Title
Pierced Lasso Topology Controls Function in Leptin

Journal
The Journal of Physical Chemistry B, 121(4)

ISSN
1520-6106

Authors
Haglund, Ellinor
Pilko, Anna
Wollman, Roy
et al.

Publication Date
2017-02-02

DOI
10.1021/acs.jpcb.6b11506

Peer reviewed

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Pierced Lasso Topology Controls Function in Leptin

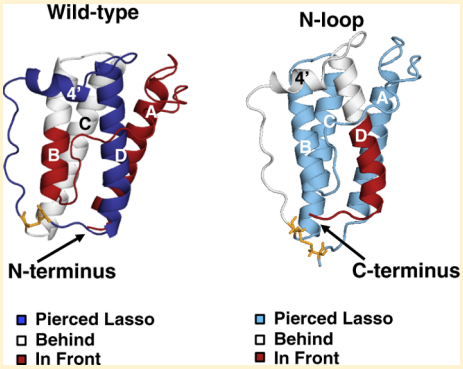
Ellinor Haglund,[†] Anna Pilko,[‡] Roy Wollman,[‡] Patricia Ann Jennings,[*,‡] and José Nelson Onuchic[*,†]

[†]Center for Theoretical Biological Physics (CTBP) and Departments of Physics and Astronomy, Chemistry and Biosciences, Rice University, Houston, Texas, United States

[‡]Department of Chemistry and Biochemistry, The University of California, San Diego (UCSD), La Jolla, California, United States

Supporting Information

ABSTRACT: Protein engineering is a powerful tool in drug design and therapeutics, where disulphide bridges are commonly introduced to stabilize proteins. However, these bonds also introduce covalent loops, which are often neglected. These loops may entrap the protein backbone on opposite sides, leading to a "knotted" topology, forming a so-called Pierced Lasso (PL). In this elegant system, the "knot" is held together with a single disulphide bridge where part of the polypeptide chain is threaded through. The size and position of these covalent loops can be manipulated through protein design in vitro, whereas nature uses polymorphism to switch the PL topology. The PL protein leptin shows genetic modification of an N-terminal residue, adding a third cysteine to the same sequence. In an effort to understand the mechanism of threading of these diverse topologies, we designed three loop variants to mimic the polymorphic sequence. This adds elegance to the system under study, as it allows the generation of three possible covalent loops; they are the original wild-type C-terminal loop protein, the fully circularized unthreaded protein, and the N-terminal loop protein, responsible for different lasso topologies. The size of the loop changes the threading mechanism from a slipknotting to a plugging mechanism, with increasing loop size. Interestingly, the ground state of the native protein structure is largely unaffected, but biological assays show that the activity is maximized by properly controlled dynamics in the threaded state. A threaded topology with proper conformational dynamics is important for receptor interaction and activation of the signaling pathways in vivo.



INTRODUCTION

Entanglement and knots constantly occur in nature, and in the microworld, knots in DNA and homopolymers are well characterized.[1−4] The more complex knots seen in proteins are harder to investigate, as proteins are heteropolymers composed of a combination of 20 different amino acids with different biophysical properties. In addition, proteins knots are often found in multimeric proteins, making the characterization of the folding mechanisms more complex. As new knotted topologies and new proteins containing knots are constantly discovered, the investigation of knots in proteins have generated intense interest. Many studies investigating this important class of proteins are becoming available. The main focus has been on the evolutionary origin, the threading mechanism—how to self-tie a protein chain into a knot, and the biological relevance and implication of a knotted topology in vivo.[5−14] Efforts to study the fully untied and unfolded chain have shown that the knot remains intact in the unfolded state much longer than was first anticipated.[3,15−17] The existence of "stable" knots in the unfolded state together with the challenge of defining an unfolded and untied chain from an unfolded and knotted chain makes it complicated to study the untied protein in vitro, leaving the threading mechanism elusive. The main focus has been the depth of the knot and the investigation of different types of knots, from (A) a slipknotted topology, where part of the chain threads through and doubles back, (B) the simplest configuration of a trefoil knot, where the protein backbone crosses the loop three times, to more complicated knots (C) with up to five crossings of the protein backbone.[10] In this work, we investigate loop topology on threading and function, using the elegant system that we discovered in the pleiotropic hormone leptin.[18,19] This system is analogous to circular permutations where the position of the disulphide bridge actually changes the topology of the protein, changing the threaded elements. Thus, switching the disulphide bridge not only has a local effect but also alters the threading and the overall topology. Additionally, a new N-/C-terminus is introduced that potentially affects the allosteric control and thus the biological activity of the protein. In the PL topologies, the threaded topology is formed by a covalent loop, where part of the polypeptide chain is threaded through, forming what we term a Pierced Lasso (PL) topology (Figure 1).[20] As it is hard to visualize a PL topology in a crystal structure, Dabrowski-Tumanski et al. developed the LassoProt web server (http://lassoprot.cent.uw.edu.pl/),[21] which enables analysis of biopolymers with PL topologies. The advantage of a PL topology, compared to other knotted proteins, is that the threaded topology can easily be manipulated, as the oxidation state of the

Received: November 15, 2016
Revised: December 27, 2016
Published: December 30, 2016

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

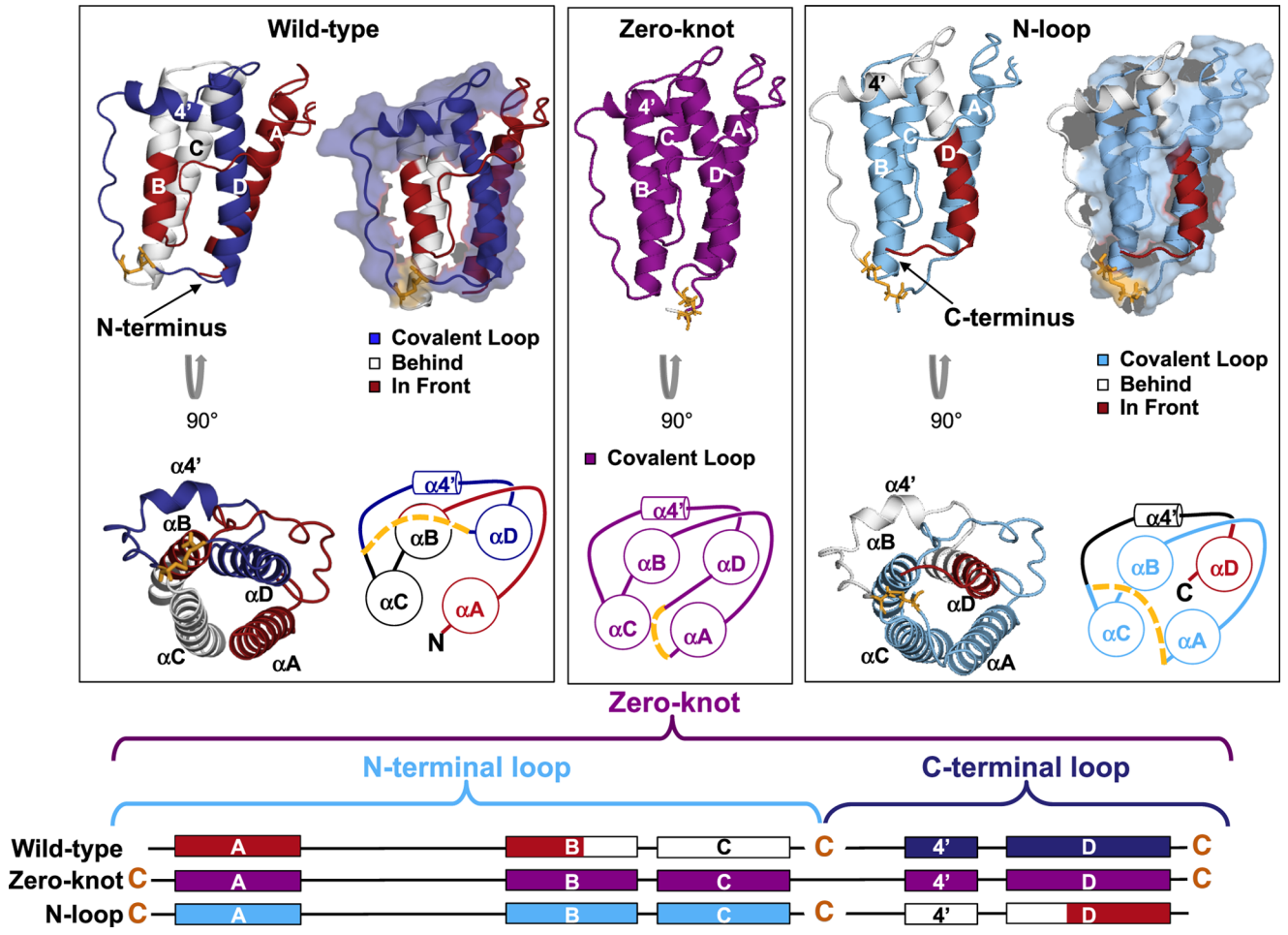


Figure 1. Cartoon representation of the different PL topologies. *Top panel:* The crystal structure (PDB code 1AX8) is used for the wild-type protein, whereas the Zero-knot and the N-loop are modeled from the wild-type structure (for modeling of the new structures see the Supporting Information). *Bottom panel:* It represents a block diagram describing the different topologies. The cysteines are highlighted in yellow, and the covalent loop is represented in deep blue (C-terminal loop), purple (Zero-knot), and light blue (N-terminal loop). In the case of the PL topologies, i.e., the wild-type and N-loop proteins, the element coming through the covalent loop is highlighted in red, and the elements behind the covalent loop is highlighted in white in the crystal structure and black in the top-view schematic figure.

