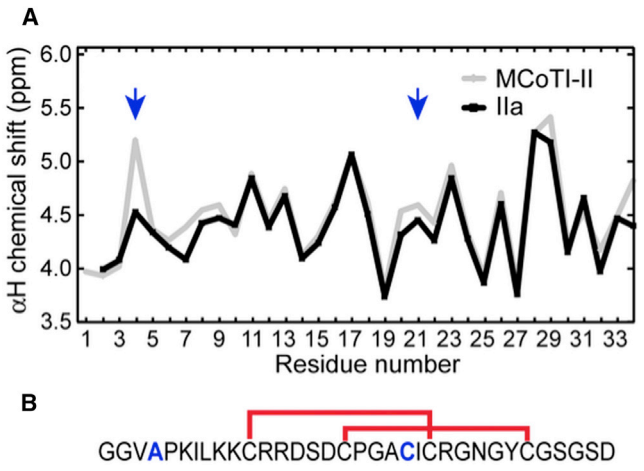

FIGURE 5 NMR analysis of MCoTI-II(C4A). (*A*) Compared are the αH chemical shifts of the oxidized form of MCoTI-II(C4A) containing two disulfide bonds (IIa) with the αH shifts of wild-type MCoTI-II. The similarity in the chemical shifts shows that the overall fold is maintained in IIa. The native disulfide bonds likely to be present in IIa are shown in (*B*). The cysteine residue mutated to an alanine, and the free cysteine are marked in (*A*) with arrows. To see this figure in color, go online.

DISCUSSION

Elucidating the folding pathways in cystine knot peptides is challenging both computationally and experimentally, given their tightly folded structures. In particular, the role of individual disulfide bonds in stabilizing the overall structure is poorly understood.

Using computer simulations, we have found that native disulfide combinations can make frequent contacts with different priority, regardless of starting (native or denatured) structural conformations. We showed that two of the three-disulfide bonds in MCoTI-II are anticorrelated when the peptide samples denatured structures populated immediately after the native state unfolds, whereas the third bond is not critical in forming the native structure. The concept of minimal frustration emphasizes that native contacts of protein structure are critical in the pathway of protein folding (1) (also examined with simulations (24)). Our results show that "minimal" is not "none" in this case, but that disulfides can hold together a somewhat-frustrated native structure in a small peptide.

In our simulations, the anticorrelation between distances 11–23 and 17–29 does not hold after the peptide makes structural transitions to more denatured states. While disulfide bonds are the key to holding the small native state in place, we hypothesize that frustration between native disulfide distances disappears when the peptide becomes more unfolded. In the limit of highly unfolded structures, any collapse of the structure to a more compact state will lead to shortened disulfide distances, which are then correlated. We represent this idea schematically in Fig. 6. If the structure unfolds to a great extent, e.g., displaying a close

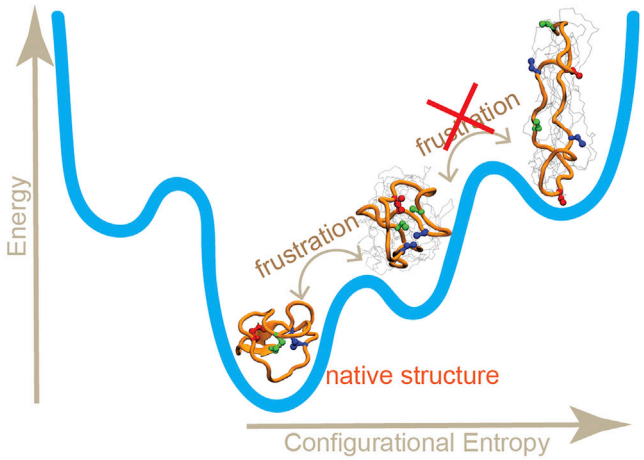

FIGURE 6 A simple scheme of the frustration in MCoTI-II represented in the folding funnel. To see this figure in color, go online.

to random coil of amino acids, the cysteines hardly play a role in the beginning of refolding. As the structure approaches the native state, some frustration appears. As mentioned in Results, the evidence from Table 2 is that the two cysteine bridges frustrate folding at most by a few *kT*, whereas the total folding free energy barrier is on the order of 17 *kT*. We thus predict that if mutations could reduce the frustration between bridges 11–23 and 17–29, the peptide might be able to fold in a few seconds rather than a minute.

