



OPEN

# Mechanical unfolding of a knotted protein unveils the kinetic and thermodynamic consequences of threading a polypeptide chain

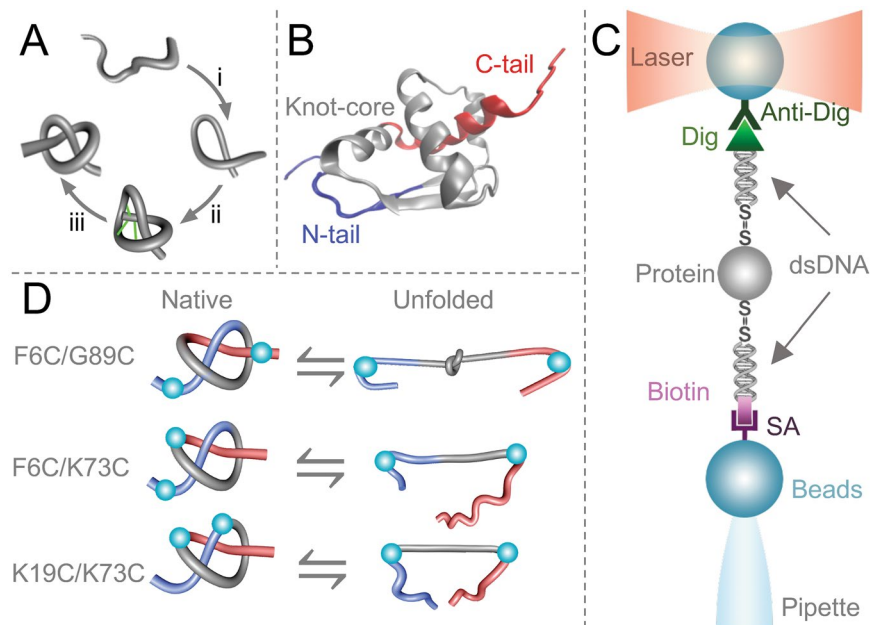
Maira Rivera<sup>1</sup>, Yuxin Hao<sup>2</sup>, Rodrigo A. Maillard<sup>2</sup> & Mauricio Baez<sup>1</sup>

Knots are remarkable topological features in nature. The presence of knots in crystallographic structures of proteins have stimulated considerable research to determine the kinetic and thermodynamic consequences of threading a polypeptide chain. By mechanically manipulating MJ0366, a small single domain protein harboring a shallow trefoil knot, we allow the protein to refold from either the knotted or the unknotted denatured state to characterize the free energy profile associated to both folding pathways. By comparing the stability of the native state with reference to the knotted and unknotted denatured state we find that knotting the polypeptide chain of MJ0366 increase the folding energy barrier in a magnitude close to the energy cost of forming a knot randomly in the denatured state. These results support that a protein knot can be formed during a single cooperative step of folding but occurs at the expenses of a large increment on the free energy barrier.

Knotted proteins have emerged as a unique class of proteins whose polypeptide chain is self-tied to form the shape of a knot. Over 600 crystallographic structures show the presence of knotted topologies<sup>1</sup>, and many other topological entanglements like slipknots and links has been identified in the protein data bank<sup>2,3</sup>. The folding mechanism of a knotted protein seems difficult because native interactions should create a set of coordinated and timely movements required to thread the polypeptide chain. The constrain resulting from threading limits the number of ways by which the polypeptide reaches the native state, thereby creating a thermodynamic burden for the configurational entropy reduction<sup>4-6</sup>. Early computational studies have suggested that specific weak interactions may be created during early or late steps of threading to help decreasing the topological barrier associated with the formation of a knot<sup>7-9</sup>. These contacts, referred as non-native contacts, occur transiently to guide the threading step, and are eliminated once the native structure of the protein is achieved (Fig. 1A). However, the role of non-native contacts during the folding of knotted proteins remains controversial<sup>6-11</sup> in part because no experimental evidence has been provided to support the studies of molecular dynamics performed with knotted proteins.