cysteines can be controlled through the chemical environment (and thereby control the formation of the covalent loop). Another advantage is that the PL can be manipulated through point mutations of the cysteines, switching the size and the position of the covalent loop and thereby switching the PL topology. This makes it possible to study the same amino acid sequence without significantly altering the secondary and tertiary structures of the protein.[20] Although not easily recognized, a PL topology is not unique to leptin[19,20] but has been found in about 18% of all proteins containing a covalent loop (in a nonredundant set) deposited in the Protein Data Bank (PDB).[21] Many of these PL topologies are secreted proteins, extracellular proteins as well as redox sensors, enzymes, and metal and cofactor binding proteins, which provide a favorable environment for the formation of the disulphide bridge. Moreover, PL topologies are diverse proteins found in all kingdoms of life, involved in a large variety of biological functions, such as cell signaling, immune responses, transporters, and inhibitors to mention a few (http://lassoprot.cent.uw.edu.pl/).[21] As disulphide bridges are commonly used in protein engineering and therapeutics, it is of great importance to investigate the biophysical/biochemical properties of the PL topologies. Presumably, PL topologies can provide a new tool to steer folding and function in proteins, as the disulphide bridge can act as an on/off switch of the knotted topology and thus the biological function.[20]

Leptin is secreted from adipose tissue and acts through the JAK/STAT signaling pathway to initiate signaling and biological activity. Wild-type leptin first binds its receptor molecule to initiate a monomeric receptor−ligand pair. This receptor−ligand complex then dimerizes to initiate signaling. The first interaction occurs through helices A and C, behind the covalent loop (receptor site II), and the second interaction occurs through loop I in front of the covalent loop (receptor site III, Figure S1). On the basis of the observation of the placement of receptor-binding elements with respect to the covalent loop, we hypothesized that the native wild-type PL topology is essential for full biological activity and optimal conformational dynamics of leptin. Polymorphism in bovine leptin shows genetic variations of an N-terminal residue adding a third cysteine to the system,[22−26] allowing for a potential switching mechanism between covalent loops in vivo. Thus, leptin is an optimal system to explore the structural and conformational features important for performing biological functions. In the case of bovine leptin, all three proteins can spontaneously form from the same amino acid sequence within the same animal (Figures

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

1 and S2). We designed all three possible loop variants from bovine leptin into the human leptin sequence, to obtain three proteins with different threaded topologies (Figure 1). We designed each protein with only two cysteines so that only one covalent loop will be formed in a given variant. In our designed systems, three different covalent loops can be formed: (A) the wild-type protein with a C-terminal loop, (B) the unthreaded/linear fully circularized protein forming a Zero-knot or (C) the N-terminal loop (Figure 1). All loop variants were investigated with computational and experimental structural and functional assays. Our results show that all our designed proteins bind to the receptor and initiate signaling; however, the observed activity is affected by the threaded topology. Interestingly, our data reveal that "proper" conformational dynamics is essential for full biological activity, as adding residues to the N-terminus of the wild-type protein increases native-state dynamics (NSD) but lowers the activity. The N-terminal loop ties down these additional residues and restores full signaling activity. In the case of the Zero-knot, when we untie the protein backbone, the biological activity decreases, indicating that the PL topology is essential for full biological activity. As there are no significant changes observed in the overall dynamics of the Zero-knot protein, we performed NSD in the bound state, where leptin is bound to receptor domains D4 and D5. The result shows a decreased dynamics around the N-terminal, the receptor-binding site III as well as in the covalent loop for the Zero-knot protein. Hence, we attribute the lower activity seen for the Zero-knot protein to a change in the structure and/or conformational dynamics upon binding to the receptor. Thus, untying the protein through fully circularizing the protein backbone probably perturbs an important crosstalk between receptor-interacting sites II and III.

MATERIALS AND METHODS

Protein Design and Purification. Wild-type leptin has two naturally occurring cysteines forming one disulphide bridge between residues C96 and C146. To re-create the different loop variants found in polymorphic bovine leptin into the human sequence, we used the QuickChange mutagenesis kit (STRATAGene) for mutating the cysteines and serines so that each construct had only two cysteines to avoid formation of more than one covalent loop. To avoid perturbation of the tertiary structure, we elongated the N-terminus and added five residues to the sequence (purchased Genescript). Our wild-type protein has the same amino acid sequence as the crystal structure (PDB code 1AX8), where the N-terminal sequence reads M, I3, Q4, K5, and V6. In this work, we added the two wild-type amino acids that are not part of the crystal structure, for example, V1 and P2, as well as two glycines and the polymorphic cysteine at position 3 to elongate the terminus, so that the new sequence reads M, G, C or S, G, V1, P2, I3, Q4, K5, and V6. The nomenclature used in this paper is based on the wild-type numbers, starting from V1. Wild-type leptin was expressed together with the elongated wild-type construct (with five added residues at the N-terminus keeping the wild-type covalent loop intact). This construct was designed as structural control of the elongation, for example, the wt^{+5} protein. The N-loop protein was designed to form an N-terminal covalent loop between the new N-terminal cysteine and C96, to mimic R4C mutation seen in bovine leptin[22−26] (Figure 1). In the case of the Zero-knot, C96 was mutated to a serine so that the covalent loop between residue C146 and the new N-terminal cysteine could form. The new constructs were expressed, and the monomeric protein was purified as described.[18]

NMR Experiments. Native $^{1}H-^{15}N$ heteronuclear single quantum coherences (HSQCs) and $^{1}H-^{15}N$ HSQC−transverse relaxation optimized spectroscopies (TROSYs) were collected at 15, 25, and 37 °C. The HSQC−TROSY experiments were carried out on a Bruker AVANCE 800 MHz magnet equipped with a triple-resonance cryoprobe, and the HSQCs were collected on a Varian 600 MHz magnet. Protein concentrations were about 100 μM for each experiment and in a buffer of 50 mM Bis-Tris at pH 6.3. We used the backbone assignments for the wild-type protein for the comparison of structural changes for the loop variants. Completely overlapping peaks are categorized as no change in structure and are colored blue in the inset figure of Figure 3, whereas peaks in red are peaks that have clearly moved in the noncrowded regions of the HSQC, that is, at high and low ppm values. All other residues are not included due to potential ambiguity and are colored white in the inset structure of Figure 3.

Equilibrium Titrations in Vitro. Equilibrium unfolding titrations were collected using average fluorescence wavelength,[27] monitoring W138 as a probe of the global unfolding reaction. Fluorescence spectra were collected with an excitation of 280 nm, and emission spectra were collected from 300 to 450 nm. Protein samples were prepared at a concentration of about 10 μM in a buffer of 10 mM MES at pH 6.3, with varying guanidinium chloride (GdmCl) concentrations ranging from 0 to 4 M. Equilibrium curves were fitted to a standard two-state equation as described for[28]

$$S = \frac{S_N + S_D \times K_{D\text{-}N}}{1 + K_{D\text{-}N}}$$

where

$$K_{D\text{-}N} = \exp\left(\frac{-\Delta G_{D\text{-}N}^{H_2O} + (m_{D\text{-}N} \times [\mathrm{GdmCl}])}{RT}\right)$$

where $m_{D\text{-}N}$ is the linear dependence of $\Delta G_{D\text{-}N}$ on denaturant concentration and $\Delta G_{D\text{-}N}^{H_2O}$ is the free energy of unfolding at 0 M GdmCl. S is the signal of the native (N) and denatured states (D).

Cell Cultures and Cell Line Generation. MCF 10A human cell line containing the leptin receptor (Figure S5) was obtained from ATCC (Manassas, VA, USA) and maintained as monolayers under standard tissue culture conditions following established protocols.[29] A genetically encoded biosensor for kinase activity, Erk kinase translocation reporter,[30] was transferred to a PiggyBac custom destination vector using Gateway Cloning Technology. The expression vector was transfected into a MCF 10A cell line together with transposase-expressing helper plasmid (kind gift from Dr. John Albeck, UC Davis) using a 3:1 ratio of FuGene HD (Promega) to DNA and allowed to incubate overnight. Drug selection was added (Blasticidin 4 μg mL^{-1}) 48 h post-transfection.

Immunoblot Analysis. MCF 10A cells were grown to 80% confluence onto six-well plates and left overnight in the assay media[29] to allow serum starvation. After incubation with protein at the indicated concentration for 1 h, whole-cell lysates from 1×10^6 treated and untreated cells were obtained using RIPPA buffer (50 mM Tris−HCl 8.0, 150 mM NaCl, 1% NP40, 0.5% deoxycolate, 0.1% sodium dodecyl sulfate (SDS)), supplemented with protease and phosphatase inhibitors

