




RESEARCH ARTICLE | AUGUST 08 2023

Effects of sequence-dependent non-native interactions in equilibrium and kinetic folding properties of knotted proteins

João N. C. Especial ; Patrícia F. N. Faisca  



J. Chem. Phys. 159, 065101 (2023)

<https://doi.org/10.1063/5.0160886>



View
Online



Export
Citation

CrossMark

500 kHz or 8.5 GHz?
And all the ranges in between.

Lock-in Amplifiers for your periodic signal measurements



Find out more

 Zurich
Instruments

Effects of sequence-dependent non-native interactions in equilibrium and kinetic folding properties of knotted proteins

Cite as: J. Chem. Phys. 159, 065101 (2023); doi: 10.1063/5.0160886

Submitted: 6 June 2023 • Accepted: 24 July 2023 •

Published Online: 8 August 2023



View Online



Export Citation



CrossMark

João N. C. Especial^{1,2}  and Patrícia F. N. Faisca^{1,2,a)} 

AFFILIATIONS

¹ Departamento de Física, Faculdade de Ciências, Ed. C8, Universidade de Lisboa, Campo Grande, Lisboa, Portugal

² BioISI—Biosystems and Integrative Sciences Institute, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, Lisboa, Portugal

^{a)} Author to whom correspondence should be addressed: pffaisca@fc.ul.pt

ABSTRACT

Determining the role of non-native interactions in folding dynamics, kinetics, and mechanisms is a classic problem in protein folding. More recently, this question has witnessed a renewed interest in light of the hypothesis that knotted proteins require the assistance of non-native interactions to fold efficiently. Here, we conduct extensive equilibrium and kinetic Monte Carlo simulations of a simple off-lattice C-alpha model to explore the role of non-native interactions in the thermodynamics and kinetics of three proteins embedding a trefoil knot in their native structure. We find that equilibrium knotted conformations are stabilized by non-native interactions that are non-local, and proximal to native ones, thus enhancing them. Additionally, non-native interactions increase the knotting frequency at high temperatures, and in partially folded conformations below the transition temperatures. Although non-native interactions clearly enhance the efficiency of transition from an unfolded conformation to a partially folded knotted one, they are not required to efficiently fold a knotted protein. Indeed, a native-centric interaction potential drives the most efficient folding transition, provided that the simulation temperature is well below the transition temperature of the considered model system.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0160886>

I. INTRODUCTION

Knotted proteins are globular proteins that embed a physical (i.e., open) knot in their native structure. The first knotted protein was identified in 1977,¹ and according to the most recent survey, $\approx 1\%$ of the protein entries deposited in the Protein Data Bank (PDB) correspond to knotted chains.² More recently, a study that investigated protein structures predicted with AlphaFold³ found that knots are present in 0.2% of the human proteomes.⁴ Knotted proteins turn out with different levels of topological complexity, from the simplest and most abundant trefoil (or 3_1) knot to the remarkably complex 6_3 knot, exclusively found in a human protein.⁵ Protein knots are further classified as deep or shallow, according to whether the number of residues that must be removed from one of the termini to unknot the chain is larger or smaller than 10, respectively.⁶

Despite being statistically rare,⁷ knotted proteins have attracted considerable attention, and during the last 20 years, a plethora

of experimental and theoretical investigations have been dedicated to the understanding of how they knot and fold (reviewed in Refs. 6 and 8), and establishing their biological role (reviewed in Ref. 9).

The very first computational study addressing knotted proteins¹⁰ investigated the folding and knotting mechanism of protein YibK, which features a deep trefoil knot. The study used a C-alpha model combined with a native-centric Gō potential, and sampling was performed with fixed-temperature Langevin Molecular Dynamics (MD) simulations. Interestingly, and quite surprisingly, it was found that in timescales typically used in molecular simulations, no knotting event was observed in any of the considered trajectories, suggesting that native interactions are not sufficient to knot the protein in a reasonable timescale. Another study, focusing on the same protein,¹¹ and using the same model and sampling methodology, reported an exceedingly small (1%–2%) knotting frequency in the attempted MD trajectories. However, a modified native-centric

Gō potential, in which non-native interactions were applied selectively to two chain segments (86–93 and 122–147) of YibK, was found to fold 100% of the attempted trajectories.¹⁰ These non-native interactions were identified by trial and error. A subsequent study framed on fixed temperature Monte Carlo (MC) simulations of a C-alpha model explored AOTCase, a protein larger than YibK, which also embeds a deep trefoil knot in its native structure.¹² Again, no knotting event was recorded when protein folding energetics was modeled through a native-centric Gō potential, but a modified Gō potential with sequence-dependent non-native interactions (i.e., non-native interactions that establish between any pair of contacting residues, and not only between selected residues as in Ref. 10) was found to produce a detectable fraction of knotted structures in partially folded conformations (when only $\approx 30\%$ of the native contacts are formed).

The situation is different for shallow knotted proteins on-lattice. In this case, a standard Gō potential efficiently folds model systems featuring shallow knots (including a 5_2 knot) in the context of MC simulations conducted at a fixed temperature, although knotted proteins require far more MC steps to fold than their unknotted counterparts.^{13–15} Moreover, off-lattice MC simulations of MJ0366, the smallest knotted protein found to date featuring a shallow trefoil knot, were also successful in correctly folding the chain, both at fixed temperature and with a replica-exchange engine.^{16,17}

The results summarized above thus seem to indicate that non-native interactions are necessary to fold proteins featuring deep knots in a biological timescale, but how they succeed in doing so is not clear. As a matter of fact, understanding the role of non-native interactions in the folding of knotted proteins has been pointed out as a major fundamental question in recent years.⁸

For several proteins, knotting proceeds through a single event in which one of the termini threads through a loop, the so-called knotting loop, formed by the remainder of the chain. It is possible that non-native interactions serve to enhance the native interactions that already exist between the regions involved in knotting as suggested for YibK,¹⁰ that they stabilize the transition state as shown by a combination of experiments and simulations on MJ0366,¹⁸ that they may assist threading events, or that they stabilize knotted intermediate states as suggested by lattice models.¹⁹

Here, we explore the role of native and non-native interactions in both equilibrium and kinetic properties of the folding transition of three knotted trefoil proteins by conducting extensive Monte Carlo simulations of a simple C-alpha off-lattice model. We consider the following two interaction potentials: a purely native-centric Gō potential and a modified Gō potential in which sequence-dependent non-native interactions are added as perturbations to the stronger native ones. Equivalent strategies have been used elsewhere.^{20–22} We found that non-native interactions can drive a downhill folding transition. Additionally, non-native interactions can substantially increase the knotting frequency above the transition temperature and in partially folded conformations below the transition temperature. Equilibrium knotted conformations are stabilized by non-native interactions that are non-local and proximal to native ones, thus enhancing them. Kinetic simulations show that the pure Gō potential is able to efficiently fold knotted proteins, provided the simulation temperature is well below the transition temperature. However, in that temperature range, non-native interactions dras-

tically increase the efficiency of the transition from unfolded to a partially folded knotted conformation.

II. MODELS AND METHODS

A. Protein representation

To represent the protein conformation, we use a simple C-alpha model in which residues are coarse-grained as hard spherical beads of uniform size, centered on the C-alpha atoms. These, in turn, constitute spherical joints that articulate rigid sticks that represent the amide planes that connect consecutive C-alpha atoms in the backbone. We adopt a radius of 1.7 Å for the beads, which is the van der Waals radius of C-alpha atoms,²³ and for the length of each stick, we adopt the distance between the C-alpha atoms of the respective bonded residues in the protein's native conformation, these being ~ 2.9 Å, for cis bonds, and 3.8 Å, for trans bonds. Two non-bonded residues are said to be in contact in the native conformation if the smallest distance between any two heavy atoms, one belonging to each residue, is ≤ 4.5 Å; this cutoff value is chosen because it is slightly larger than twice the average van der Waals radius of heavy atoms in proteins.