For deeply knotted proteins like  $\alpha$ -knotted methyltransferases YibK from *Haemophilus influenzae* and AOTCase from *Xanthomonas campestris*, the explicit contribution of proper non-native interactions to the protein energetics is likely required to form a knot efficiently<sup>7,8</sup>. For these proteins, it is suggested that non-native interactions favor the threading of one end of the chain through a loose loop during early stages of folding. Conversely, for small knotted proteins having shallow knots, such as MJ0366 from *Methanocaldococcus jannaschii*, the thermodynamic bias imposed by the optimization of native contacts could be enough to overcome a large topological barrier of folding<sup>12</sup>. Molecular dynamic simulations performed with native-centric potentials describe that the threading in MJ0366 occurs from the C-terminus via a twisted loop that is stabilized by native interactions<sup>9,12</sup>. Because a twisted loop takes place late in the folding pathway, threading occurs as a diffusive movement in a confined molecular environment. The application of more realistic force fields decrease the folding energy barrier but maintain the overall mechanism described above<sup>9,11,13</sup>. However, these molecular-level descriptions obtained from *in silico* folding simulations have been difficult to verify experimentally.

<sup>1</sup>Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, Santiago, Chile. <sup>2</sup>Department of Chemistry, Georgetown University, Washington, DC, 20057, USA. e-mail: [rodrigo.maillard@georgetown.edu](mailto:rodrigo.maillard@georgetown.edu); [mauricio.baez@ciq.uchile.cl](mailto:mauricio.baez@ciq.uchile.cl)



**Figure 1.** Structure of MJ0366 and experimental design. **(A)** To form a trefoil knot, the polypeptide chain first forms the threading loop (i) that is threaded by one of the protein's termini (ii). Transient contacts (green lines), formed during threading must break to form the native knot (iii). **(B)** Structure of MJ0366 (PDB 2EFV) showing the knot core of 66 residues (gray), which is threaded by an amino terminal tail (blue) and a carboxy terminal tail (red). The protein representation was generated using VMD version 1.9.4<sup>58</sup> (<https://www.ks.uiuc.edu/Research/vmd/>). **(C)** Optical-tweezers experimental setting. A double cysteine mutant is covalently attached to two dsDNA handles. The system is held between two polystyrene beads coated with either streptavidin (SA) or anti-digoxigenin antibodies (Anti-Dig). **(D)** To unfold MJ0366 and preserve the knot in the unfolded state, we applied force from the N- and C-tails of the protein (construct F6C/G89C). To simultaneously unfold the protein and untie the knot, we applied force either from the N-tail and one residue located in the core of the knot (construct F6C/K73C), or from two residues located in the core of the knot (construct K19C/K73C). Cyan beads indicate the pulling points in the construct. The simplified representations of knots were generated by using KnotPlot (<http://www.knotplot.com>).

Several experiments performed with chaotropic agents show the presence of intermediate states during the folding of  $\alpha/\beta$ -knot methyltransferases YbeA and YibK<sup>14,15</sup>, the artificially knotted protein 2ouf-knot<sup>16</sup>,  $\alpha$ -haloacid-dehalogenase DehI<sup>17</sup>, and Ubiquitin carboxyl-terminal hydrolase isozymes (UCH) L1<sup>18,19</sup> and L3<sup>20</sup>. However, the interpretation of such kind of data is difficult because bulk spectroscopic methodologies cannot directly determine whether the folding reaction started from an unknotted or knotted denatured state, or when a knot is formed during their folding mechanism. Moreover, indirect observations indicate that knots seem to be prevalent topologies in the denatured state of deeply knotted proteins<sup>21,22</sup>. To overcome these problems, *in vitro* translation and mechanical unfolding experiments designed specifically to untie the polypeptide chain of YibK, YbeA and UCH-L1 showed that a decrease of the folding rate constant when threading the polypeptide chain is mandatory to reach the native state<sup>23,24</sup>. Besides the intramolecular non-native interactions evolved to overcome the free energy barrier of knotted proteins, it has been proposed that the cellular machinery, like chaperonins and the ribosome, can assist the folding of knotted proteins *in vivo* by promoting the formation of a knot in confined spaces<sup>23,25,26</sup>, by stabilizing key intermediates and establishing new folding routes<sup>10,26–28</sup>, or by modulating the collapse by hydrophobic interactions<sup>29</sup>. These results support that knotted proteins must overcome a topological energy barrier derived from the threading of the polypeptide chain. However, a complete description of the kinetic and thermodynamic consequences related with this process has not been determined experimentally.