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

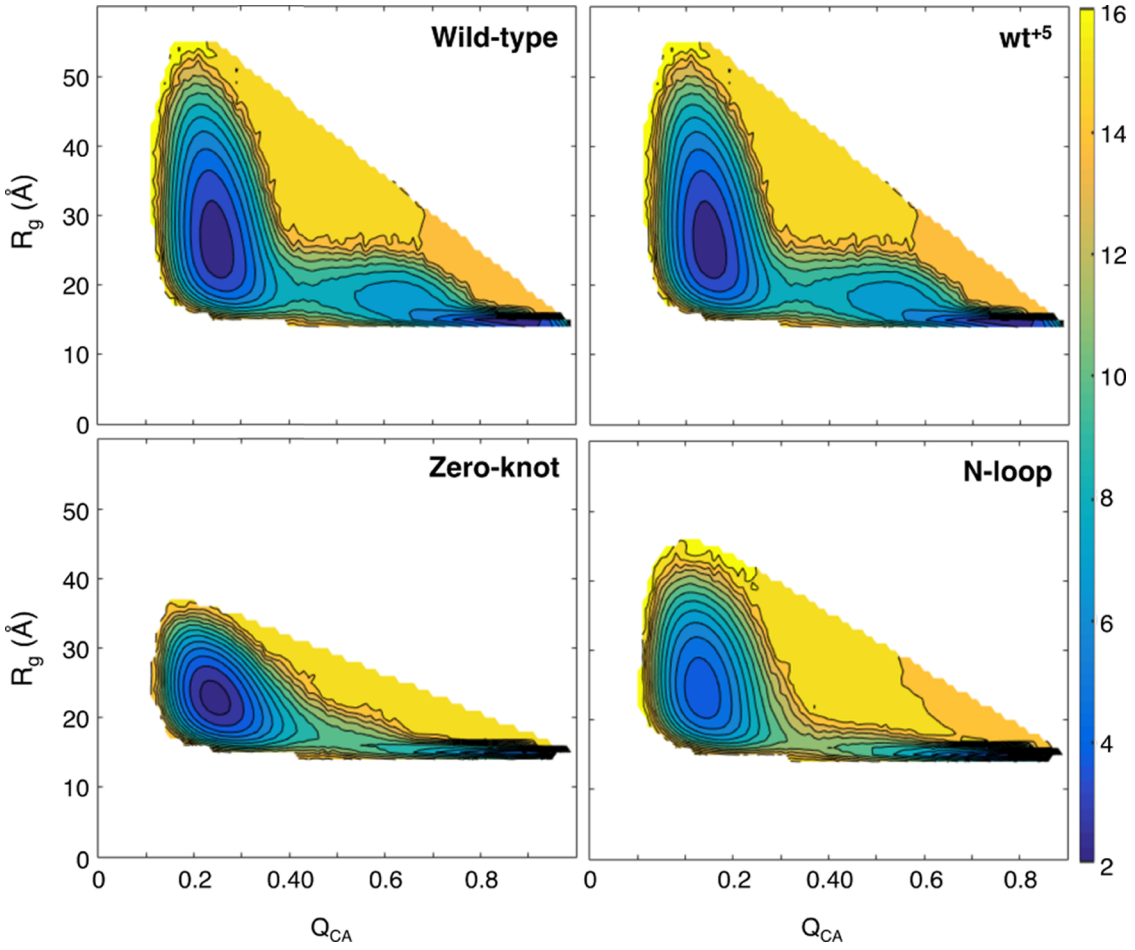


Figure 2. Folding free-energy landscape. The radius of gyration (R_g) is plotted vs the degree of nativeness (Q_{CA}), representing the population of states. Overall, the plots agree well except for R_g of the denatured basin at low Q_{CA}, and the broader transition state (TS) seen in the wild-type and wt^{+5}. The broader TS observed is an effect of loop size, as the smaller loop traps the denatured protein in a threaded state. R_g in the denatured basin varies inversely with loop size, where the largest effect is seen for the Zero-knot where the loop is 147-residue long. Colors represent the population of states from low in yellow to high in blue.

(SIGMA-Aldrich). The lysates were briefly sonicated on ice, heated at 100 °C for 5 min, separated by SDS-polyacrylamide gel electrophoresis, and transferred to low autofluorescence poly(vinylidene difluoride) membranes (Millipore) for immunoblotting. The results presented here were run on the same gel; however, some unrelated variants were also run on this gel. Therefore, the lanes for the unrelated proteins were removed for clarity. The following primary antibodies were used: Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) (D13.14.4E) XP Rabbit mAb antibody (Cell Signaling Technology Cat# 4370S, RRID:AB_2281741), Ob-R (M-18) antibody (Santa Cruz Biotechnology Cat# sc-1834, RRID:AB_2136231), and Actin (C-11) (Santa Cruz Biotechnology Cat# sc-1615, RRID:AB_630835). Antibody signals were analyzed as integrated intensities of regions defined around the bands of interests on a BioRad ChemiDoc XRS+ imaging system.

Structure-Based Models (SBMs). Coarse-grained Cα models were used to study the loop variants of leptin. Each amino acid is represented as a single bead, and attractive interactions are given to residue pairs close in the native state. These native interactions are identified based on a shadow map.[31,32] The energy function used was

$$
\begin{aligned}
V(r_{ij}) = & \sum_{\text{bonds}_{ij}} k_b(r_{ij} - r_{ij}^N)^2 + \sum_{\text{angles}_{ijk}} k_a(\theta_{ijk} - \theta_{ijk}^N)^2 \\
& + \sum_{\text{dihedrals}_{ijkl}} (k_d^1[1 - \cos(\varphi_{ijkl} - \varphi_{ijkl}^N] \\
& + k_d^2[1 - \cos(3(\varphi_{ijkl} - \varphi_{ijkl}^N))]) \\
& + \sum_{\text{contacts}} \varepsilon\left(5\left(\frac{\sigma}{r_{ij}}\right)^{12} - 6\left(\frac{\sigma}{r_{ij}}\right)^{10}\right) \\
& + \sum_{\text{no bonded atoms}} \varepsilon\left(\frac{\sigma}{r_{ij}}\right)^{12}
\end{aligned}
$$

where the last two terms correspond to attractive and repulsive interactions, respectively. r_{ij}^N denotes the native distance between atoms i and j along the sequence. The local topology of the chain is described by the native angles, θ_{ijk}^N, between the bonds connecting residue pairs ij and jk and by the native dihedrals, φ_{ijk}^N, or torsional angles between the planes defined by residues ijk and jkl. The strengths of the interactions are given in reduced energy units by constants $k_b = 2 \times 10^4 \varepsilon$ nm^{-2}, $k_a = 40\varepsilon$ rad^{-2}, $k_d^1 = \varepsilon$, and $k_d^2 = 0.5\varepsilon$, where ε is the reduced energy

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

unit, $\sum$ = 4 Å. Details of the model are presented elsewhere.[33−35]

Our models[36−38] are parameterized based on empirical values given from the protein structure (PDB code 1AX8). The available crystal structure does not describe the coordinates of residues 25−38 in loop I and were reconstructed according to ref 18. This "gap" in the crystal structure matches the receptor-binding site III, crucial for the dimerization of the receptor complex, which initiates leptin signaling. To mimic an oxidized state for the protein, we modified the contact between the two cysteines into a covalent bond at the appropriate position for each loop construct.

Molecular Dynamics (MD) Simulations. To create the input files for our simulations, we used the web server SMOG (http://smog-server.org/).[31,32,39] The GROMACS 4 package was used to perform the MD simulations.[40] All results are presented with reduced units, and an integration step of t = 0.0005 was used throughout. Stochastic dynamics was used with coupling constant 2 to maintain the temperature. The apparent folding temperatures are estimated from each maximum peak in each specific heat curve. For a formed native contact, the energy gain is measured by epsilon (ε) and thus the temperatures and energies reported in this article are measured in units of ε. Several constant temperature runs, including transitions from fully denatured to folded native states, were performed, and corrected folding mechanisms (Q_{CA} vs $q_{(segment)}$) were then created using the Weighted Histogram Analysis Method to create free-energy profiles ($F(Q)$). To evaluate the nativeness of a structure, we used Q_{CA}, the number/fraction of native contacts within a structure, as described in ref 19. Additionally, we also characterize the complexity introduced from the PL topology with the radius of gyration (R_g) of the configuration from the native state.

In Silico NSD. All-atom structure-based simulations[33,39] were performed to characterize the NSD of the loop variants of leptin. This model is an extension of the Cα model, where all heavy atoms are additionally taken into account.[33,39] Thus, two additional terms are added to maintain the conformation of the backbone and amino acid side chains. Details of the All-Atom model are presented elsewhere.[33,39] All NSD simulations are performed far below T_f, to ensure fully folded structures. The first 500 frames of each trajectory were deleted to ensure that the system was equilibrated. On the basis of the obtained trajectories, we calculate and diagonalize the (mass-weighted) covariance matrix for the backbone of the protein using GROMACS standardized tools. All structures are fitted to the native state of leptin available in PDB (1AX8). The slow component of the dynamics described by the first four eigenvectors was analyzed as described.[19] Principal component analysis is commonly used to identify essential motions, where a set of collective internal fluctuations make up the most important contributions to protein dynamics on a large number of conformations taken from a MD trajectory and are useful in combination with folding analysis.[41] The principal components were calculated by making the projections of the trajectory on the eigenvectors. To characterize the amplitudes of the motions of each atom with respect to the native structure, we calculated the root-mean-square deviation based on the first, second, third, and fourth eigenvectors, as seen in Figure 2. The crystal structure of receptor domains D4 and D5 (PDB code 3V6O without the antibody) was modeled together with the respective loop construct for the NSD of the bound monomeric receptor complex.[42]

RESULTS AND DISCUSSION

The elegant topology we discovered in leptin is an ideal system to study threading of a protein terminus across a closed loop, as the threaded topology is controlled through a single disulphide bridge.[18] The disulphide bridge can easily be manipulated by the chemical environment and can act as an on/off switch to control the threaded motif.[20] Polymorphism in bovine leptin shows genetic variations at position number 4, where the wild-type arginine is spontaneously substituted with a cysteine.[22−26] In this case, three proteins with a different covalent loop can be formed within the same animal: (A) the wild-type protein with a C-terminal loop of 50 residues, where helix C and half of helix B are threaded through the covalent loop, (B) the unthreaded/linear fully circularized protein forming a Zero-knot, and (C) the N-terminal loop of 97 residues, where part of helix D is threaded through the loop (Figure 1). To gain a broader understanding of how a protein can spontaneously self-tie, we generated all three loop variants in the human leptin sequence: (A) the human wild-type protein, (B) the wt^{+5} protein (our control protein with five extra residues at the N-terminus), (C) the Zero-knot protein, and (D) the N-loop protein (for a more detailed description, see the Materials and Methods section below) (see Figure 1).