B. Interaction potential

1. Gō potential

For a purely native-centric Gō potential, the total energy of a conformation defined by bead coordinates $\{\vec{r}_i\}$ is given by

$$E_{Go}(\{\vec{r}_i\}) = \varepsilon \sum_{i,j \geq i+2}^N \varphi \left(\frac{|\vec{r}_i - \vec{r}_j| - |\vec{r}_i^{nat} - \vec{r}_j^{nat}|}{w} \right) \times \left(\chi_{ij}^{nat} \chi_{ij}^{nat} + \chi_{ji}^{nat} \chi_{ji}^{nat} + \frac{1}{2} \right) \Delta_{ij}^{nat}, \quad (1)$$

where N is the chain length measured in the number of beads, Δ_{ij}^{nat} is an element of the $N \times N$ native contact map matrix that takes the value 1 if the $i - j$ contact is present in the native conformation and is 0 otherwise, ε is a uniform intramolecular energy parameter (taken as -1 in this study, in which energies and temperatures are shown in reduced units), φ is the potential well associated with the native contacts, w is the half-width of this potential well, and the chirality of contact $i - j$ in the conformation under consideration is

$$\chi_{ij} = \Theta((\vec{r}_i - \vec{r}_j) \cdot [(\vec{r}_{j+1} - \vec{r}_j) \times (\vec{r}_{j-1} - \vec{r}_j)]) - \frac{1}{2}. \quad (2)$$

The chirality of the $i - j$ contact in the native conformation is

$$\chi_{ij}^{nat} = \Theta((\vec{r}_i^{nat} - \vec{r}_j^{nat}) \cdot [(\vec{r}_{j+1}^{nat} - \vec{r}_j^{nat}) \times (\vec{r}_{j-1}^{nat} - \vec{r}_j^{nat})]) - \frac{1}{2}. \quad (3)$$

In Eqs. (2) and (3), Θ is Heaviside's unit step function, which takes the value 1 if its argument is greater than zero and the value 0 otherwise. The chirality factor in (1) favors the native conformation *vis a vis* its mirror conformation, thereby ensuring chirality coherence among all contacts and convergence of the simulations toward the native ensemble for temperatures below the transition temperature.

In this study, we use inverse quadratic potential wells, and Eq. (1) becomes

$$E_{Go}(\{\vec{r}_i\}) = \epsilon \sum_{i,j \geq i+2}^N \left[\left(\frac{|\vec{r}_i - \vec{r}_j| - |\vec{r}_i^{nat} - \vec{r}_j^{nat}|}{w} \right)^2 + 1 \right]^{-1} \times \left(\chi_{ij} \chi_{ij}^{nat} + \chi_{ji} \chi_{ji}^{nat} + \frac{1}{2} \right) \Delta_{ij}^{nat}. \quad (4)$$

The half-width of the potential well, w , determines the degree of cooperativity of the folding transition, with a wider well leading to less cooperative transitions occurring at higher transition temperatures. We measure the degree of cooperativity of the transition by the ratio of the full width at half maximum (*FWHM*) of the C_V peak to the temperature at which the peak occurs, the melting temperature, T_m . A typical two-state transition that has been well-characterized experimentally is that of the B1 domain of protein G (PDB id: 2GB1²⁴), and its *FWHM*/ T_m ratio at pH 5.4 has been determined to be $\sim 4.4\%$.²⁵ Hence, the half-width of the potential well is adjusted to obtain a simulated *FWHM*/ T_m ratio between 4% and 5%. This is not just a technical detail of the model since very narrow transitions, which lead to artificially large cooperativity, are poorly sampled in equilibrium simulations that use a replica exchange procedure as done here.^{16,26,27} The width proposed above has been successfully used in previous simulations employing a similar potential and sampling.^{16,17,28,29}

When this potential is used we consider that a native contact is formed if the distance between the centers of the respective beads differs from the distance between their C-alpha atoms in the native conformation by less than the half-width of the potential wells, w .

2. Modified Gō potential

In order to consider the effects of heterogeneous native and non-native interactions on protein folding energetics, we consider a modified potential that is a linear combination of the pure Gō potential with a sequence-dependent (SD) potential:

$$E(\{\vec{r}_i\}) = (1 - \alpha)E_{Go}(\{\vec{r}_i\}) + \alpha \rho E_{SD}(\{\vec{r}_i\}), \quad (5)$$

with

$$E_{SD}(\{\vec{r}_i\}) = \sum_{i,j \geq i+2}^N \left[\left(\frac{|\vec{r}_i - \vec{r}_j| - r_{ij}}{w} \right)^2 + 1 \right]^{-1} \times \left[\left(\chi_{ij} \chi_{ij}^{nat} + \chi_{ji} \chi_{ji}^{nat} + \frac{1}{2} \right) \Delta_{ij}^{nat} + (1 - \Delta_{ij}^{nat}) \right] \epsilon_{ij}. \quad (6)$$

Since we want the modified Gō potential to represent a minimal perturbation of the pure Gō potential, the interaction parameters ϵ_{ij} in Eq. (6) are calculated through a least-squares fit to the native contact map matrix whose elements are Δ_{ij}^{nat} (see Appendix). The parameter $0 \leq \alpha \leq 1$ in Eq. (5) controls the strength of heterogeneous interactions, including the non-native ones, and the constant ρ , defined as

$$\rho = \frac{E_{Go}(\{\vec{r}_i^{nat}\})}{E_{SD}(\{\vec{r}_i^{nat}\})}, \quad (7)$$

normalizes the SD potential term to the same native energy as the Gō potential to ensure that the native energy remains unchanged on adding this perturbative term, whatever the value of α .

If $i - j$ is a native contact, $r_{ij} = |\vec{r}_i^{nat} - \vec{r}_j^{nat}|$ and, if $i - j$ is a non-native contact, r_{ij} is approximated by the arithmetic mean of all native contact distances. To identify which native and non-native contacts are formed in a sampled conformation, we consider that contact $i - j$ is formed if the distance between the centers of the beads i and j differs from r_{ij} by less than the half-width of the potential wells, w .

C. Monte Carlo sampling

The conformational space is explored with the Metropolis³⁰ Monte Carlo method by using a move set that comprises crankshaft [Fig. 1(a)] and pivot [Fig. 1(b)] moves.

These elementary moves can be performed in either of the following two ways: (1) by limiting the amplitude of the rotation so that no bead or stick is allowed to overlap or move across another, and (2) by not limiting the rotation and allowing such crossings to potentially occur as long as no steric clashes are present in the resulting trial conformation. A simulation that only performs moves of the first kind preserves the linear topology of the chain in all Monte Carlo steps and is designated as a linear topology preserving (LTyP) simulation. Conversely, a simulation that allows moves that take the chain across itself is designated as non-LTyP. The non-LTyP move set may alter the knot type.