Here, we use optical tweezers to study the folding mechanism of MJ0366, a homodimer containing a  $3_1$  or trefoil shallow knot into each monomer. Based on the experimental approach reported by Ziegler *et al.* in 2016<sup>24</sup>, we designed three different pulling geometries to either untie or tie the knot in the unfolded state. The mechanical folding or unfolding of MJ0366 was consistent with a simple two-state mechanism for all pulling geometries used in this study. This behavior allowed us to compare the energy cost of forming a knot during the folding of MJ0366 with the energy cost of forming a knot randomly in its unfolded state. We find that knotting the polypeptide chain of MJ0366 strongly increase the energy barrier of folding in a magnitude close to the energy cost of forming a knot randomly in the unfolded state. These observations are discussed in the context mechanism of MJ0366 proposed by *in silico* molecular dynamic and other knotted proteins.

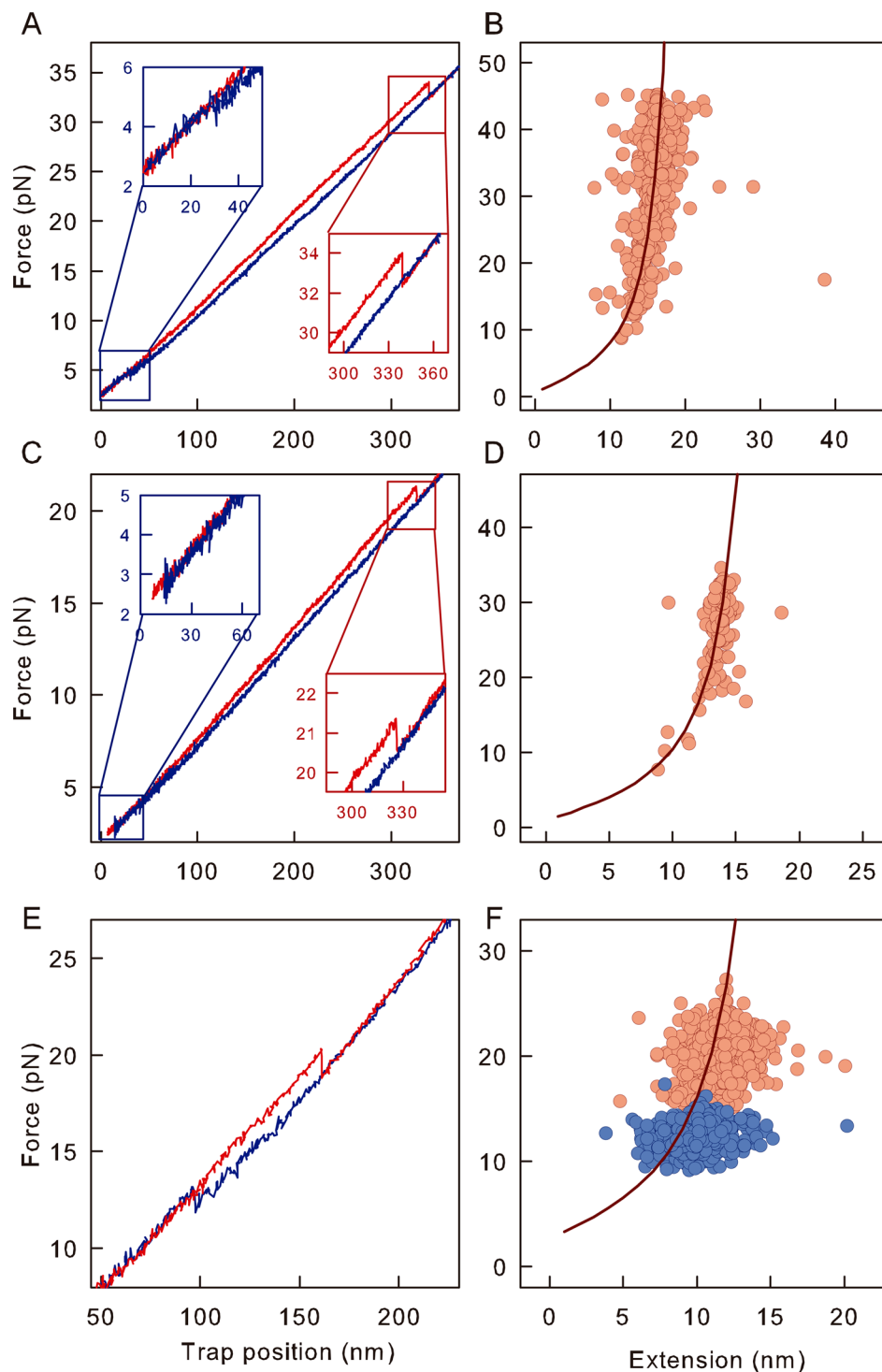
## Results

MJ0366 is a homodimer containing a  $3_1$  knot into each monomer (Fig. 1B). The knot core of MJ0366 spans 66 residues (Lys11 to Asn76; gray in Fig. 1B) which is threaded by an N-terminal tail of 10 residues (N-tail; blue in Fig. 1B) and a C-terminal tail of 6 residues (C-tail; red in Fig. 1B)<sup>1,30</sup>. To probe the folding barrier associated with the formation of a knot in MJ0366, we used optical tweezers to mechanically unfold MJ0366 from specific points of its structure (Fig. 1C). This experimental strategy allowed us to compare the free energy profiles of a protein construct designed to preserve the knot in the unfolded state (Fig. 1D, construct F6C/G89C) with protein constructs designed to untie the knot in the denatured state (Fig. 1D, construct F6C/K73C and K19C/K73C). Since MJ0366 has a cysteine at position 81<sup>30</sup>, the three constructs used in this study were characterized in the background of the mutant C81A. Like the wild type protein, all constructs were homodimers in solution (Supplementary Fig. S1A) characterized by indistinguishable unfolding stabilities determined by chemical denaturation under equilibrium conditions (Supplementary Fig. S1B and Table S1).