Polymorphism in Bovine Leptin Modifies the PL Topology. The satiety hormone leptin is found among eukaryotes, especially in mammals where almost 80% of all leptin sequences are found.[43] The sequence homology is high, with the most conserved regions involving helices A and C together with loops I and IV. This also represents the receptor-interacting sites II and III, respectively.[42,44] The two cysteines and part of helices D and B are also conserved (Figure 1). These regions are in agreement with residues important for folding and threading,[18] suggesting that the PL topology is conserved across the leptin family (Figure S2, top panel). However, polymorphism in bovine leptin shows diversity in residue 4, where the wild-type arginine is sporadically substituted with a cysteine residue. In these animals, calf mortality is higher, and bovine health is compromised if they survive to adulthood. However, the population of topologies is unclear, as these leptin variants are cleared from the blood rapidly.[22−26] Residue 4 is part of a nonconserved region in the sequence, where glutamine, arginine, and tryptophan are the most common residues (Figure S2, top panel). Substituting the positively charged arginine with a polar cysteine could potentially break the salt bridges formed by this N-terminal residue or affect the receptor interactions, as helix A is part of the binding interface. However, as glutamine is as common as arginine in other leptin orthologs at this position, one can speculate that this substitution has negligible effects on protein stability and/or receptor interaction from a simple side-chain perspective. Nonetheless, a third cysteine within the same amino acid sequence can possibly form three different combinations of disulphide bridges, thereby introducing new topological variants, switching the position and size of the covalent loop (Figure 1). In line with this observation, structurally homologous, but sequence divergent, PL topologies discovered to date all have an N-terminal loop, save leptin.[19] To test the importance of loop size as well as the biological implications of an N- versus C-terminal loop in leptin, we switched the covalent loop in human leptin through point mutations. Native $^{1}H-^{15}N$ HSQC−TROSY spectra were collected to probe the structural effect from manipulation of

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

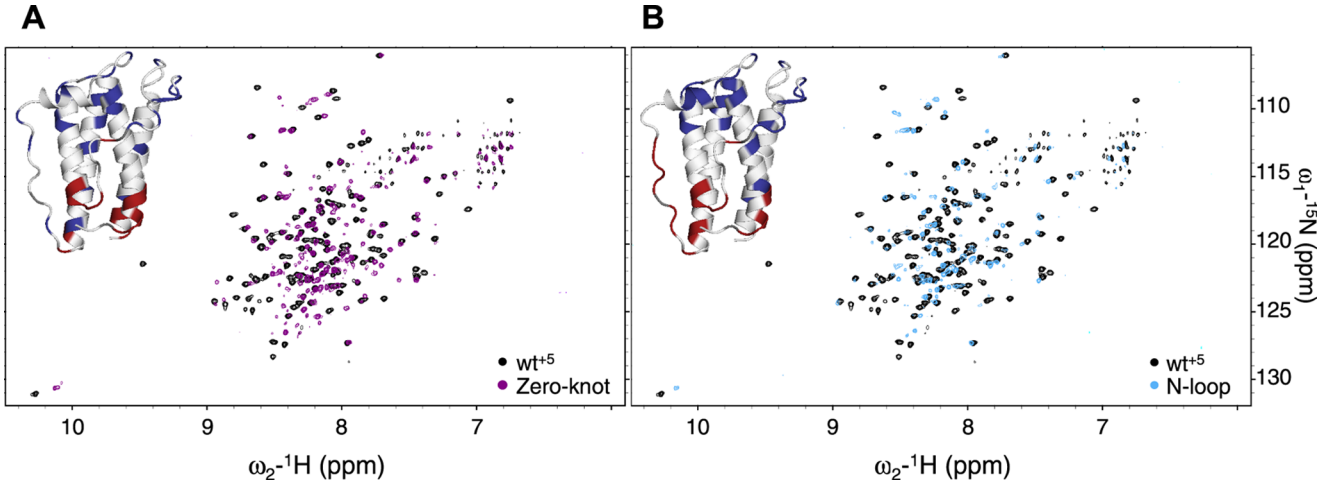


Figure 3. Evaluation of the loop variants. Overlay of the $^{1}H-^{15}N$ TROSY–HSQC data at 15 °C shows a well-dispersed spectrum. The Zero-knot is represented in (A) and the N-loop in (B). There are peaks that shift (red) and peaks that agree well (blue) with the wt^{+5} control protein and the loop variants as mapped to the inset structure. It is clear that the large shifts are at the bottom of the four-helix bundle, in the vicinity of the manipulated original disulphide bridge, whereas the top of the bundle remains intact. The heterogeneity observed in the loop variants indicates that the size of the covalent loop affects the conformational dynamics of the protein.

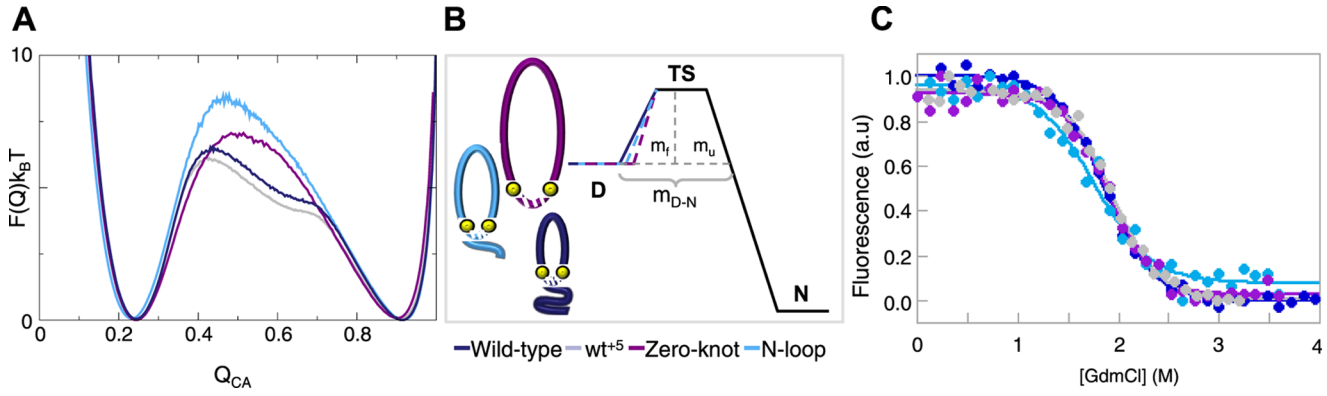


Figure 4. Elucidating the free-energy landscape. The wild-type protein is represented in deep blue, wt^{+5} in gray, Zero-knot in purple, and N-loop in light blue. (A) Two-dimensional plot of the free-energy landscape from SBM simulations, showing no significant effects from switching the covalent loop in silico. (B) Cartoon showing the three different configurations in the denatured state, where the size of the covalent loop seems to have an effect on the denatured state, as the covalent loop stays intact, and thus changes R_g in the denatured ensemble (Figure 2). (C) Thermodynamics data from equilibrium titrations in GdmCl, showing that the curves coincide, indicating that the switch in covalent loop has no significant effect. Overall, there are no significant effects on the thermodynamics upon switching the covalent loop in leptin.

the covalent loop (Figure 3). Whereas thermodynamic unfolding data suggest similarities in the native fold (see Figure 4C), our NMR spectrum indicates that while folded these proteins show different levels of heterogeneity. The thermodynamic stability and cooperativities suggest that the heterogeneous populations lie in the native basin with similar stabilities and relatively low barriers for interconversion. The observed heterogeneity in all NMR spectra is a result of changes in loop size, with larger loops showing more variation in peak intensities. The threaded element restricts the fluctuations in the wild-type protein the most as one-third of the protein is trapped on each side of the covalent loop. In the case of the N-terminal protein, the covalent loop is bigger (about 100-residue long), and only 15 residues are threaded through, making it easier to thread and unthread affecting the heterogeneity of the system. In the Zero-knot, only the termini are restricted and nothing is threaded, resulting in the most heterogeneous spectrum. All spectra were collected at 15 °C to maximize the observed signal, while maintaining well-dispersed spectra. Comparisons of the spectra obtained for the Zero-knot and N-loop proteins with respect to our control protein (the wt^{+5} protein, see the Materials and Methods section for a more detailed description) indicate that the largest chemical shift perturbations observed occur in the vicinity of the original wild-type disulphide bridge, at the bottom of the four-helix bundle, whereas the top of the bundle remains relatively intact (Figure 3).

Folding Free-Energy Landscape. Small molecules fold upon minimally frustrated funnel-like free-energy landscapes allowing for fast and robust attainment of the native state.[45,46] MD simulations are commonly used to study the folding events of small- to mid-sized proteins and more recently also larger systems, such as the ribosome,[47] and larger molecular machines.[48] Coarse-grained models have been able to capture nonmonotonic folding events such as backtracking[34,49−51] as well as more complicated folding events such as knotting of small proteins.[18,19,52,53] In this work, we used SBMs to study the effect of loop size and the position of the covalent loop in leptin. Folding and threading mechanisms of the loop variants were investigated using structure-based C_{CA} models. We