1. Equilibrium simulations

To sample canonically distributed conformational states, we use Metropolis³⁰ Monte Carlo simulations combined with replica-exchange (or parallel tempering)^{31,32} (MC-RE). For complex structures, such as those of deeply knotted proteins, the preservation of linear topology provides correct equilibrium results, but only after long relaxation. Indeed, as shown in Ref. 33, they require up to two orders of magnitude more Monte Carlo steps to equilibrate than non-LTyP simulations of the same model system. Thus, in the case of equilibrium simulations, we use the non-LTyP variant of the

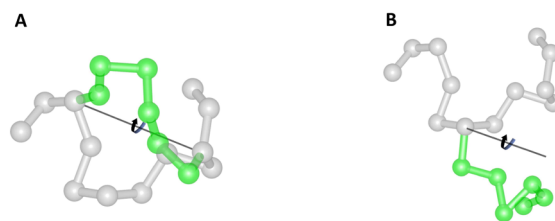


FIG. 1. The move set. In the crankshaft move (a) an axis passes through two randomly selected beads, and the beads between these are rotated around the axis. In the pivot move, (b) an axis passes through one randomly selected bead, and the beads between that and a randomly selected terminus are rotated around the axis.

move-set. Further details on the RE-MC simulations can be found elsewhere.³³

The Weighted Histogram Analysis Method (WHAM)³⁴ was used to analyze the data generated by the MC-RE simulations and produce maximum likelihood estimates of the density of states from which expected values for thermodynamic properties were calculated as functions of temperature. In particular, the heat capacity, C_V , is defined in reduced units as $C_V = (\langle E^2 \rangle - \langle E \rangle^2) / T^2$. The melting temperature, T_m , was determined as the temperature at which the C_V peaks, and the width of this peak at half-maximum was determined to calculate the $FWHM/T_m$ ratio.

The fraction of native contacts, Q , was chosen as the reaction coordinate, and WHAM was also used to project the density of states along Q to obtain free energy and knotting probability (p_K) profiles along this coordinate at a specific temperature.

In both equilibrium and kinetic simulations (see below), the topological state (knotted or unknotted) of a sampled conformation was determined using the Koniaris–Muthukumar–Taylor (KMT) algorithm.³⁵ The temperature at which $p_K = 0.5$ is regarded the central temperature of the knotting transition and designated as T_k .

2. Kinetic simulations

The physical process of protein folding is subject to excluded volume. Therefore, to properly capture the sequence of conformational transitions that establish a pathway starting from an initial unfolded conformation to a final knotted one, it is necessary to use the LTyP variant of the move-set. Thus, kinetic simulations are performed with the LTyP variant of the move set at fixed temperatures.

In this work, we are interested in assessing folding efficiency and knotting efficiency as a function of temperature. In both cases, the MC simulation starts from an unfolded conformation that was obtained by performing a MC simulation at a high temperature ($T = 4.0$). Such conformations are unknotted, and their Q is smaller than the Q at which the knotting probability (at T_k and for the Gō potential) is 0.01. When the purpose of the simulation is to access folding efficiency, it stops when the native conformation is found. The criterion used to determine if a conformation is native is quite stringent: a conformation is deemed native if it is knotted, and its Q is larger than that at which the knotting probability (at T_k and for the Gō potential) is 0.99.

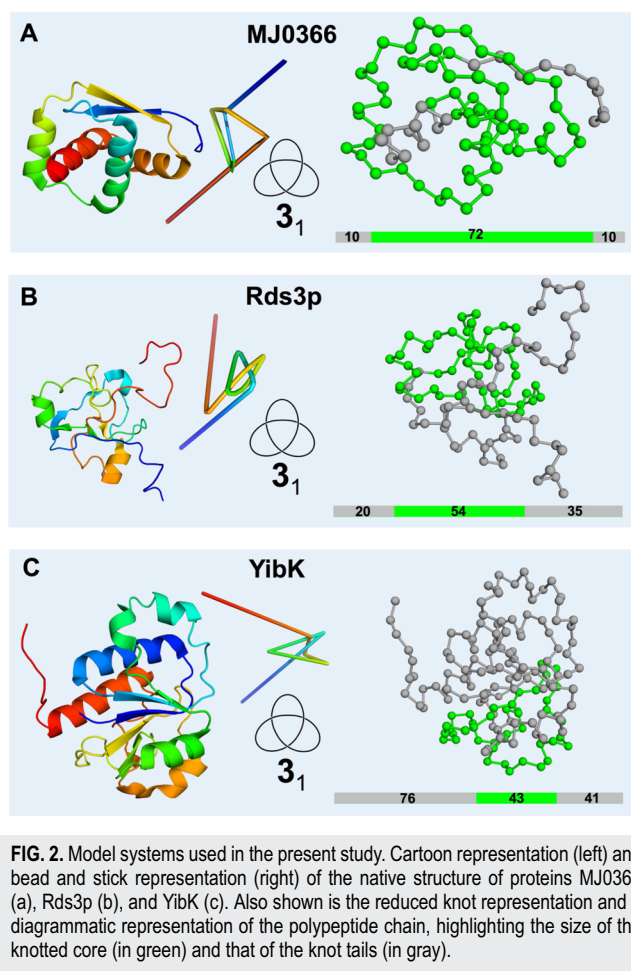
To investigate how folding and knotting efficiencies depend on the temperature, we conducted kinetic simulations at several temperatures. For all considered model systems, 24 kinetic simulations were performed at temperatures below T_m , and for proteins, MJ0366 and Rds3p, eight additional simulations were performed at T_m and temperatures above. All temperatures were separated by constant increments equal to the $FWHM$ of the C_V peak. Since the number of MC steps does not correspond to a physical time, instead of considering the folding time (or, alternatively, folding rate), we calculated the number of unfolding-to-folding (or knotting) transitions that occurred per million mcs (Mmcs) at each considered temperature. A model system that folds (or knots) more efficiently will exhibit a higher number of transitions at a certain temperature.

III. RESULTS

A. Model systems

We consider three proteins featuring a trefoil knot in their native structures.

Protein MJ0366 from *Methanocaldococcus jannaschii* (PDB id: 2efv)³⁶ is the smallest knotted protein found to date. It has 92 residues, and the knotted core comprises residues 11 to 82. Since both knot tails are short (10 residues), the knot is classified as shallow. The shallow knot is easily perceived in the native structure, being formed by threading the C-terminal helix, alpha-4, through the knotting loop formed by helices alpha-1 and alpha-2 [Fig. 2(a)]. The second protein is Rds3p (PDB id: 2k0a),³⁷ a metal-binding protein from *Saccharomyces cerevisiae*. Rds3p is 109 residues long, and the knotted core extends from residue 21 to residue 74. The N-tail of the knot is 20 residues long, and the C-tail comprises 35 residues [Fig. 2(b)]. Therefore, the knot is classified as deep. The third protein is YibK (PDB id: 1j85),³⁸ a methyltransferase from *Haemophilus influenzae*. This is the largest protein investigated in this study, featuring 160 residues with the knotted core extending from residue 77 to residue 119. In this case, the N-tail is 76 residues long, and



the C-tail is 41 residues long [Fig. 2(c)]. The knot is classified as deep.