For all protein constructs, force-extension curves obtained at constant pulling velocity display one unfolding transition, indicating the cooperative unfolding of a single MJ0366 subunit irrespectively of the pulling direction (Fig. 2A,C and E). The application of the Worm-Like Chain (WLC) model<sup>31</sup> to unfolding transitions plots allow us to determine the contour length upon unfolding,  $L_c$  (Supplementary Fig. S2). The experimental  $L_c$  values for F6C/K73C and K19C/K73C (23 nm and 21 nm, respectively), are in quantitative agreement with the expected molecular extension between the pulling points in the proteins (24 nm and 20 nm, respectively; Fig. 2B,D), indicating complete unfolding during their mechanical perturbation. For F6C/G89C, which is mechanically pulled from the N- and C-tails, we obtained  $L_c = 23 \pm 2$  nm (Fig. 2F and Supplementary Fig. S2A), a value that is 7 nm shorter than the one expected for the full-length protein,  $L_{c, \text{theoretical}} = 30.3$  nm (84 residues between the pulling points). Similar differences of 4–6 nm have been reported in single molecule studies designed to tight a trefoil ( $3_1$ ) knot in the unfolded state<sup>24,32–34</sup>. It is important to highlight that MJ0366 is a homodimer in solution, and no information regarding monomer association or dimer dissociation steps can be extracted from our experiments due to the tethering constrains used to mechanically unfold MJ0366 (Supplementary Fig. S3). However, the value of  $L_c$  corroborates the presence of a knot in its denatured state, and directly shows that the interface contacts of the dimer are not required to form a knot in each monomer. In agreement, small conformational changes have been reported for the wild type<sup>35</sup> or double cysteine dimer when their interface is disrupted by chemical denaturation (Supplementary Fig. S1B).