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

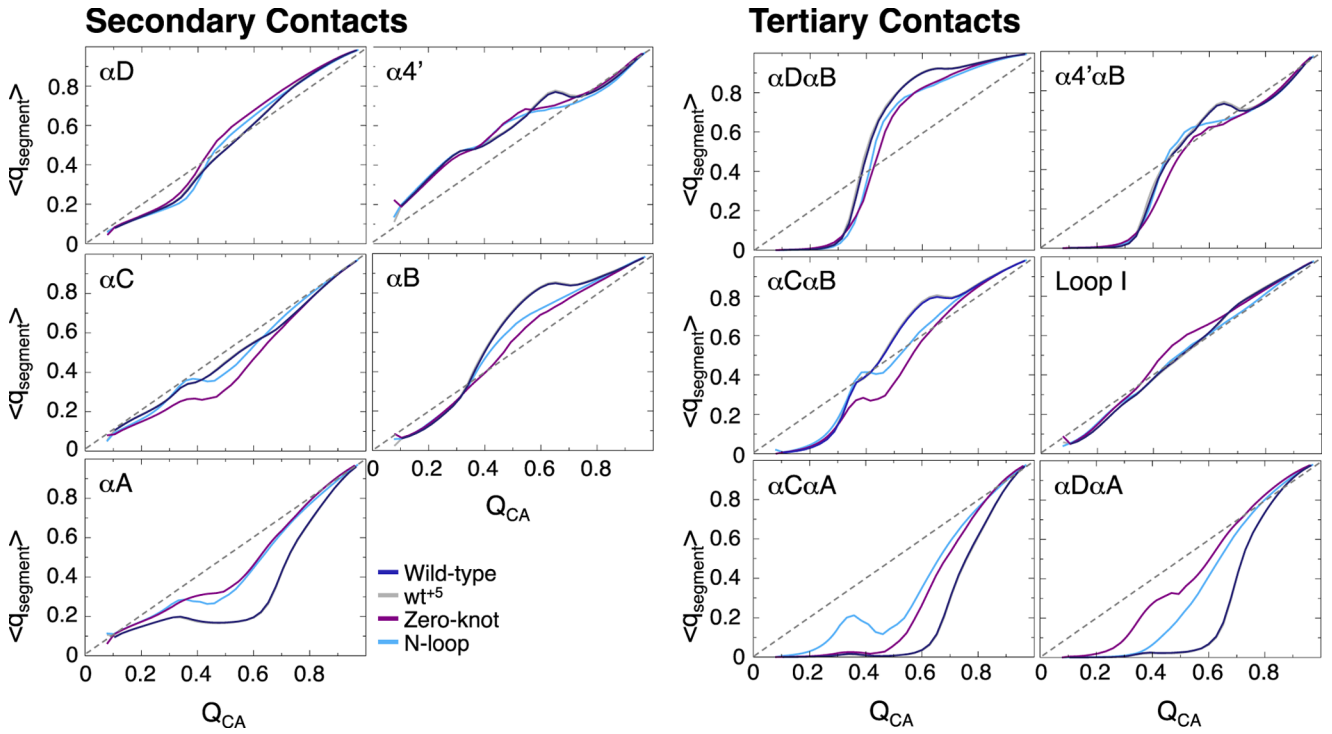


Figure 5. Secondary and tertiary formation of contacts. The individual formation of contacts within the α-helices is plotted in the left panel, and the tertiary contact formations between the secondary structures are plotted in the right panel. Q_{CA} represents the number of all native contacts, whereas $q_{\langle segment \rangle}$ represents the contact within each segment. The plot reveals that the switch of the covalent loop has a stabilizing effect seen in all plots of helix A for the Zero-knot and N-loop proteins, where the covalent loop retains the N-terminal helix. In all other cases, the curves coincide well and show no effects from switching the covalent loop.

reconstructed the new protein structures using the wild-type crystal as a model (PDB code 1AX8), and to generate the new disulphide bridge, we used the online software program Macro Molecular Builder (MMB 2.15, https://simtk.org/home/rnatoolbox).

To obtain information about the folding free-energy landscape, we define the reaction coordinate(s), the radius of gyration (R_g), and Q_{CA} (the fraction of native contacts formed at any point on the folding route). This approach allows for a qualitative and quantitative interpretation of the folding route shown in plots of R_g as a function of Q_{CA} in Figure 2[34,39,54−56] The similarity in the plots obtained for all variants indicates that the folding route is robust and implies that all proteins fold in a similar manner. The denatured basin of the Zero-knot and N-loop proteins, with a larger covalent loop, shows a smaller R_g than that of the other two proteins. This is expected in the oxidized state of the proteins, as the covalent loop is larger in these cases and kept intact in the denatured state. This will restrict the available search space of the denatured ensemble for the Zero-knot and N-loop proteins (Figures 2 and 4B). The wild-type and wt^{+5} proteins show a broader TS than the Zero-knot and N-loop proteins. A broader TS was observed previously for the wild-type oxidized protein.[18] This broadening of the TS was attributed to a trapped, unfolded but threaded conformation, which does not fully unfold, thus collapsing back to the native state instead of the fully unfolded and unthreaded state.[18] Hence, the observed broader TS in the case of both wt^{+5} and wild-type proteins can be attributed to these unsuccessful unfolding attempts, expanding R_g in this region.[18] In the case of the Zero-knot, the chain is never threaded, so a narrower TS is expected. As the covalent loop is larger in the N-loop protein and the "knot" is shallower, 15 residues cross the backbone compared to 45 residues in the wild-type and wt^{+5} proteins, and a "trapped" unfolded but threaded state is not observed. This indicates that it is easier for the N-loop protein to unfold and unthread than for the proteins with a smaller covalent loop. Soler et al. showed that the probability of keeping a knot in the denatured state depends on the depth of the knot, that is, the deeper the knot, the higher the probability of keeping the knot intact in the denatured state.[57] Thus, the trapped unfolded but threaded state observed in the wild-type and wt^{+5} proteins can likely be attributed to the depth of the knot in leptin. The population of an unfolded but threaded state also affects the shape of the free-energy profile, flattening the TS around a Q_{CA} of 0.5−0.7 for the wild-type and wt^{+5} proteins (Figure 4A).[18]

Each protein was simulated at its individual equilibrium temperature (at the folding T_f, shown in Figure 4A). The plot shows a robust folding free-energy landscape, with only minor changes of the barrier height, with about $2k_BT$ between all proteins ($F(Q)/k_BT$ = 6.1 for the wild-type protein, 6.4 for the wt^{+5} protein, 7.0 for the Zero-knot protein, and the $8.2k_BT$ for the N-loop protein). The decreased barrier is an effect of the trapped state observed in the wild-type and wt^{+5} proteins, and when these states are excluded, the barrier goes up and coincides with the other two plots. Taken together, the free-energy profile is not significantly affected by the size or position of the covalent loop. Our results are in agreement with those of Soler et al.,[57] where the depth of the knot seems to be important for the unfolding and unthreading of leptin, as deeper knot stays intact much longer in the denatured state, making it harder to fully unfold and untie the chain. Unthreading the knot and fully circularizing the protein backbone do not significantly affect the folding free-energy

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

landscape other than truncating R_g of the denatured basin, as the covalent loop gets bigger. Even though the effect is smaller, a reduction in R_g in the denatured state is also seen in the N-terminal protein. This is an effect of extending the covalent loop from 50 to 97 residues (Figure 4B).

Although the data presented in Figure 4A represent the average formation of all secondary and tertiary contacts, it is important to determine whether the order of events is affected by the changed topology. To address this issue, we plotted the contact formation between the secondary and tertiary segments as Q_{CA} versus $q_{\langle segment \rangle}$ in Figure 5. First, upon comparing the wild-type protein (represented in deep blue) to our control protein (the wt^{+5} protein, represented in gray), no significant changes are observed in the order of folding of specific segments as a function of total contacts formed as the curves coincide well in all plots. Therefore, we can conclude that the introduction of five extra residues at the N-terminal end has no significant effects on the folding route. Comparing the wild-type protein (represented in deep blue) to the Zero-knot protein (represented in purple) and the N-loop protein (represented in light blue) reveals that the switch of the covalent loop affects the formation of helix A. As helix A is now part of the covalent loop in both of the latter variants, it is stabilized by the formation of the covalent loop, thus shifting from a late folding event to an earlier event. This also affects the tertiary contact formations with neighboring helices, seen in the plots of contact formation between αCαA and αDαA, where the Zero-knot and N-loop proteins have more contacts formed at lower total contacts, Q_{CA}. If the contacts between helices A and C are formed too early on the folding route of the N-loop protein, backtracking is observed. This means that the formed contacts between helices A and C has to break, so that helix B can fold into place, and thus make contact with the new threaded element, helix D. Overall, the general folding route in leptin is not affected by the switch of the covalent loop, except for the earlier formation of helix A as a result of tethering it within the covalent loop.

Elucidating the Thermodynamic Behavior in Vitro. The thermodynamic stability of the loop variants of leptin was assessed by the change in exposure of a single tryptophan as a function of denaturant concentration (GdmCl). The equilibrium titrations are shown in Figure 4C, where the wild-type protein is represented in deep blue, the wt^{+5} protein in gray, the Zero-knot protein in purple, and the N-loop protein in light blue. The observed denaturant dependence of the unfolding transitions is nearly coincident, except for a small shift in the $m_{D\text{-}N}$ value and stability for the N-loop protein, changing from 1.8 for the wild-type protein to 1.6 kcal/mol/[D]. The decreased $m_{D\text{-}N}$ results in a change in stability ($\Delta G_{D\text{-}N}$), changing from 3.6 to 2.8 kcal mol^{-1} (Table 1), as there is no significant shift of C_m (the molar concentration midpoint in GdmCl) from 1.9 to 1.8 M GdmCl. Because of the large error in the fit of the N-loop protein, we conclude that the $m_{D\text{-}N}$ value is within experimental error (Table 1). However, switching the position of the covalent loop has a crucial effect on the denatured ensemble as all proteins are oxidized, with a resulting differential constraint on the entropy of the system depending on the size of the covalent loop (Figure 4B). Thus, changing the size of the covalent loop will affect the denatured basin (Figure 2), truncating R_g for the Zero-knot and N-loop proteins compared to that for the wild-type and wt^{+5} proteins.

Table 1. Thermodynamic Unfolding Data for the Loop Variants of Leptin from GdmCl Titrations[a]

	$m_{D\text{-}N}$ (kcal mol^{-1} m^{-1})	C_m (M)	$\Delta G_{D\text{-}N}$ (kcal mol^{-1})
wild-type	1.8 ± 0.09	1.9 ± 0.01	3.4
wt^{+5}	1.9 ± 0.13	1.9 ± 0.02	3.6
Zero-knot	1.8 ± 0.13	1.9 ± 0.02	3.4
N-loop	1.6 ± 0.21	1.8 ± 0.04	2.8

[a]The unfolding equilibrium data were fitted to a standard two-state equation (see Materials and Methods).