Our experiments confirmed that the third disulfide bond (between residues 4 and 21) is not critical for forming the native structure and that the two disulfide bonds (between residues 11–23 and 17–29) that stabilize the frustrated structure form remarkably quickly. The native disulfide bonds in the IIa isomer of MCoTI-II(C4A) were confirmed by NMR spectroscopy and are consistent with the simulations. The stability associated with the native structure is presumably the reason why the IIa isomer accumulates in the oxidative folding and forms so quickly. Interestingly, the free thiol from Cys21 did not interfere with the formation and/or stability of this isomer. Isomers with single disulfide bonds did not accumulate to a significant extent and, therefore, could not be isolated for characterization. This lack of accumulation of one-disulfide species is consistent with these isomers not containing native structure and consequently, being unstable.

Our folding experiments have been done on a linear version of MCoTI-II because it is thought that in vivo the cystine knot is likely to form before cyclization (25). Previous studies have shown that a disulfide isomer equivalent to IIa has been observed in folding studies starting with a reduced, cyclic form of MCoTI-II (23,26), indicating that the IIa isomer is involved in the folding of both the linear and cyclic forms. A topologically equivalent isomer is also present in folding studies on the cyclic form of the related cyclotide, kalata B1 (23,26,27). The cyclic two-disulfide isomer of MCoTI-II appears to be the immediate precursor to the native fold, in contrast to the kalata B1 two-disulfide isomer. Another difference between the folding of MCoTI-II and kalata B1 is that cyclization has been shown to facilitate the folding of kalata B1 in vitro (28), but we have shown that even without the cyclic backbone MCoTI-II can efficiently form the disulfide bonds between 11–23 and 17–29. The primary sequences differ significantly between MCoTI-II and kalata B1, despite both containing the CCK fold, and these sequence differences are likely to be responsible for the differences in the disulfide folding pathways.

MCoTI-II has previously been shown to have greater flexibility in the loop formed upon cyclization (i.e., loop 6) and loop 1 (Fig. 1 *c*), compared to the other parts of the molecule, which are quite rigid (29). Larger fluctuations in distance were observed between 4 and 21 than those between the other two native disulfide matches in our MD simulations (Fig. S2, *A* and *B*). As Cys^4 is neighboring to loop 1 in the sequence, the large fluctuation in 4–21 distance agrees with the previous observation of high structural mobility in this loop 1 in both simulations and experiments on MCoTI-II (26,29). Interestingly, most cyclotides have a highly conserved loop 1 sequence, with the exception of the trypsin inhibitor cyclotides (including MCoTI-II), which have a unique sequence that serves as a binding site for trypsin (30). It is interesting to speculate that a tradeoff might be made in case of trypsin, such that the evolved sequence of loop 1 loses the stability of the native disulfide bond involving Cys4 while acquiring new functionality.

Cyclotides have attracted much attention as novel pharmaceuticals because of their extreme stability and various biological activities (6). In particular, the MCoTI cyclotides are of critical importance to potential pharmaceutical application as they are able to penetrate cell membranes and in some cases target intracellular receptors (31–33). Developing efficient methods for folding cyclotides is a critical step because of their roles as drug leads. Although our results unveil the importance of disulfide bridges in MCoTI-II to bring together frustrated structure elements, how the disulfide bridges affect real-time folding of MCoTI-II is yet to be determined. Qin et al. (34) have recently developed a computational method that enables the simulation of the disulfide bond formation. It is expected that a real-time folding process of MCoTI-II with the formation of disulfide bridges as considered here will further reveal the critical role of disulfide bridges in protein folding.

CONCLUSIONS

In this study, we found that the MCoTI-II peptide has an inherent dynamical propensity for correct disulfide sites coming together for eventual bond formation. From correlation analysis we found that frustration arises in a partly unfolded cyclic peptide between the Cys^{11}-Cys^{23} and Cys^{17}-Cys^{29} pairs, when both shorten to form the respective disulfide bonds: the distances do not tend to shorten to a nativelike value simultaneously. The observed frustration indicates that both disulfide bonds need to be formed to snap the peptide into the native structure. In contrast, the Cys^4-Cys^{21} bond is not important for structural integrity, and it can form with either one of the other two disulfide bonds. More unfolded structures do not show evidence for frustration, so frustration appears only when the peptide is conformationally near the native state. The simulation results were further supported by experiment: when Cys^4 is mutated and 4–21 cannot form a bridge, a stable MCoTI-II form that is similar to the native form is observed in NMR spectroscopy. Thus, disulfide bonds can do even more than stabilize a native structure in the absence of a large hydrophobic core: they can hold together a native structure that is frustrated—although not all disulfides in the peptide are required, to achieve this.