B. Non-native interactions and the nature of the folding transitions

The effects of non-native interactions in decreasing the free energy barrier associated with a two-state folding transition have been studied theoretically²⁰ and observed in molecular simulations of several model systems.²⁰ In particular, a downhill scenario is predicted by the free energy landscape theory when the entropic and enthalpic contributions to the free energy compensate each other to the point where no significant ($>3RT$) barrier appears along the folding reaction coordinate.^{39–41}

In Fig. 3, we report the temperature dependence of the heat capacity, C_V , energy, E , and knotting probability, p_K , for proteins MJ0366 (left), Rds3p (center), and YibK (right) for the pure Gō and modified Gō potentials with increasing strength ($\alpha = 0.05, 0.10, 0.15$, and 0.20) of non-native interactions. Additionally, we also show the projection of free energy on the reaction coordinate Q at

temperature T_k . This temperature is nearly identical to T_m except for MJ0366 with $\alpha = 0.20$ due to the higher peak that occurs in its C_V at much lower temperatures. For the three considered model systems, the adopted parametrization of the pure Gō potential delivers a two-state folding transition. We observe the flattening of the heat capacity curves and the progressive lowering of the free-energy barrier as the strength of non-native interactions increases, such that for $\alpha = 0.15$ and $\alpha = 0.10$, the transition becomes effectively downhill for MJ0366 and YibK, respectively. In the case of MJ0366, the heat capacity curve exhibits a low-temperature peak that corresponds to a structural rearrangement, whereby the protein in a near-native conformation gets rid of residual non-native interactions. Interestingly, for protein Rds3p, the two-state transition is conserved despite the significant decrease in the free energy barrier. Consistently, the dependence of E on T conserves its sigmoidal shape in the case of Rds3p, but it becomes progressively less sigmoidal for the other model systems as non-native interactions become more prominent. The dependence of E on T also shows that non-native interactions contribute to energetically stabilize the partially folded (or denatured) conformations sampled at high temperatures above the transition midpoint, which explains why the temperature at which

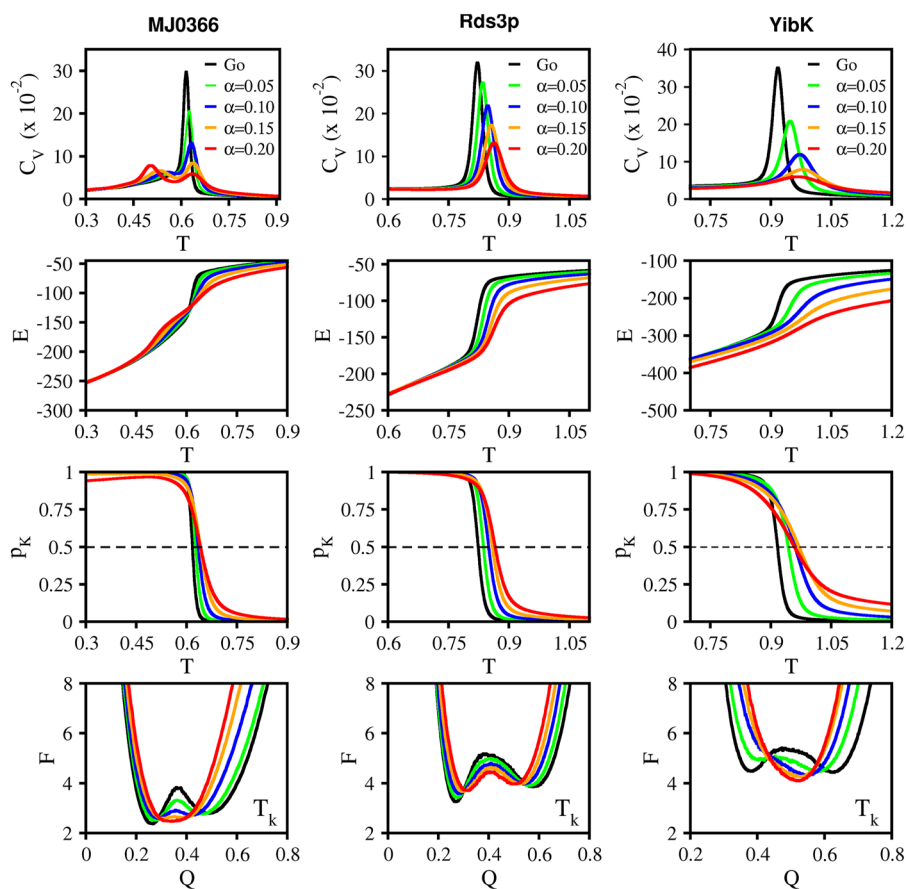


FIG. 3. Folding transition. Dependence of the heat capacity, C_V , energy, E , and knotting probability, p_K , on temperature T , for proteins MJ0366 (left), Rds3p (middle), and YibK (right). The bottom row shows the projection of the free energy, F , on the reaction coordinate Q at temperature T_k , which is the temperature at which $p_K = 0.5$.

the heat capacity peaks increases through the addition of non-native interactions.

It is also interesting to note that in the free energy profiles, the free energy minima corresponding to the thermally denatured ensemble shifts to the right as the strength of non-native interactions increases; this means that the thermally denatured ensemble becomes progressively more structurally consolidated (as the corresponding free energy minima is located at higher Q) through the addition of non-native interactions. At T_k the downhill transition relaxes to an equilibrium ensemble with approximately half of the native interactions established, but the two-state transitions exhibit a more structurally consolidated native ensemble.

C. Non-native interactions and knotting probability

The analysis of the knotting probability curves in Fig. 3 shows that their sigmoidal shape is fairly conserved across the considered interaction potentials (i.e., considered α values). However, the probability to find knotted conformations at high temperatures ($T > T_k$) increases significantly as the strength of non-native interactions is also increased (Fig. 3). This behavior is consistent with the energetic stabilization driven by non-native interactions at high temperatures, which compensate the entropic cost of knotting, thus making it more favorable in that temperature regime.

We also analyzed the dependence of p_K on Q at temperature T_k (Fig. 4). We designate by Q_k the reaction coordinate for which the knotting probability is 0.5. The fraction of native contacts Q_k varies between 0.35 (for MJ0366 with potential $\alpha = 0.20$) and 0.54 (for YibK with Gō potential). Interestingly, when the folding transition becomes downhill, the probability to find knotted conformations increases significantly below the transition midpoint (i.e., for $Q < Q_k$), for proteins MJ0366 and YibK, presumably because non-native interactions substitute native ones, enabling stable knotted conformations to form with a smaller number of native interactions.

D. Non-native interactions that stabilize the knot

A conditional probability contact map is an $N \times N$ matrix whose elements represent the probability of any contact being formed in an equilibrium ensemble of conformations with the same topological state (knotted or unknotted), and a fraction of native

contacts Q at temperature T . To identify the native and non-native contacts that most likely contribute to the stabilization of knotted conformations, we considered ensembles of conformations with Q_k at T_k . By subtracting the conditional contact map of unknotted conformations from that of knotted ones and taking only the positive values, a differential contact map is calculated that can be used to identify the native and non-native interactions having the highest differential probability. The latter, which are reported in Fig. 5, are the ones that most likely contribute to knot stabilization.

The analysis of Fig. 5 shows that the number of non-native interactions that stabilize knotted conformations is higher for the two proteins with deep knots. This is in line with previous observations that indicated that proteins with shallow knots do not require the assistance of non-native interactions to efficiently tie their backbones. Additionally, as the strength of non-native interactions increases the number of native local interactions (located close to the main diagonal) that stabilize the knotted conformations decreases. This is consistent with the fact that knotting requires interactions between residues that are located far away from each other in the polypeptide chain. The observed non-native interactions are thus mainly non-local being in tandem, or located nearby, native ones. This suggests that non-native interactions may operate as a system of redundant interactions that assist knotting, e.g., by keeping the protein in a conformation favorable for threading to occur. For the three proteins, it is possible to identify clusters of native and neighboring non-native interactions that are broadly conserved across different α (i.e., for different non-native interaction strengths). For YibK, these non-native interactions establish between residues located on the two stretches (75–87) and (119–122) of the polypeptide chain. In the case of protein MJ0366, the equivalent interactions involve residues located on the stretches (17–18) and (53–56). The interactions that stabilize the knot in the case of Rds3p appear to be more plastic than in the other two proteins, although there is some level of conservation for the regions (22–34) and (63–76). Figure 6 highlights these regions in the three-dimensional native structures of the three proteins.