Conformational Dynamics Tuned for Optimal Activity in Vivo. Protein activity in vivo is affected by a number of biophysical properties, such as expression levels, stability of the protein, affinity for the receptor as well as the dynamics of the native state. Protein dynamics is important as it affects the affinity for the receptor and can be observed in a wide range of timescales using sophisticated experimental methods, such as X-ray diffraction, NMR, or single-molecule fluorescence, or by computational efforts, such as MD, and sometimes quantum mechanical when electron or proton transfer is involved in the dynamics.[58] To quantify the NSD in silico, we performed all-atom structure-based simulations far below the folding temperature, where the protein is effectively always in the folded basin. These simulations can obtain information about the conformational dynamics of a system on a nanosecond-to-millisecond timescale. We calculated the essential dynamics of the backbone of each loop variant by projecting the trajectory onto the first four principal components (eigenvectors 1−4; Figures 6 and S3). Each new eigenvector, obtained by the diagonalization of the covariance matrix of positional fluctuations, corresponds to a collective motion of the system, where the first eigenvector corresponds to the largest concerted motions of the system and the fourth eigenvector corresponds to the collective quasi-constraint (usually referred to as near-constraint) vibrations.[59]

While investigating the NSD of leptin and four other four-helix bundles with a PL topology, we discovered that the NSD changed between the oxidized and reduced states, making the disulphide bridge act as an on/off switch controlling the dynamics of the system.[19] This was not observed in any of the nonthreaded topologies in our studies, as the NSD was the same in both oxidized and reduced proteins, except in the close proximity of breaking the disulphide bridge.[19] This suggests that the threaded topology is important for the conformational dynamics of the native state of PL proteins. Interestingly, leptin, unlike the other members of the PL family, is the most dynamic overall, and it is the oxidized, threaded state that has optimal dynamics for signaling activity. This enhanced dynamics may allow leptin to more easily engage in the diversity of receptors (five isoforms) it interacts with in vivo. The leptin−receptor interaction occurs through two different binding sites on the opposite side of the covalent loop in leptin, diagonally across from each other. First, leptin interacts with the receptor through binding of helices A and C (binding site II), in the back of the covalent loop, to initiate the monomeric complex (Figure S1). Second, the monomeric complex dimerizes through loop I (binding site III) and forms the active quaternary complex. Looking at the conformational dynamics of leptin, its most dynamic regions in the oxidized threaded state agree well with receptor-binding sites II and III, for example, helices A and C and loop I, respectively. This indicates that conformational dynamics is important for the interaction with the leptin receptor. Helix A is the most dynamic region in the wild-type

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

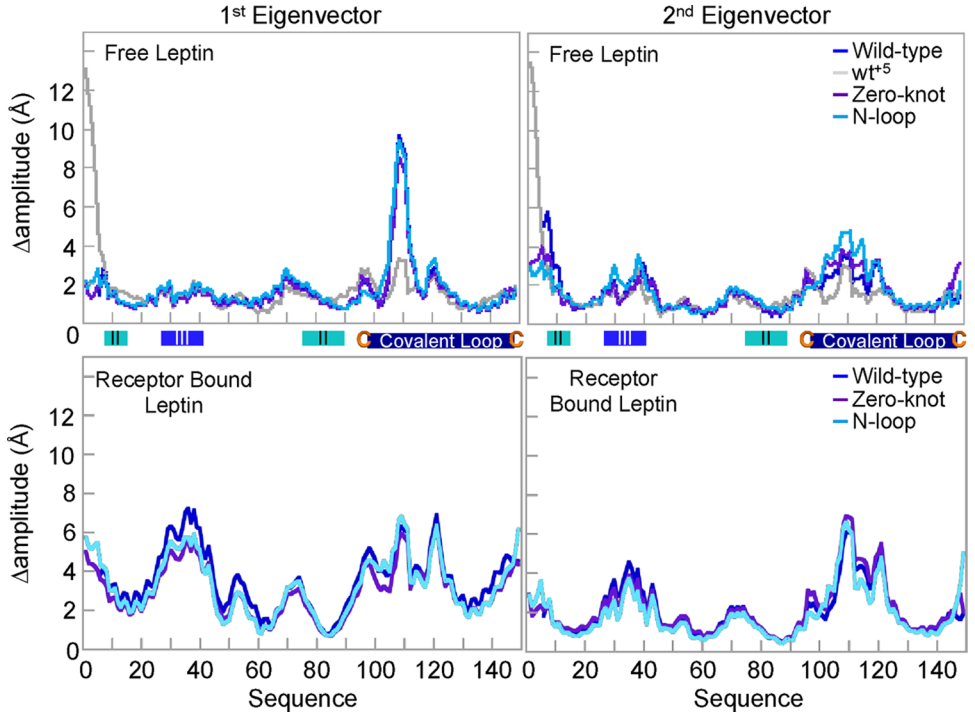


Figure 6. Conformational dynamics from SBM. The plots represent the first principal components for the first two eigenvectors of the essential dynamics of the protein backbone. The *top panel* shows the conformational dynamics of leptin in its free form and the *bottom panel* shows the dynamics in the monomeric receptor complex bound form. II and III receptor-binding sites and the position of the wild-type covalent loop are represented as rectangles in cyan, blue, and deep blue between the *top* and *bottom panels*. Surprisingly, there are no significant effects in the NSD upon switching the PL or unthreading the protein (Zero-knot). The data show that adding five residues to the N-terminus of the protein has a large effect on the NSD around helix A as well as in the covalent loop (residues C96−C146, represented in gray). When bound to domains D4 and D5, the Zero-knot protein shows a small decrease in its conformational dynamics at the N-terminus, in the region around binding site III, important for interaction with D3 and formation of the active quaternary complex, as well as in the covalent loop. There are no significant shifts seen for the other two eigenvectors plotted in Figure S3.

protein, which is part of binding site II.[19] Interestingly, shifting the position and size of the covalent loop and/or breaking the threaded topology, as seen in the Zero-knot protein, have no significant effect on the overall NSD, as these curves agree well (Figures 6 and S3). Breaking the covalent bond between the C-terminal cysteine and C96 does not show any significant effect on the dynamics. This is probably due to the fact that helix D is no longer part of the covalent loop but is still restricted as it is now the threaded element in the N-loop protein. However, the fact that we had to lower the temperature to obtain well-behaved proteins in vitro, seen from our $^{1}H-^{15}N$ TROSY−HSQC spectra (Figure 3), implies that switching the loop introduces new conformational dynamics, on a time scale different from that observed by SBM.

Piercing the Lasso: Switching the Mechanism. Molecular chaperones can enable successful folding by releasing trapped conformations and thereby facilitate correct folding to the native state. This is advantageous in the case of knotted proteins, where chaperones increase the number of successful folding events as well as speed up the folding reaction, as was demonstrated in the methyltransferase trefoil knots, YibK and YbeA.[60] However, biology has found a way to overcome the topological traps of threading along the folding free-energy landscape, as proteins can spontaneously self-tie without the use of molecular chaperones in vivo.[17] There are two possible mechanisms for a polypeptide chain to thread through a closed loop: (A) either by a so-called slipknotting mechanism or by (B) plugging the terminus across the loop.[61,62] The knotting mechanism remains elusive even though large efforts have been made to understand the underlying mechanism of threading. Through evaluation of the threading mechanism in structurally homologous PLB, with diverse functions, we found that the size of the covalent loop determines the dominant mechanism of threading, where PL proteins with a smaller loop dominantly slipknots, whereas proteins with a larger loop plugs their terminus across the closed loop.[19] However, these findings were a result of studies on proteins with diverse biological functions and no sequence homology. To investigate the threading mechanism further, we designed three different loop variants of human leptin, mimicking the polymorphic bovine leptin sequence (Figure 1). In this case, the same amino acid sequence is used, with the exception of the positions of the covalent loop (see the Materials and Methods section below for a description of the molecular modeling of the structures). Inspection of the folding trajectories of our N- versus C-terminal loop shows that the terminus prefers to plug through the larger covalent loop in the N-loop protein (95%), whereas the wild-type protein and the wt^{+5} protein dominantly slipknots through the smaller covalent loop (96%). In the case of wild-type leptin, the dominant route to thread the N-terminus across the loop is through slipknotting. Interestingly, visual inspection of the structure shows no preference for plugging or slipknotting of the terminus due to the depth of the knot as 50 residues are in front of the covalent loop and 50 residues are threaded through the loop. That is, no matter which mechanism is used, plugging or threading, the protein has to thread 50 residues through the loop. Therefore, switching the loop from the smaller wild type to the larger loop in the N-loop

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

protein also affects the number of residues being dragged across the covalent loop. Here, we see a switch in the threading mechanism with the position of the covalent loop where the protein now has a 97-residue-long covalent loop where only 15 residues are plugged through.

Does the Covalent Loop Affect the Biological Activity? The activity of proteins can be tested through stimulation of human cell lines containing the target receptor for a specific ligand. In this work, we used a human MCF 10A cell line containing the leptin receptor and tracked the activity by monitoring the activation of the mitogen-activated protein kinase (MAPK) signaling pathways and the phosphorylation of extracellular signal-regulated kinase (Erk). We analyze the activity both through single time point (via western blot detection of phosphorylation of Erk; Figure 7) and through

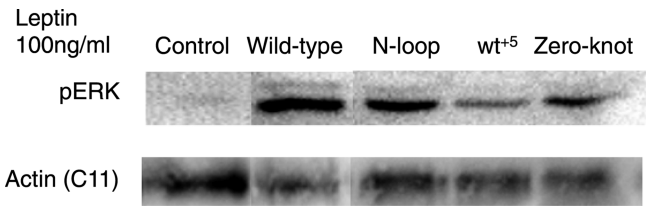


Figure 7. Biological activity assays through western blot. The activities were tested in human MCF 10A cells containing the leptin receptor. Single time points were analyzed, and the activation was monitored through the MAPK signaling pathways and the phosphorylation of extracellular signal-regulated kinase (Erk).

dynamic signaling (represented by a multivariate vector that contains a single cell's response at multiple time points;[63−67] Figure S4). Interestingly, all leptin variants show activity, even though the level of activation is affected by the PL topology. Switching the position of the loop from the C- to the N-terminus of the protein does not affect the overall activity. Thus, the effect of "locking" helix A to the four-helix bundle is not as dramatic as expected, indicating that our connecting loop, i.e., the new loop connecting the new N-terminal cysteine with C96, was long enough not to perturb the essential dynamics of helix A. Even though helix D is not in direct contact with the receptor, it appears that the dynamics of the covalent loop, and consequently helix D, is essential for the overall dynamics of the protein,[18] as helix D builds up the rigid part of the covalent loop in the wild-type protein important for the threading mechanism.