In closing, we offer the hypothesis that frustrated structure will be more common in small peptides than in large proteins. Large proteins can have long chains connecting two components of an active site, capable of positioning both components accurately even though the spatial resolution of a single amino acid is only ~0.3 nm. The shorter chains in a small peptide do not have enough slack for such accurate positioning of two components, but small amounts of frustration can be overcome by covalent disulfide bond linkage.

SUPPORTING MATERIAL

Four figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(16)30108-4.

AUTHOR CONTRIBUTIONS

M.G., N.L.D., and K.S. designed the research; Y.Z., P.S.B., and D.W. carried out the research; Y.Z., P.S.B., and M.G. analyzed the data; and M.G., N.L.D., and Y.Z. wrote the article.

ACKNOWLEDGMENTS

The Anton machine at PSC was generously made available by D. E. Shaw Research. The 700-MHz NMR data were recorded at the Centre for Advanced Imaging (located at The University of Queensland). We thank Dr. Hu Qiu, Dr. Lela Vukovic, and Mr. Wen Ma for their comments on the article.

N.L.D. is an Australian Research Council Future Fellow. The work was supported by National Institutes of Health grants No. 9P41GM104601 (to K.S., Y.Z.) and No. 2R01-GM093318 (to M.G.). Anton time was provided by the National Center for Multiscale Modeling of Biological Systems through grant No. PSCA15066P from the National Institutes of Health and the Pittsburgh Supercomputing Center.

REFERENCES

1. Bryngelson, J. D., J. N. Onuchic, …, P. G. Wolynes. 1995. Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins.* 21:167–195.
2. Jäger, M., Y. Zhang, …, J. W. Kelly. 2006. Structure-function-folding relationship in a WW domain. *Proc. Natl. Acad. Sci. USA.* 103:10648–10653.
3. Capaldi, A. P., C. Kleanthous, and S. E. Radford. 2002. Im7 folding mechanism: misfolding on a path to the native state. *Nat. Struct. Biol.* 9:209–216.
4. Creighton, T. E. 1992. Protein Folding. W.H. Freeman, New York.
5. Felizmenio-Quimio, M. E., N. L. Daly, and D. J. Craik. 2001. Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from *Momordica cochinchinensis*. *J. Biol. Chem.* 276:22875–22882.
6. Gould, A., Y. Ji, …, J. A. Camarero. 2011. Cyclotides, a novel ultrastable polypeptide scaffold for drug discovery. *Curr. Pharm. Des.* 17:4294–4307.
7. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28.
8. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
9. MacKerell, A. D., Jr., M. Feig, and C. L. Brooks, 3rd. 2004. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* 25:1400–1415.
10. MacKerell, A. D., D. Bashford, …, M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
11. MacKerell, A. D., Jr. 2004. Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.* 25:1584–1604.
12. Phillips, J. C., R. Braun, …, K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.
13. Feller, S. E., Y. Zhang, …, B. R. Brooks. 1995. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* 103:4613–4621.
14. Martyna, G. J., D. J. Tobias, and M. L. Klein. 1994. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* 101:4177–4189.
15. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.
16. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of Cartesian equations of motion of a system with constraints—molecular dynamics of N-alkanes. *J. Comput. Phys.* 23:327–341.
17. Shaw, D. E., K. J. Bowers, …, B. Towles 2009. Millisecond-scale molecular dynamics simulations on Anton. *In* Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis, SC 2009. Article No. 39. Association for Computing Machinery (ACM) Digital Library, New York. http://dx.doi.org/10.1145/1654059.1654099.
18. Lippert, R. A., C. Predescu, …, D. E. Shaw. 2013. Accurate and efficient integration for molecular dynamics simulations at constant temperature and pressure. *J. Chem. Phys.* 139:164106.
19. Shan, Y., J. L. Klepeis, …, D. E. Shaw. 2005. Gaussian split Ewald: a fast Ewald mesh method for molecular simulation. *J. Chem. Phys.* 122:54101.
20. Hunter, M. J., and E. A. Komives. 1995. Deprotection of *S*-acetamidomethyl cysteine-containing peptides by silver trifluoromethanesulfonate avoids the oxidation of methionines. *Anal. Biochem.* 228:173–177.
21. Hernandez, J. F., J. Gagnon, …, D. Le Nguyen. 2000. Squash trypsin inhibitors from *Momordica cochinchinensis* exhibit an atypical macrocyclic structure. *Biochemistry.* 39:5722–5730.
22. Mylne, J. S., L. Y. Chan, …, D. J. Craik. 2012. Cyclic peptides arising by evolutionary parallelism via asparaginyl-endopeptidase-mediated biosynthesis. *Plant Cell.* 24:2765–2778.
23. Čemažar, M., N. L. Daly, …, D. J. Craik. 2006. Knots in rings. The circular knotted protein *Momordica cochinchinensis* trypsin inhibitor-II folds via a stable two-disulfide intermediate. *J. Biol. Chem.* 281:8224–8232.
24. Best, R. B., G. Hummer, and W. A. Eaton. 2013. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. USA.* 110:17874–17879.
25. Jennings, C., J. West, …, M. Anderson. 2001. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from *Oldenlandia affinis*. *Proc. Natl. Acad. Sci. USA.* 98:10614–10619.
26. Cemazar, M., A. Joshi, …, D. J. Craik. 2008. The structure of a two-disulfide intermediate assists in elucidating the oxidative folding pathway of a cyclic cystine knot protein. *Structure.* 16:842–851.
27. Daly, N. L., R. J. Clark, and D. J. Craik. 2003. Disulfide folding pathways of cystine knot proteins. Tying the knot within the circular backbone of the cyclotides. *J. Biol. Chem.* 278:6314–6322.
28. Daly, N. L., S. Love, …, D. J. Craik. 1999. Chemical synthesis and folding pathways of large cyclic polypeptides: studies of the cystine knot polypeptide kalata B1. *Biochemistry.* 38:10606–10614.