IV. NON-NATIVE INTERACTIONS AND FOLDING EFFICIENCY

Here, we explore the potential role of non-native interactions in assisting the folding transition of knotted proteins. In order to

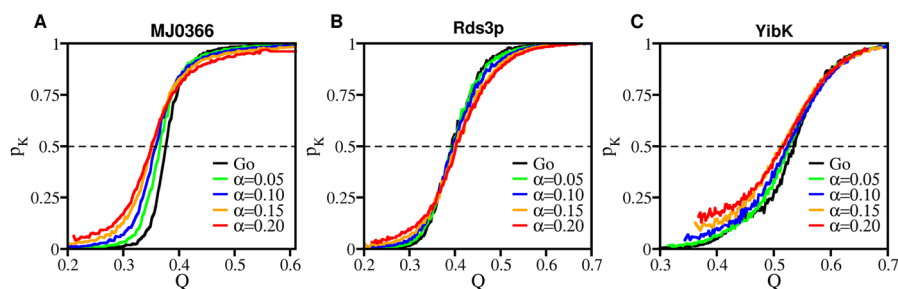


FIG. 4. Knotting probability. Dependence of the knotting probability, p_K , on the fraction of native contacts at T_k . The reaction coordinate for which the knotting probability is 0.5 is designated as Q_k .

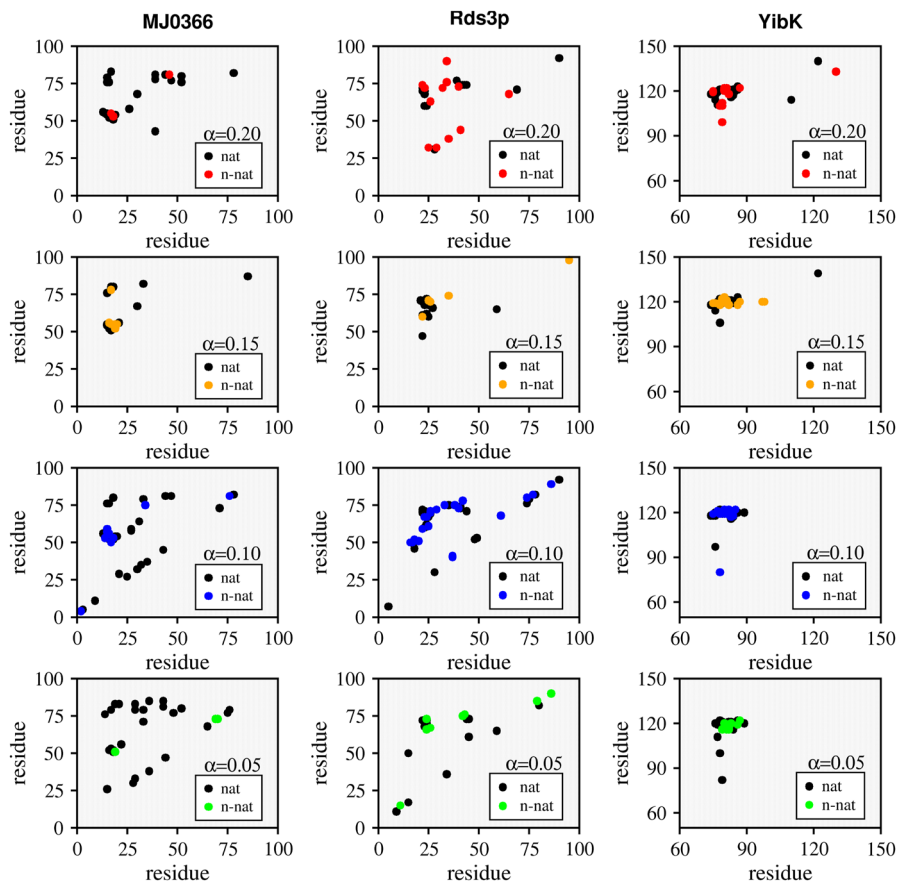


FIG. 5. Interactions that stabilize the knot. Contact maps showing the native (black dots) and non-native (colored dots) contacts that appear with the highest differential probability in the ensemble of knotted conformations with Q_k at T_k for proteins MJ0366 (left), Rds3p (middle), and YibK (right) for several strengths ($\alpha = 0.05, 0.10, 0.15,$ and 0.20) of non-native interactions.

do so, it is necessary to conserve the linear topology of the chain by means of MC simulations at fixed temperature. We assess the role of non-native interactions in assisting the formation of a fully folded conformation and, in the case of protein YibK, in assisting

the formation of a knotted conformation (as explained in the Sec. II, in this case, we are interested in the transition to the first knotted conformation independently of its degree of structural consolidation). Structural consolidation is assessed according to the criteria

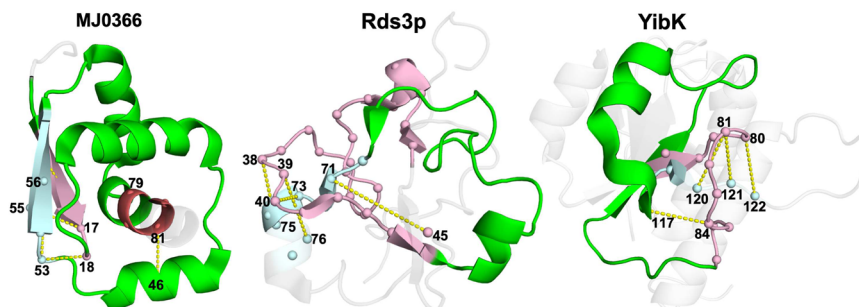


FIG. 6. Interactions that stabilize the knot. Three-dimensional structures of MJ0366 (a), Rds3p (b), and YibK (c). In the cartoon representation, only the knotted core is colored. Different colors are used to identify different chain segments of the knotted core containing residues involved in the most frequent native and non-native interactions that stabilize knotted conformations. Selected non-native interactions are represented as dotted yellow lines in the three-dimensional representation.

described in the Sec. II—a conformation is considered unfolded if it is unknotted and its Q is smaller than 0.25 for MJ0366, 0.26 for Rds3p, and 0.34 for YibK, and considered as folded if it is knotted and its Q is greater than 0.58 for MJ0366, 0.61 for Rds3p, and 0.72 for YibK. To carry out these tasks, we define *foldicity* as being the number of folding transitions occurring per million mcs (Mmcs) at each considered temperature, and *knoticity* as being the number of knotting transitions occurring per million mcs (Mmcs) at each considered temperature. Apart from considering the effects of non-native interactions added in a sequence-specific manner, in the case of protein YibK, we also consider the effect of specific non-native interactions, namely, those that have been identified in Sec. III D as stabilizers of knotted conformations, and those that have been identified in the study by Wallin *et al.*¹⁰ as being critical to achieve 100% folding efficiency. Since the addition of non-native interactions in a sequence-dependent manner can change the nature of the folding transition from two-state to downhill (Fig. 3), we consider the highest value of the parameter α (which controls the strength of non-native interactions) that conserves the two-state nature of the transition.

The results reported in Figs. 7(a)–7(c) show that the three model systems are able to fold under a pure native-centric Gō potential, but this is only observed below a certain threshold temperature that is well below the melting temperature, T_m , in the case of protein YibK. Furthermore, there is an optimal temperature at which folding proceeds at maximum efficiency. Thus, in the context of the adopted model and sampling protocol, non-native interactions are not necessary to efficiently fold a knotted protein, provided the temperature range of the simulations is adequate.