Fully circularizing the protein will break the PL topology and thus possibly affect the dynamics of the protein. Our in silico data suggest that the conformational dynamics, in the free form, is not affected by the threaded topology (top of Figure 6 and Figure S3). However, circularizing the protein decreases the biological activity, even though we do not observe significant effects in protein stability, thermodynamic behavior, or conformational dynamics of the native state. One might speculate that the structure and/or dynamics are different in the unthreaded state after binding to the receptor, as the activity is affected. Additionally, Gertler et al.[68] showed that mutating residues in loop I will switch the fully active agonist into an antagonist with maintained affinity for the receptor, implying that the effect is in the second, dimerizing, step when loop I interacts with domain 3 on the receptor to initiate the quaternary complex (Figure S2). To test our hypothesis that the lowered activity is an effect of the change in conformational dynamics in the bound monomeric state, we performed native-state dynamic simulations in a bound state, for example, one leptin molecule is bound to domains D4 and D5 of the leptin receptor, forming the monomeric complex (bottom of Figure 6 and Figure S3). Our results show that the Zero-knot protein has decreased dynamics in three areas, that is, at the N-terminus of the protein, around receptor-binding site III, and at the position of the wild-type covalent loop, around residue C96. As the III receptor-binding site is located in the middle of a very flexible loop area between helices A and B, where there are no coordinates in the crystal structure, thus creating a gap (residues 25−38), the results for this region are based on our reconstruction filling the gap.[18] Our NSD for the bound state shows that the Zero-knot protein has decreased dynamics in the first two eigenvectors at the N-terminus and in the covalent loop, and suggests that receptor-binding site III is affected. Taken together, unthreading the protein backbone probably perturbs the crosstalk between the receptor-binding sites II and III in the bound monomeric state, thus perturbing the dimer interface and lowering the biological activity.

Optimal Dynamics Is Essential for Biological Activity. The disulphide bridge in leptin plays an important role in the PL topology, acting as an on/off switch for the threaded topology.[20] Breaking the bond and untying the protein backbone have a dramatic effect on protein stability, conformational dynamics, and activity, as the reduced protein is less dynamic with lower activity than the oxidized protein.[18] Breaking the bond not only changes the local dynamics but surprisingly dramatically decreases the overall dynamics in regions far away from the disulphide bridge, making the reduced protein less adaptable for optimal receptor interactions. This indicates that the conformational dynamics is a unique feature for leptin−receptor interaction and activation of the JAK/STAT signaling pathway. To avoid perturbation of the four-helix bundle in the designed variants in the present work, we extended the N-terminus of the protein by adding five residues. Surprisingly, adding residues to the N-terminus dramatically changes the dynamics of helix A (making it more dynamic) as well as the dynamics in the covalent loop (decreasing the dynamics) seen as a shift in all four eigenvectors (the wt^{+5} protein is represented in gray in Figures 6 and S3). The wt^{+5} protein has a much lower biological activity than the wild-type protein, despite no changes in protein stability or folding mechanism (Figure 7). In addition, these five residues are present in the N-loop and Zero-knot proteins, thus it is unlikely that it is the composition of the extension that alters signaling. Therefore, we attribute this observed decrease in activity in the wt^{+5} protein to the increased NSD. That is, adequate dynamics is required for optimal activation and receptor interaction in the case of leptin, where helix A and the covalent loop have to be dynamic enough to be malleable to fit the receptor interface but not too dynamic to obtain full biological activity.

CONCLUSIONS

In this work, new PL topologies are introduced by a simple switch of the position of the covalent loop leading to three different topologies: the wild-type, Zero-knot, and N-loop proteins. To evaluate the biological implication of the threaded/unthreaded topology, we performed in silico simulations and thermodynamic titrations together with biological activity assays. Our data reveal that leptin requires a threaded topology with proper conformational dynamics for optimal signaling. In a previous work, we showed that the

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

reduced (fully unthreaded) protein has a much lower activity, NSD, and protein stability, compared to the threaded wild-type protein.[18] Here, it is hard to determine whether the effect is due to the decreased dynamics and/or stability or whether it is due to topological changes as the backbone is unthreaded in the reduced state. In the current work, we show that the overall conformational dynamics increases as the N-terminus is extended by five residues (seen for the wt^{+5} protein). This effect is seen not only in helix A but also in the covalent loop. Activity assays show that the wt^{+5} protein loses its activity, almost switching the protein from an active agonist to an inactive antagonist. This shows that leptin is optimized to be malleable to fit its receptor but not too dynamic or too rigid to be able to properly bind and interact with its receptor (Figure 7). In the case of the wt^{+5} protein, we can attribute the decreased activity to the increased dynamics of the receptor-binding region when untethered, as there are no significant effects on the native structure, and the protein stability is kept intact. Interestingly, neither the size nor the position of the covalent loop is important for biological overall activity. Nonetheless, while breaking the threaded topology via fully circularizing the protein backbone, the overall activity is reduced (Figure 7). As the Zero-knot protein shows no significant effects on the native structure or its conformational dynamics in its free unbound state (Figures 3 and 6), we performed NSD simulations in the bound monomeric state, for example, when leptin is bound to the receptor domains D4 and D5 (PDB code 3V6O). Thus, we hypothesize that the threaded topology is important in the crosstalk between receptor sites II and III in the bound monomeric state affecting the association with the second receptor complex. Our NSD simulation from the bound monomeric state showed that the conformational dynamics is in fact affected when we break the symmetry of the threaded topology. Therefore, we attribute the lowered activity to the effect on the conformational dynamics seen in the bound quaternary complex, as breaking the threaded topology in leptin changes the structure and/or the conformational dynamics of loop I in the monomeric state, lowering the affinity in the second step of activation making it harder to form the fully active quaternary complex.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.6b11506.

Leptin receptor complex; evolutionary comparison of the leptin sequence; conformational dynamics from SBM; image of human MCF 10 cells; western blot of the MCF 10A cell lysate; tryptophan fluorescence from native and denatured proteins (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: pajennings@ucsd.edu. Phone: (858) 534-6417 (P.A.J.).

*E-mail: jonuchic@rice.edu. Phone: (713) 348-4197 (J.N.O.).

ORCID

Patricia Ann Jennings: 0000-0002-7478-2223

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would like to thank Xuemei Huang for help with NMR measurements. We would also like to thank Jeffrey K. Noel for help with the construction of the loop variants of leptin for our SBM, Heiko Lammert for helpful scientific discussion, and Lannie Nguyen for help with protein purification. Work at the Center for Theoretical Biological Physics was sponsored by the National Science Foundation (Grants PHY-1427654 and CHE-1614101), by the Cancer Prevention and Research Institute of Texas (CPRIT—grant R1110), and by the Welch Foundation (Grant C-1792). Work at the University of California was sponsored by the National Science Foundation (Grants MCB-1212312 and PHY-1614407).

■ REFERENCES

(1) Bao, X. R.; Lee, H. J.; Quake, S. R. Behavior of complex knots in single DNA molecules. *Phys. Rev. Lett.* **2003**, *91*, No. 265506.

(2) Metzler, R.; Reisner, W.; Riehn, R.; Austin, R.; Tegenfeldt, J.; Sokolov, I. M. Diffusion mechanisms of localised knots along a polymer. *Europhys. Lett.* **2006**, *76*, 696.

(3) Sulkowska, J. I.; Sulkowski, P.; Szymczak, P.; Cieplak, M. Tightening of knots in proteins. *Phys. Rev. Lett.* **2008**, *100*, No. 058106.

(4) Dean, F. B.; Stasiak, A.; Koller, T.; Cozzarelli, N. R. Duplex DNA knots produced by *Escherichia coli* topoisomerase I. Structure and requirements for formation. *J. Biol. Chem.* **1985**, *260*, 4975−4983.

(5) Potestio, R.; Micheletti, C.; Orland, H. Knotted vs. unknotted proteins: evidence of knotpromoting loops. *PLoS Comput. Biol.* **2010**, *6*, No. e1000864.

(6) Sulkowska, J. I.; Rawdon, E. J.; Millett, K. C.; Onuchic, J. N.; Stasiak, A. Conservation of complex knotting and slipknotting patterns in proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, E1715−E1723.

(7) Taylor, W. R. Protein knots and fold complexity: some new twists. *Comput. Biol. Chem.* **2007**, *31*, 151−162.

(8) Taylor, W. R. Protein knots and fold complexity: some new twists. *Comput. Biol. Chem.* **2007**, *31*, 151−162.

(9) Taylor, W. R.; Lin, K. Protein knots: A tangled problem. *Nature* **2003**, *421*, 25.

(10) Virnau, P.; Mirny, L. A.; Kardar, M. Intricate knots in proteins: function and evolution. *PLoS Comput. Biol.* **2006**, *2*, e122.

(11) Liang, C.; Mislow, K. Knots in Proteins. *J. Am. Chem. Soc.* **1994**, *116*, 11189−11190.

(12) Faísca, P. F. Knotted proteins: a tangled tale of structural biology. *Comput. Struct. Biotechnol. J.* **2015**, *13*, 459−468.

(13) Mansfield, M. L. Are there knots in proteins? *Nat. Struct. Biol.* **1994**, *1*, 213−214.

(14) Lim, N. C.; Jackson, S. E. Molecular knots in biology and chemistry. *J. Phys.: Condens. Matter* **2015**, *27*, No. 354101.

(15) Bornschlögl, T.; Anstrom, D. M.; Mey, E.; Dzubiella, J.; Rief, M.; Forest, K. T. Tightening the knot in phytochrome by single-molecule atomic force microscopy. *Biophys. J.* **2009**, *96*, 1508−1514.

(16) Dzubiella, J. Sequence-specific size, structure, and stability of tight protein knots. *Biophys. J.* **2009**, *96*, 831−839.

(17) Mallam, A. L.; Rogers, J. M.; Jackson, S. E. Experimental detection of knotted conformations in denatured proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 8189−8194.

(18) Haglund, E.; Sulkowska, J. I.; He, Z.; Gen-Sheng, F.; Jennings, P. A.; Onuchis, J. N. The unique cysteine knot regulates the pleotropic hormone leptin. *PLoS One* **2012**, *7*, No. e45654.

(19) Haglund, E.; Sulkowska, J. I.; Noel, J. K.; Heiko, L.; Onuchic, J. N.; Jennings, P. A. Pierced lasso bundles are a new class of knot-like motifs. *PLoS Comput. Biol.* **2014**, *10*, No. e1003613.

(20) Haglund, E. Engineering covalent loops in proteins can serve as an on/off switch to regulate threaded topologies. *J. Phys.: Condens. Matter* **2015**, *27*, No. 354107.