29. Daly, N. L., L. Thorstholm, …, D. J. Craik. 2013. Structural insights into the role of the cyclic backbone in a squash trypsin inhibitor. *J. Biol. Chem.* 288:36141-8.

30. Craik, D. J. 2009. Circling the enemy: cyclic proteins in plant defence. *Trends Plant Sci.* 14:328–335.

31. Greenwood, K. P., N. L. Daly, …, D. J. Craik. 2007. The cyclic cystine knot miniprotein MCoTI-II is internalized into cells by macropinocytosis. *Int. J. Biochem. Cell Biol.* 39:2252–2264.

32. Ji, Y., S. Majumder, …, J. A. Camarero. 2013. In vivo activation of the p53 tumor suppressor pathway by an engineered cyclotide. *J. Am. Chem. Soc.* 135:11623–11633.

33. Contreras, J., A. Y. O. Elnagar, …, J. A. Camarero. 2011. Cellular uptake of cyclotide MCoTI-I follows multiple endocytic pathways. *J. Control. Release.* 155:134–143.

34. Qin, M., W. Wang, and D. Thirumalai. 2015. Protein folding guides disulfide bond formation. *Proc. Natl. Acad. Sci. USA.* 112:11241–11246.

Biophysical Journal, Volume *110*

Supplemental Information

Disulfide Bridges: Bringing Together Frustrated Structure in a Bioactive Peptide

Yi Zhang, Klaus Schulten, Martin Gruebele, Paramjit S. Bansal, David Wilson, and Norelle L. Daly

Biophysical Journal

Supporting Material

Disulfide Bridges: Bringing Together Frustrated Structure in a Bioactive Peptide

Yi Zhang,[1] Klaus Schulten,[1,2] Martin Gruebele[1,2,*], Paramjit S. Bansal,[3] David Wilson,[3] and Norelle L. Daly[3,*]

[1]Beckman Institute for Advanced Science and Technology, Center for Biophysics and Quantitative Biology, and [2]Department of Physics and Department of Chemistry, University of Illinois at Urbana-Champaign, Champaign, Illinois; and [3]Centre for Biodiscovery and Molecular Development of Therapeutics, Australian Institute of Tropical Health and Medicine, James Cook University, Cairns, Queensland, Australia