The results reported in Figs. 7(a)–7(c) also show the somewhat surprising result that the deep knot in protein Rds3p does not hamper its folding efficiency, which is actually higher than that of protein MJ0366 (featuring a shallow knot) across the considered temperature range. It was suggested, based on results obtained in the scope of lattice models, that local structure (i.e., order) inside the hydrophobic core strongly inhibits entanglements and knots.⁴² The better folding efficiency of Rds3p may thus be explained by the fact that it features a lower secondary structural content in the core in comparison to MJ0366.

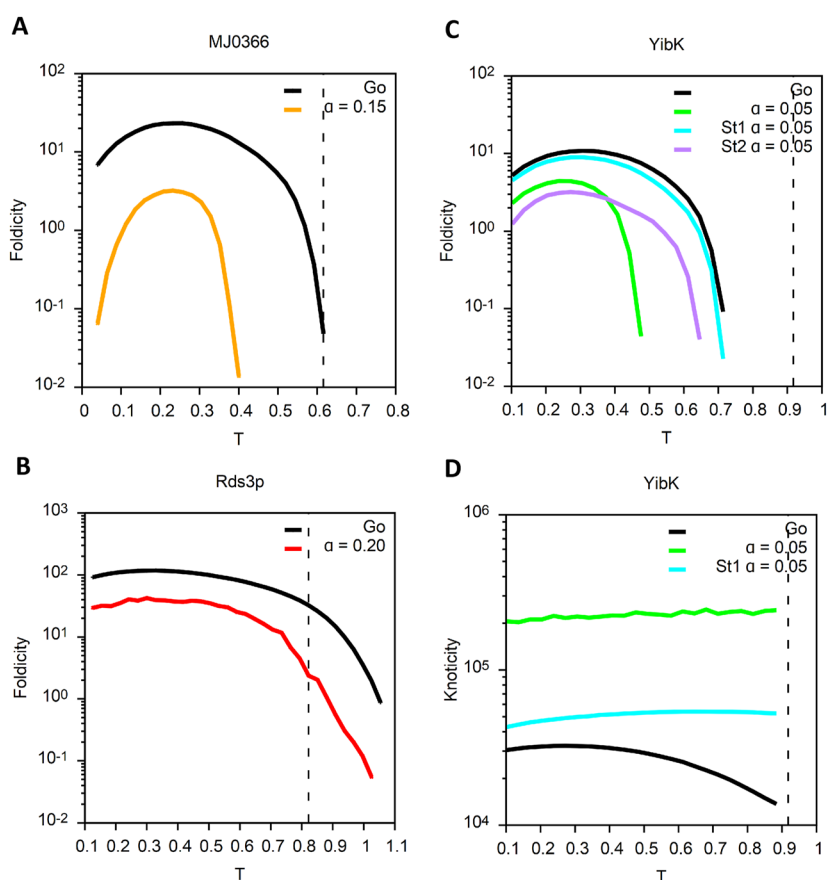


FIG. 7. Folding and knotting efficiency. Dependence of foldicity on temperature for proteins MJ0366 (a), Rds3p (b), and YibK (c). Also shown is the dependence of knoticity on temperature for protein YibK (d). In the case of YibK, St1 refers to the modified variant of the Gō potential that considers specific non-native interactions only between stretches (75–87) and (119–122), whereas St2 considers only the stretches (86–93) and (122–147). The vertical-dashed line indicates the melting temperature, T_m , for the Gō potential.

On the other hand, the addition of sequence-specific non-native interactions clearly hampers folding, with foldicity decreasing up to one order of magnitude in the considered model systems. This may be due to the fact that non-native interactions persist in nearly native (i.e., stable) conformations, thus precluding the formation of native ones; however, the latter must be established in order for the adopted folding criterion to be satisfied. Consequently, more MC steps are required for a simulation to finish, which reduces the total number of folding transitions observed per Mmcs. Focusing on the results obtained for YibK, we note that the non-native interactions established between the two stretches, (75–87) and (119–122), of the polypeptide chain, do not significantly alter the knotting efficiency relative to the native centric potential. However, foldicity decreases for the specific non-native interactions considered in Ref. 10. This likely resides in the specificities of the model and simulation protocol adopted in that case.

A completely different scenario is observed when one looks into the effects of non-native interactions on knoticity of YibK. In this case, the addition of non-native interactions contributes to increase the knotting efficiency, and the latter is substantial (about one order of magnitude larger across the considered temperature range) for the sequence-dependent potential. We observed that on average the first formed knotted conformations have $Q \approx 0.3$. The added non-native interactions are thus capable of energetically stabilizing partially folded conformations and compensating for the entropy loss due to knotting. We thus conclude that non-native interactions assist knotting but are not a pre-requisite to efficiently fold knotted proteins.

V. CONCLUSIONS

Determining the role of non-native interactions in folding dynamics, kinetics, and mechanisms is a classic problem in protein folding, which has been extensively investigated in the scope of molecular simulations of protein models with different resolutions. It has been suggested that non-native interactions participate directly,⁴³ or indirectly,⁴⁴ in the formation of the folding nucleus, that they accelerate or decelerate the folding rate, depending on whether they are more populated in the transition state or unfolded state,^{20,45} that they may be determinant for protein folding cooperativity,⁴⁶ that they interplay with native geometry to determine folding kinetics,^{47,48} and that they determine the folding pathway at the fine level of contact formation.⁴⁹

More recently, this problem has witnessed a renewed interest in light of the hypothesis that knotted proteins require them to fold efficiently. Determining the role of native and non-native interactions for the spontaneous folding of knotted proteins has been pointed out as an important fundamental question in the literature.⁸

Here, we explored the role of non-native interactions in the thermodynamics and kinetics of the folding transition of three knotted proteins featuring a trefoil knot in their native structure. In one protein, MJ0366, the knot is shallow, for the other two proteins, Rds3p and YibK, the knot is deep. We observed that non-native interactions can lower the free energy barrier to folding as observed for some unknotted proteins. The results from equilibrium simulations further indicate that non-native interactions energetically stabilize knotted conformations in a particular way. They are mainly non-local interactions being located in the vicinity of native

interactions. This observation suggests that non-native interactions operate as a system of redundant interactions that assist the knotting step. In line with this finding, the (non-equilibrium) kinetic simulations of deeply knotted protein YibK show that knotting efficiency, measured as the number of knotting transitions per million MC steps, increases substantially in the presence of sequence-specific non-native interactions. The knotted conformations are partially folded conformations with an average $Q \approx 0.3$. These results are consistent with those reported in Ref. 12 for protein AOTCase, featuring a deep trefoil knot, in which a sequence-dependent interaction potential was shown to substantially increase the knotting frequency in partially folded conformations, attaining a maximum value in conformations with $Q \approx 0.3$.

The results from the kinetic simulations also show that in the context of the present model and simulation protocol, native interactions alone can efficiently fold a knotted protein, provided the simulation temperature is below (or well below) the equilibrium melting temperature (for protein YibK). This is a critical finding of the present study, which may be invoked to rationalize the lack of folding success in studies of the same protein carried out by others.^{10,11} They also suggest that folding a protein with a deep knot is not necessarily less efficient than folding a protein with a shallow knot. Indeed, protein Rds3p folds more efficiently than MJ0366, presumably as a consequence of having a hydrophobic core exhibiting a lower degree of local structural order (i.e., featuring less secondary structural content).⁴² Overall, our results indicate that a sequence-specific interaction that accounts for non-native interactions may facilitate the transition from unknotted to knotted in partially folded conformations, but it may not facilitate the transition to the final folded and knotted conformations. This is likely a consequence of persistent non-native interactions in nearly folded conformations becoming highly stabilized in the temperature range where knotting is favored.