(21) Dabrowski-Tumanski, P.; Niemyska, W.; Pasznik, P.; Sulkowska, J. I. LassoProt: server to analyze biopolymers with lassos. *Nucleic. Acids Res.* **2016**, *44*, W383.
(22) Brickell, J. S.; Pollott, G. E.; Clempson, A. M.; Otter, N.; Wathes, D. C. Polymorphisms in the bovine leptin gene associated with perinatal mortality in Holstein-Friesian heifers. *J. Dairy Sci.* **2010**, *93*, 340−347.
(23) Giblin, L.; Butler, S. T.; Kearney, B. M.; Waters, S. M.; Callanan, M. J.; Berry, D. P. Association of bovine leptin polymorphisms with energy output and energy storage traits in progeny tested Holstein-Friesian dairy cattle sires. *BMC Genet.* **2010**, *11*, 73.
(24) Nkrumah, J. D.; Li, C.; Yu, J.; Hansen, C.; Keisler, D. H.; Moore, S. S. Polymorphisms in the bovine leptin promoter associated with serum leptin concentration, growth, feed intake, feeding behavior, and measures of carcass merit. *J. Anim. Sci.* **2005**, *83*, 20−28.
(25) Reicher, S.; Ramos-Nieves, J. M.; Hileman, S. M.; Boisclair, Y. R.; Gootwine, E.; Gertler, A. Nonsynonymous natural genetic polymorphisms in the bovine leptin gene affect biochemical and biological characteristics of the mature hormone. *J. Anim. Sci.* **2012**, *90*, 410−418.
(26) Nkrumah, J. D.; Li, C.; Basarab, J. B.; Guercio, S.; Meng, Y.; Murdoch, B.; Hansen, C.; Moore, S. S. Association of a single nucleotide polymorphism in the bovine leptin gene with feed intake, feed efficiency, growth, feeding behaviour, carcass quality and body composition. *Can. J. Anim. Sci.* **2003**, *84*, 211−219.
(27) Rees, A. R.; Sternberg, M. J. E.; Wetzel, R. *Protein Engineering: A Practical Approach*; IRL Press at Oxford University Press: New York, 1992, ISBN-10: 0199631387.
(28) Chrunyk, B. A.; Evans, J.; Lillquist, J.; Young, P.; Wetzel, R. Inclusion body formation and protein stability in sequence variants of interleukin-1 beta. *J. Biol. Chem.* **1993**, *268*, 18053−18061.
(29) Debnath, J.; Muthuswamy, S. K.; Brugge, J. S. Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods* **2003**, *30*, 256−268.
(30) Regot, S.; Hughey, J. J.; Bajar, B. T.; Carrasco, S.; Covert, M. W. High-sensitivity measurements of multiple kinase activities in live single cells. *Cell* **2014**, *157*, 1724−1734.
(31) Noel, J. K.; Whitford, P. C.; Onuchic, J. N. The shadow map: a general contact definition for capturing the dynamics of biomolecular folding and function. *J. Phys. Chem. B* **2012**, *116*, 8692−8702.
(32) Noel, J. K.; Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. SMOG@ctbp: simplified deployment of structure-based models in GROMACS. *Nucleic Acids Res.* **2010**, *38*, W657−661.
(33) Baxter, E. L.; Zuris, J. A.; Wang, C.; Vo, P. L.; Axelrod, H. L.; Cohen, A. E.; Paddock, M. L.; Nechushtai, R.; Onuchic, J. N.; Jennings, P. A. Allosteric control in a metalloprotein dramatically alters function. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 948−953.
(34) Clementi, C.; Nymeyer, H.; Onuchic, J. N. Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* **2000**, *298*, 937−953.
(35) Lammert, H.; Schug, A.; Onuchic, J. N. Robustness and generalization of structure-based models for protein folding and function. *Proteins* **2009**, *77*, 881−891.
(36) Doshi, U.; McGowan, L. C.; Ladani, S. T.; Hamelberg, D. Resolving the complex role of enzyme conformational dynamics in catalytic function. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 5699−5704.
(37) Kalodimos, C. G.; Biris, N.; Bonvin, A. M. J. J.; Levandoski, M. M.; Guennuegues, M.; Boelens, R.; Kaptein, R. Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science* **2004**, *305*, 386−389.
(38) Rasmussen, B. F.; Stock, A. M.; Ringe, D.; Petsko, G. A. Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature* **1992**, *357*, 423−424.
(39) Whitford, P. C.; Noel, J. K.; Gosavi, S.; Schug, A.; Sanbonmatsu, K. Y.; Onuchic, J. N. An all-atom structure-based potential for proteins: bring minimal models with all-atom empirical forcefields. *Proteins* **2009**, *75*, 430−441.
(40) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(41) Estácio, S. G.; Fernandes, C. S.; Krobath, H.; Faisca, P. F.; Shakhnovich, E. I. Robustness of atomistic Go models in predicting native-like folding intermediates. *J. Chem. Phys.* **2012**, *137*, No. 085102.
(42) Carpenter, B.; Hemsworth, G. R.; Wu, Z.; Maamra, M.; Strasburger, C. J.; Ross, R. J.; Artymiuk, P. J. Structure of the human obesity receptor leptin-binding domain reveals the mechanism of leptin antagonism by a monoclonal antibody. *Structure* **2012**, *20*, 487−497.
(43) Finn, R. D.; et al. Pfam: the protein families database. *Nucleic Acids Res.* **2014**, *42*, D222−D230.
(44) Mancour, L. V.; Daghestani, H. N.; Dutta, S.; Westfield, G. H.; Schilling, J.; Oleskie, A. N.; Herbstman, J. F.; Chou, S. Z.; Skiniotis, G. Ligand-induced architecture of the leptin receptor signaling complex. *Mol. Cell* **2012**, *48*, 655−661.
(45) Leopold, P. E.; Montal, M.; Onuchic, J. N. Protein folding funnels: a kinetic approach to the sequence-structure relationship. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8721−8725.
(46) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545−600.
(47) Ratje, A. H.; et al. Head swivel on the ribosome facilitates translocation by means of intra- subunit tRNA hybrid sites. *Nature* **2010**, *468*, 713−716.
(48) Jana, B.; Morcos, F.; Onuchic, J. N. From structure to function: the convergence of structure based models and co-evolutionary information. *Phys. Chem. Chem. Phys.* **2014**, *16*, 6496−6507.
(49) Chavez, L. L.; Gosavi, S.; Jennings, P. A.; Onuchic, J. N. Multiple routes lead to the native state in the energy landscape of the beta-trefoil family. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10254−10258.
(50) Gosavi, S.; Chavez, L. L.; Jennings, P. A.; Onuchic, J. N. Topological frustration and the folding of interleukin-1 beta. *J. Mol. Biol.* **2006**, *357*, 986−996.
(51) Fisher, K. M.; Haglund, E.; Noel, J. K.; Hailey, K. L.; Onuchic, J. N.; Jennings, P. A. Geometrical frustration in interleukin-33 decouples the dynamics of the functional element from the folding transition state ensemble. *PLoS One* **2015**, *10*, No. e0144067.
(52) Noel, J. K.; Onuchic, J. N.; Sulkowska, J. I. Knotting a protein in explicit solvent. *J. Phys. Chem. Lett.* **2013**, *4*, 3570−3573.
(53) Sulkowska, J. I.; Noel, J. K.; Onuchic, J. N. Energy landscape of knotted protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 17783.
(54) Bryngelson, J. D.; Wolynes, P. G. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7524−7528.
(55) Clementi, C.; Jennings, P. A.; Onuchic, J. N. How native-state topology affects the folding of dihydrofolate reductase and interleukin-1beta. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5871−5876.
(56) Onuchic, J. N.; Wolynes, P. G. Theory of protein folding. *Curr. Opin. Struct. Biol.* **2004**, *14*, 70−75.
(57) Soler, M. A.; Faisca, P. F. Effects of knots on protein folding properties. *PLoS One* **2013**, *8*, No. e74755.
(58) Gruebele, M. Protein dynamics in simulation and experiment. *J. Am. Chem. Soc.* **2014**, *136*, 16695−16697.
(59) Amadei, A.; Linssen, A. B.; Berendsen, H. J. Essential dynamics of proteins. *Proteins* **1993**, *17*, 412−425.
(60) Mallam, A. L.; Jackson, S. E. Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. *Nat. Chem. Biol.* **2012**, *8*, 147−153.
(61) Sulkowska, J. I.; Noel, J. K.; Ramirez-Sarmiento, C. A.; Rawdon, E. J.; Millet, K. C.; Onuchic, J. N. Knotting pathways in proteins. *Biochem. Soc. Trans.* **2013**, *41*, 523−527.
(62) Wallin, S.; Zeldovich, K. B.; Shakhnovich, E. I. The folding mechanics of a knotted protein. *J. Mol. Biol.* **2007**, *368*, 884−893.

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718

(63) Hao, N.; Budnik, B. A.; Gunawardena, J.; O'Shea, E. K. Tunable signal processing through modular control of transcription factor translocation. *Science* **2013**, *339*, 460−464.
(64) Hoffmann, A.; Levchenko, A.; Scott, M. L.; Baltimore, D. The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science* **2002**, *298*, 1241−1245.
(65) Purvis, J. E.; Lahav, G. Encoding and decoding cellular information through signaling dynamics. *Cell* **2013**, *152*, 945−956.
(66) Santos, S. D.; Verveer, P. J.; Bastiaens, P. I. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nat. Cell Biol.* **2007**, *9*, 324−330.
(67) Selimkhanov, J.; Taylor, B.; Yao, J.; Pilko, A.; Albeck, J.; Hoffmann, A.; Tsimring, L.; Wollman, R. Systems biology. Accurate information transmission through dynamic biochemical signaling networks. *Science* **2014**, *346*, 1370−1373.
(68) Gertler, A.; Elinav, E. Novel superactive leptin antagonists and their potential therapeutic applications. *Curr. Pharm. Des.* **2014**, *20*, 659−665.

DOI: 10.1021/acs.jpcb.6b11506
J. Phys. Chem. B 2017, 121, 706−718