*Correspondence: mgruebel@illinois.edu; norelle.daly@jcu.edu.au

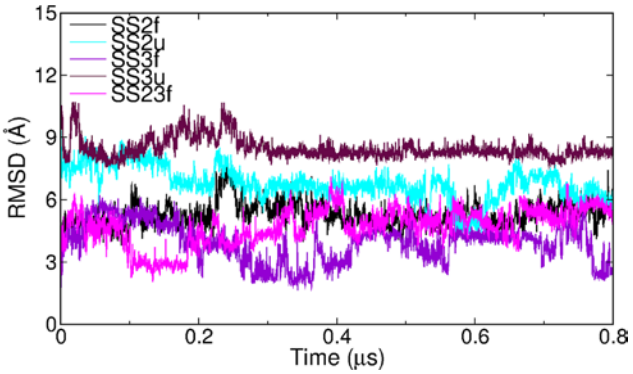

Fig. S1 RMSDs measured for Cα atoms for structures of MCoTI-II in different conditions; RMSD values were determined all relative to the native structure.

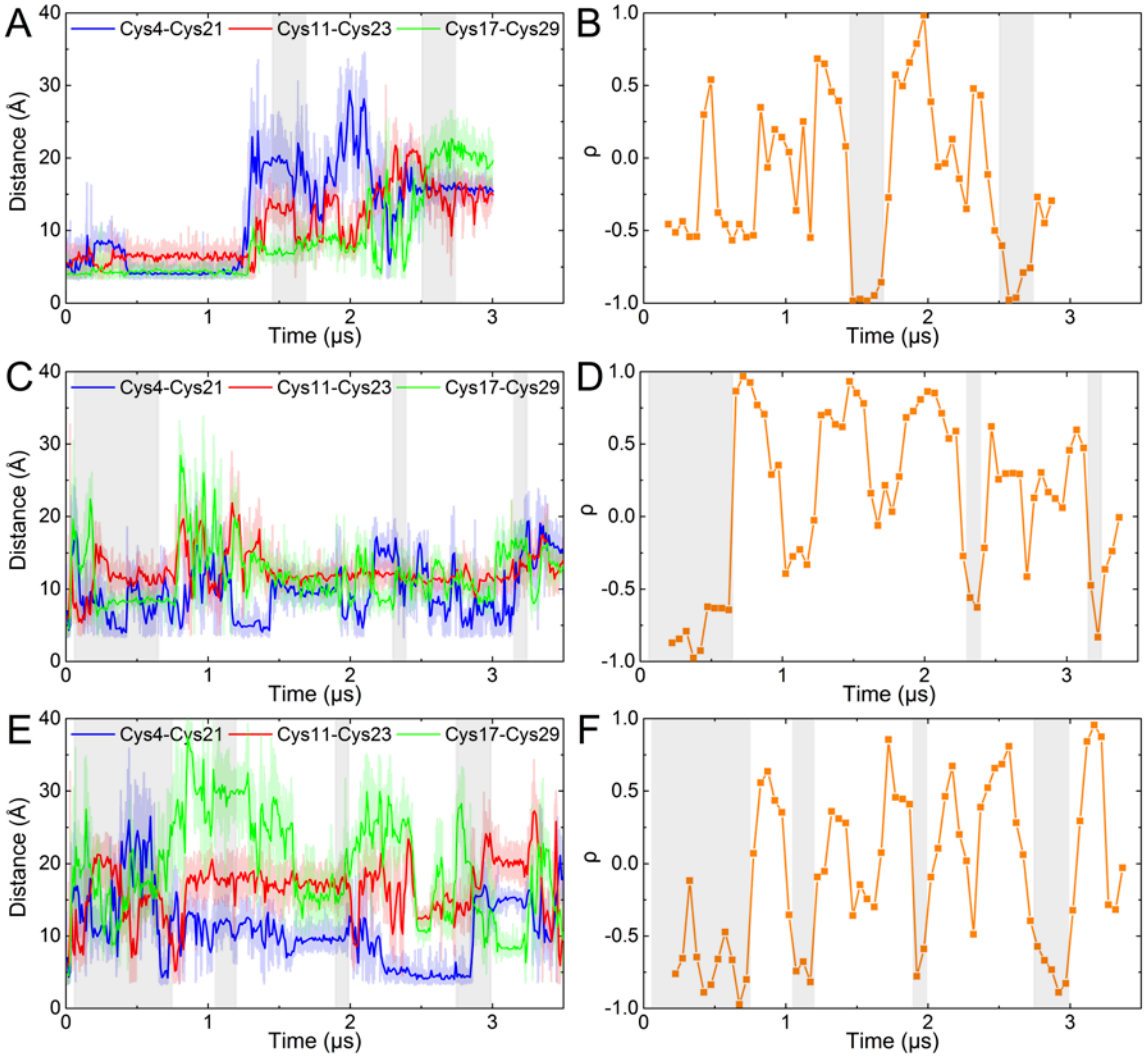

Fig. S2 Sulfur-sulfur distances and correlations monitored in simulations of different conditions. A. Sulfur-sulfur distances monitored in simulation SS0f(b). Average distances taken over 0.01 μs windows are shown as bold traces; light traces show raw distance data. Grey shading highlights the time when Cys11-Cys23 and Cys17-Cys29 distances are anti-correlated. B. Pearson correlation coefficient between Cys11-Cys23 and Cys17-Cys29 distances in simulation SS0f(b) from panel A. A gliding window of size 0.3 μs and step 0.05 μs is applied to monitor correlation as a function of time. The same grey shading as in panel A cover the time when large negative correlation coefficients are present. C. and D. Same as A. and B. for simulation SS0u(b). E. and F. Same as A. and B. for simulation SS0u(c).