ACKNOWLEDGMENTS

Work supported by Grant No. UID/MULTI/04046/2020 Research Unit grant from FCT, Portugal (through BioISI). J.N.C.E. acknowledges financial support from FCT, Portugal, through PhD Grant No. SFRH/BD/144345/2019. P.F.N.F. and J.N.C.E. would like to acknowledge the contribution of the COST Action CA17139. A part of this work was performed on the computational resources of INCD (<http://www.incd.pt>) funded by FCT and UE under Project No. LISBOA-01-0145-FEDER-022153. Access was granted by FCT through Project No. 2022.26279.CPCA.A0.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

P.F.N.F. designed the research. J.N.C.E. performed the calculations. J.N.C.E. and P.F.N.F. analyzed the data. P.F.N.F. and J.N.C.E. prepared the figures and wrote the paper.

João N. C. Especial: Data curation (equal); Formal analysis (equal); Investigation (supporting); Methodology (equal); Software (lead); Writing – review & editing (supporting). **Patrícia F. N. Faisca:** Conceptualization (lead); Formal analysis (equal); Investigation (lead); Methodology (equal); Supervision (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (lead).

DATA AVAILABILITY

The data that support the findings of this study are available within the article.

APPENDIX: CALCULATION OF INTERACTION PARAMETERS ε_{ij}

Let us consider a protein formed by N residues. Its native structure is defined by the set of residue coordinates $\{\vec{r}_i^{nat}\}$, and its primary structure is defined by $\{\sigma_i\}$, with σ_i being the chemical species of residue $i = 1, \dots, N$. Since there are 20 different amino acids we may assign an integer, say $l = 1, \dots, 20$, to each chemical species and σ_i will be an integer in this range.

In a sequence-dependent (SD) potential, the interaction energy between residues i and j , at the bottom of the potential well, ε_{ij} , depends only on the chemical species, $\sigma_i = l$ and $\sigma_j = m$. Let us call it I_{lm} . All interaction energies thus belong to a 20×20 matrix and the well bottom interaction energy of contact $i - j$ is an element of this matrix: $\varepsilon_{ij} = I_{\sigma_i \sigma_j} = I_{lm}$. On the other hand, the pure Gō potential has a well bottom contact energy of $-\Delta_{ij}^{nat}$. To determine the SD potential that best approximates the Gō potential we must minimize the quadratic deviation,

$$\eta = \sum_{i,j}^N (\varepsilon_{ij} + \Delta_{ij}^{nat})^2. \quad (A1)$$

In the above sum, each native contact that involves chemical species l and m contributes a term $(I_{lm} + 1)^2$, and each non-native contact that involves the same chemical species contributes a I_{lm}^2 term. Let p_{lm} be the total number of native contacts that involve the pair of chemical species l and m , and q_{lm} be the total number of non-native contacts that involve those chemical species. For a particular protein (i.e., for a particular native structure and corresponding sequence) it is straightforward to evaluate p_{lm} and q_{lm} . Indeed, it suffices to take into account that (a) all pairs of non-bonded residues may potentially form a contact, (b) those pairs that are in contact in the native structure are native contacts, and (c) all pairs that may potentially form a contact in any other conformation and that are not native contacts are non-native ones. We can thus determine for each pair of non-bonded residues, i and j , their chemical species $\sigma_i = l$ and $\sigma_j = m$ and also if they form a native or a non-native contact should they come into contacting distance. Hence, taking into account the definitions of p_{lm} and q_{lm} , Eq. (A1) can be rewritten as

$$\eta = \sum_{l,m} p_{lm} (I_{lm} + 1)^2 + q_{lm} I_{lm}^2. \quad (A2)$$

Given that the I_{lm} are the degrees of freedom of this minimization problem, the global minimum is achieved when

$$\forall_{lm} : \frac{\partial \eta}{\partial I_{lm}} = 0, \quad (A3)$$

and thus, when

$$I_{lm} = -\frac{p_{lm}}{p_{lm} + q_{lm}}. \quad (A4)$$

This solution actually has a rather simple combinatorics interpretation: If l and m are two amino acid species (not necessarily distinct), $p_{lm} + q_{lm}$ is the total number of geometrically possible contacts that may occur between residues of these species and p_{lm} is the number of such contacts that are present in the native conformation. The ratio $p_{lm}/(p_{lm} + q_{lm})$ is thus the fraction of possible contacts that are native and the inter-amino acid contact energies are just the symmetric of these fractions.

REFERENCES

- J. S. Richardson, "β-sheet topology and the relatedness of proteins," *Nature* **268**, 495–500 (1977).
- P. Dabrowski-Tumanski, P. Rubach, D. Goundaroulis, J. Dorier, P. Sułkowski, K. C. Millett, E. J. Rawdon, A. Stasiak, and J. I. Sułkowska, "KnotProt 2.0: A database of proteins with knots and other entangled structures," *Nucleic Acids Res.* **47**(D1), D367–D375 (2018).
- J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Židek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, "Highly accurate protein structure prediction with AlphaFold," *Nature* **596**, 583–589 (2021).
- A. P. Perlinska, W. H. Niemyska, B. A. Gren, M. Bukowicki, S. Nowakowski, P. Rubach, and J. I. Sułkowska, "AlphaFold predicts novel human proteins with knots," *Protein Sci.* **32**, e4631 (2023).
- A. P. Perlinska, W. H. Niemyska, B. A. Gren, P. Rubach, and J. I. Sułkowska, "New 6₃ knot and other knots in human proteome from AlphaFold predictions," *bioRxiv:2021.474018* (2022).
- P. F. N. Faisca, "Knotted proteins: A tangled tale of structural biology," *Comput. Struct. Biotechnol. J.* **13**, 459–468 (2015).
- R. C. Lua and A. Y. Grosberg, "Statistics of knots, geometry of conformations, and evolution of proteins," *PLoS Comput. Biol.* **2**, e45 (2006).
- S. E. Jackson, A. Suma, and C. Micheletti, "How to fold intricately: Using theory and experiments to unravel the properties of knotted proteins," *Curr. Opin. Struct. Biol.* **42**, 6–14 (2017).
- S. E. Jackson, "Why are there knots in proteins?," in *Topology and Geometry of Biopolymers*, Contemporary Mathematics (AMS, 2020), Vol. 746, pp. 129–153 (2020).
- S. Wallin, K. B. Zeldovich, and E. I. Shakhnovich, "The folding mechanics of a knotted protein," *J. Mol. Biol.* **368**, 884–893 (2007).
- J. I. Sułkowska, P. Sułkowski, and J. Onuchic, "Dodging the crisis of folding proteins with knots," *Proc. Natl. Acad. Sci. U. S. A.* **106**, 3119–3124 (2009).
- T. Skrbčić, C. Micheletti, and P. Faccioli, "The role of non-native interactions in the folding of knotted proteins," *PLoS Comput. Biol.* **8**, e1002504 (2012).
- P. F. N. Faisca, R. D. M. Travasso, T. Charters, A. Nunes, and M. Cieplak, "The folding of knotted proteins: Insights from lattice simulations," *Phys. Biol.* **7**, 16009 (2010).
- M. A. Soler and P. F. N. Faisca, "Effects of knots on protein folding properties," *PLoS One* **8**, e74755 (2013).
- M. A. Soler, A. Nunes, and P. F. N. Faisca, "Effects of knot type in the folding of topologically complex lattice proteins," *J. Chem. Phys.* **141**, 025101 (2014).

- ¹⁶M. A. Soler, A. Rey, and P. F. N. Faisca, "Steric confinement and enhanced local flexibility assist knotting in simple models of protein folding," *Phys. Chem. Chem. Phys.* **18**, 26391–26403 (2016).
- ¹⁷J. Especial, A. Nunes, A. Rey, and P. F. N. Faisca, "Hydrophobic confinement modulates thermal stability and assists knotting in the folding of tangled proteins," *Phys. Chem. Chem. Phys.* **21**, 11764–11775 (2019).
- ¹⁸C. Paissoni, S. Puri, I. Wang, S.-Y. Chen, C. Camilloni, and S.-T. D. Hsu, "Converging experimental and computational views of the knotting mechanism of a small knotted protein," *Biophys. J.* **120**, 2276–2286 (2021).
- ¹⁹J. N. C. Especial and P. F. N. Faisca, "A specific set of heterogeneous native interactions yields efficient knotting in protein folding," *J. Phys. Chem. B* **125**(27), 7359–7367 (2021).
- ²⁰C. Clementi and S. S. Plotkin, "The effects of nonnative interactions on protein folding rates: Theory and simulation," *Protein Sci.* **13**(7), 1750–1766 (2004).
- ²¹C. Cheng, J. Wu, G. Liu, S. Shi, and T. Chen, "Effects of non-native interactions on frustrated proteins folding under confinement," *J. Phys. Chem. B* **122**(31), 7654–7667 (2018).
- ²²P. Mouro, V. de Godoi Contessoto, J. Chahine, R. Junio de Oliveira, and V. Pereira Leite, "Quantifying nonnative interactions in the protein-folding free-energy landscape," *Biophys. J.* **111**, 287–293 (2016).
- ²³J. Tsai, R. Taylor, C. Chothia, and M. Gerstein, "The packing density in proteins: Standard radii and volumes," *J. Mol. Biol.* **290**(1), 253–266 (1999).
- ²⁴A. M. Gronenborn, D. R. Filpula, N. Z. Essig, A. Achari, M. Whitlow, P. T. Wingfield, and G. M. Clore, "A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein g," *Science* **253**(5020), 657–661 (1991).
- ²⁵P. Alexander, S. Fahnstock, T. Lee, J. Orban, and P. Bryan, "Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: Why small proteins tend to have high denaturation temperatures," *Biochemistry* **31**(14), 3597–3603 (1992).
- ²⁶J. Kim, T. Keyes, and J. E. Straub, "Generalized replica exchange method," *J. Chem. Phys.* **132**, 224107 (2010).
- ²⁷L. Prieto and A. Rey, "Influence of the native topology on the folding barrier for small proteins," *J. Chem. Phys.* **127**, 175101 (2007).
- ²⁸B. Fernández del Río and A. Rey, "Behavior of proteins under pressure from experimental pressure-dependent structures," *J. Phys. Chem. B* **125**(23), 6179–6191 (2021).
- ²⁹M. Larriva, L. Prieto, P. Bruscolini, and A. Rey, "A simple simulation model can reproduce the thermodynamic folding intermediate of apoflavodoxin," *Proteins: Struct., Funct., Bioinf.* **78**(1), 73–82 (2010).
- ³⁰N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.* **21**(6), 1087–1092 (1953).
- ³¹Y. Sugita and Y. Okamoto, "Replica-exchange molecular dynamics method for protein folding," *Chem. Phys. Lett.* **314**(1–2), 141–151 (1999).
- ³²D. J. Earl and M. W. Deem, "Parallel tempering: Theory, applications, and new perspectives," *Phys. Chem. Chem. Phys.* **7**, 3910–3916 (2005).
- ³³J. N. C. Especial, A. Rey, and P. F. N. Faisca, "A note on the effects of linear topology preservation in Monte Carlo simulations of knotted proteins," *Int. J. Mol. Sci.* **23**, 13871 (2022).
- ³⁴J. D. Chodera, W. C. Swope, J. W. Pitera, C. Seok, and K. A. Dill, "Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations," *J. Chem. Theory Comput.* **3**(1), 26–41 (2007).
- ³⁵W. R. Taylor, "A deeply knotted protein structure and how it might fold," *Nature* **406**, 916–919 (2000).
- ³⁶V. Thiruselvam, T. Kumarevel, P. Karthe, S. Kuramitsu, S. Yokoyama, and M. N. Ponnuswamy, "Crystal structure analysis of a hypothetical protein (MJ0366) from *Methanocaldococcus jannaschii* revealed a novel topological arrangement of the knot fold," *Biochem. Biophys. Res. Commun.* **482**(2), 264–269 (2017).
- ³⁷A.-M. M. van Roon, N. M. Loening, E. Obayashi, J.-C. Yang, A. J. Newman, H. Hernández, K. Nagai, and D. Neuhaus, "Solution structure of the U2 snRNP protein Rds3p reveals a knotted zinc-finger motif," *Proc. Natl. Acad. Sci. U. S. A.* **105**(28), 9621–9626 (2008).
- ³⁸K. Lim, H. Zhang, A. Tempczyk, W. Krajewski, N. Bonander, J. Toedt, A. Howard, E. Eisenstein, and O. Herzberg, "Structure of the YibK methyltransferase from *Haemophilus influenzae* (HI0766): A cofactor bound at a site formed by a knot," *Proteins: Struct., Funct., Genet.* **51**(1), 56–67 (2003).
- ³⁹M. Gruebele, "Downhill protein folding: Evolution meets physics," *C. R. Biol.* **328**(8), 701–712 (2005).
- ⁴⁰J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, "Funnels, pathways, and the energy landscape of protein folding: A synthesis," *Proteins: Struct., Funct., Genet.* **21**(3), 167–195 (1995).
- ⁴¹S. S. Cho, P. Weinkam, and P. G. Wolynes, "Origins of barriers and barrierless folding in BBL," *Proc. Natl. Acad. Sci. U. S. A.* **105**(1), 118–123 (2008).
- ⁴²T. Wüst, D. Reith, and P. Virnau, "Sequence determines degree of knottedness in a coarse-grained protein model," *Phys. Rev. Lett.* **114**, 028102 (2015).
- ⁴³E. I. Shakhnovich, L. Li, and L. A. Mirny, "Kinetics, thermodynamics and evolution of non-native interactions in a protein folding nucleus," *Nat. Struct. Biol.* **7**, 336–342 (2000).
- ⁴⁴W. L. Treptow, M. A. A. Barbosa, L. G. Garcia, and A. F. Pereira de Araújo, "Non-native interactions, effective contact order, and protein folding: A mutational investigation with the energetically frustrated hydrophobic model," *Proteins: Struct., Funct., Bioinf.* **49**(2), 167–180 (2002).
- ⁴⁵A. Zarrine-Afsar, S. Wallin, A. M. Neculai, P. Neudecker, P. L. Howell, A. R. Davidson, and H. S. Chan, "Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding," *Proc. Natl. Acad. Sci. U. S. A.* **105**(29), 9999–10004 (2008).
- ⁴⁶M. Enciso and A. Rey, "Improvement of structure-based potentials for protein folding by native and nonnative hydrogen bonds," *Biophys. J.* **101**, 1474–1482 (2011).
- ⁴⁷Z. Zhang and H. S. Chan, "Competition between native topology and nonnative interactions in simple and complex folding kinetics of natural and designed proteins," *Proc. Natl. Acad. Sci. U. S. A.* **107**(7), 2920–2925 (2010).
- ⁴⁸H. Kroboth and P. F. N. Faisca, "Interplay between native topology and nonnative interactions in the folding of tethered proteins," *Phys. Biol.* **10**(1), 016002 (2013).
- ⁴⁹P. F. N. Faisca, A. Nunes, R. D. M. Travasso, and E. I. Shakhnovich, "Non-native interactions play an effective role in protein folding dynamics," *Protein Sci.* **19**(11), 2196–2209 (2010).