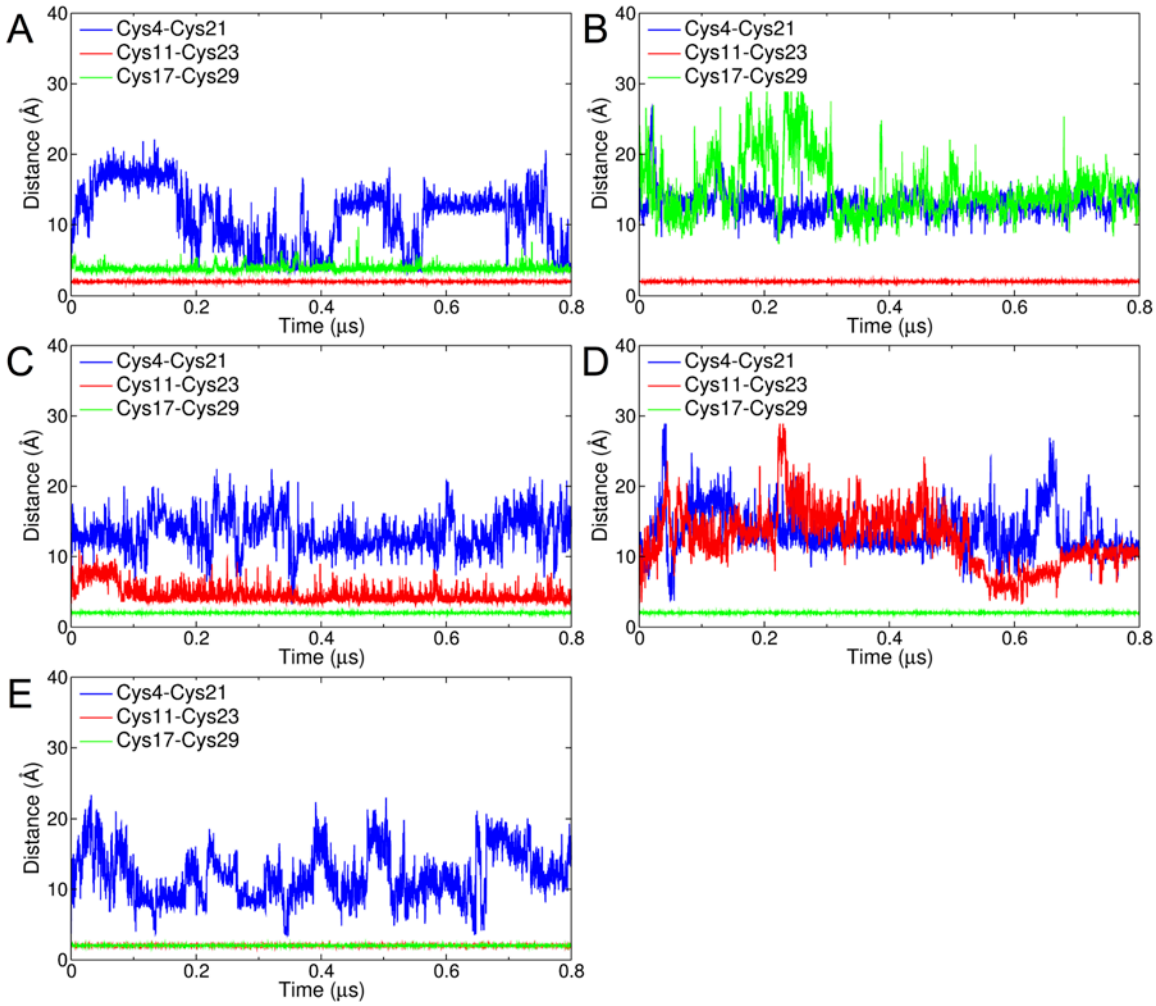

Fig. S3 Sulfur-sulfur distances monitored in simulations of different conditions: A. SS2f B. SS2u C. SS3f D. SS3u and E. SS23f.

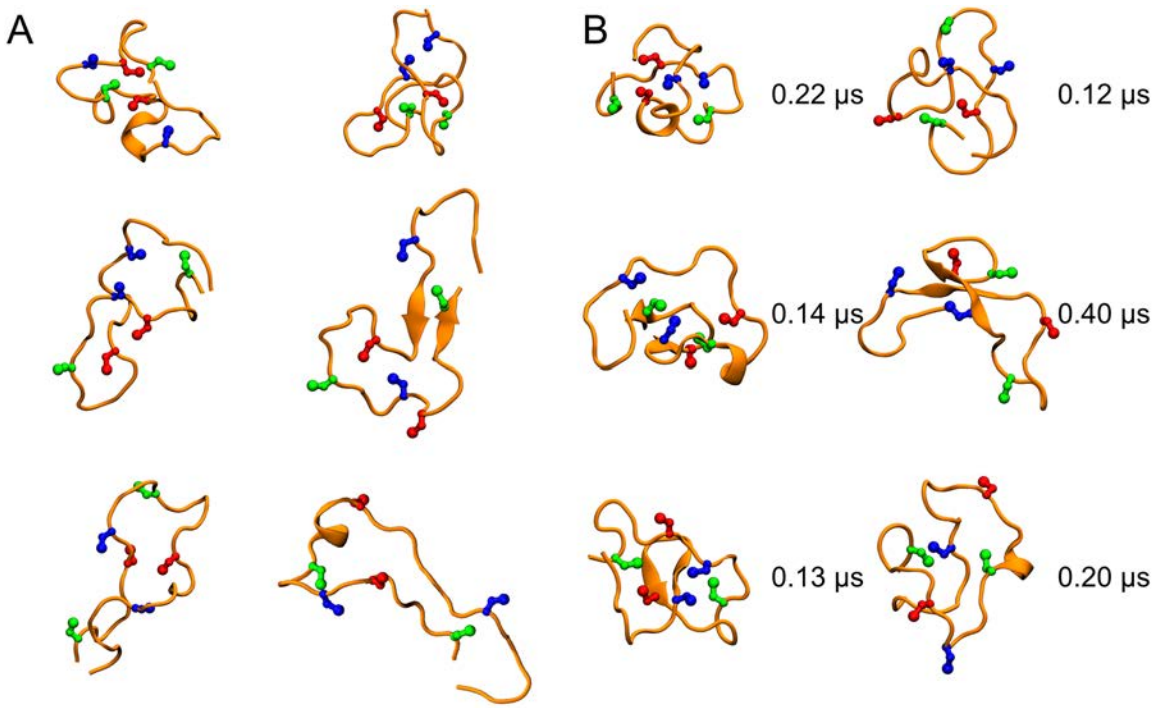

Fig. S4 A sampling of both anti-correlated and correlated structures in SS0f(a, b) and SS0u(a, b, c) . A. Randomly selected structures from highly mobile states right after the initial unfolding transitions in simulations, where distance 11-23 and distance 17-29 is anti-correlated. B. Representative structures of peptide trapped in less native-like microstates, labeled on the right with approximate lifetime for each microstate. The structures in these microstates were stabilized by secondary structures (alpha helix/beta sheet) and were more compact and less mobile compared with structures in panel A. The anti-correlation between distance 11-23 and distance 17-29 disappears in these microstates.