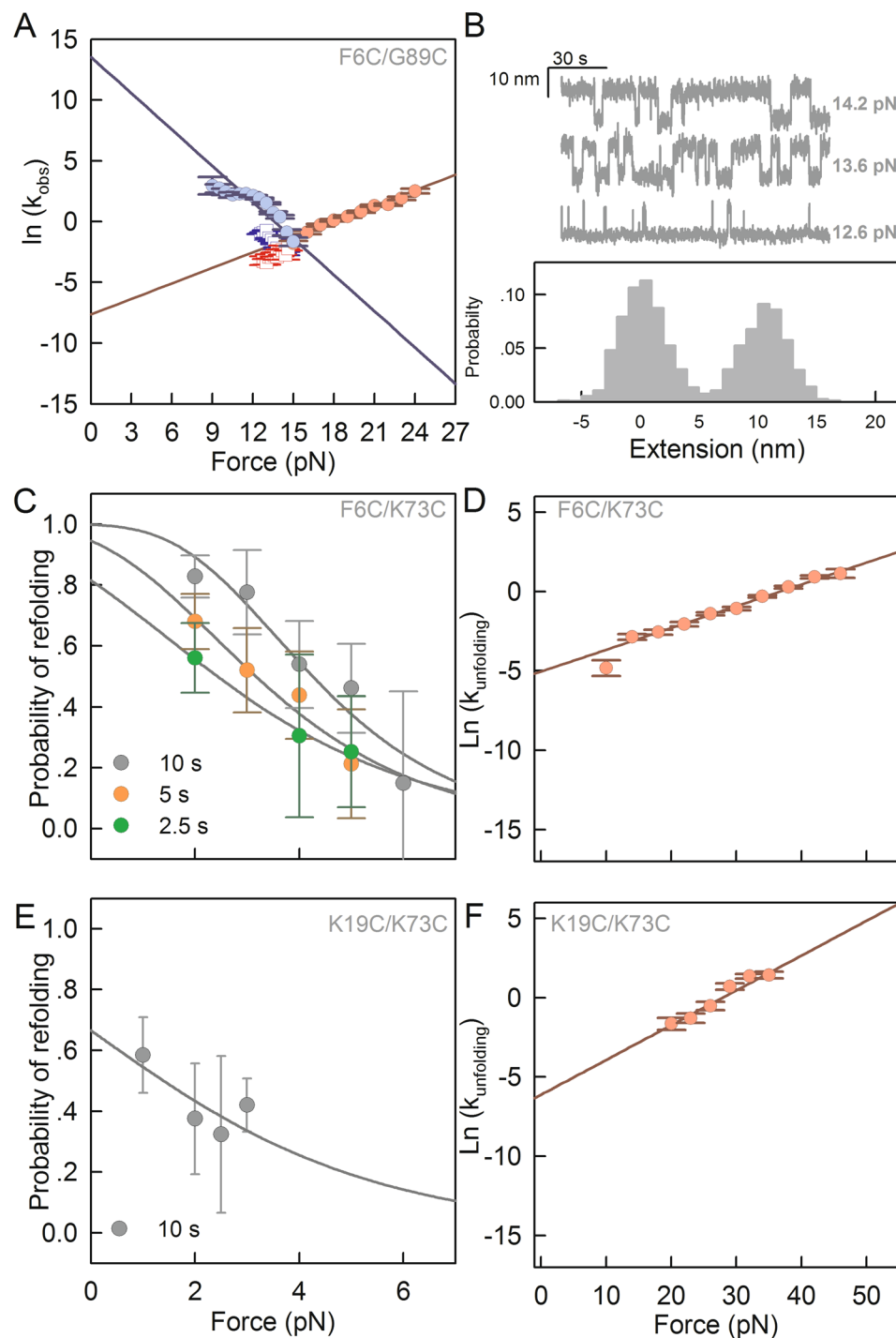
The force-dependent rate constants,  $k(F)$ , were determined by transforming force distributions (Supplementary Fig. S4) using methods developed by Dudko<sup>36</sup>. Dudko *et al.*<sup>37</sup>, also developed an analytical model that incorporates the Kramer's theory to Bell's model to describe how the energy barriers and the position of the transition state are affected by force. Experimentally this is reflected in non-linear dependencies of the natural logarithm of the observed kinetic rate constants with force. However, in our case the dependence of the natural logarithm of the kinetic constants with force was linear (Fig. 3A,D and F). Therefore, we fitted our data using the model described by Bell<sup>38</sup>. The Bell model assumes the existence of a single kinetic barrier that the protein needs to energetically overcome to unfold or refold, and that the position of the transition state, ( $\Delta x^\ddagger$ ), does not depend on force. For F6C/G89C, we obtained an unfolding rate constant at zero force ( $k_U^0$ ) of  $4 \cdot 10^{-4} \text{ s}^{-1}$  and a distance to the transition state from the folded state ( $\Delta x_{F \rightarrow U}^\ddagger$ ) of  $1.8 \pm 0.5$  nm (Fig. 3A, red circles and Table 1). Similarly, we obtained a refolding rate constant at zero force ( $k_F^0$ ) of  $\sim 10^5 \text{ s}^{-1}$  and a distance from the unfolded state to the transition state ( $\Delta x_{U \rightarrow F}^\ddagger$ ) of  $4.1 \pm 0.9$  nm (Fig. 3A, blue circles and Table 1). We corroborated the results obtained from unfolding and refolding force distributions by conducting time-dependent extension traces at various constant forces (Fig. 3B). At any given constant force, F6C/G89C fluctuates between a compacted or folded state and an extended or unfolded state. The observed change in extension between these two states at 13.6 pN is  $\sim 10$  nm (Fig. 3B), which matches the end-to-end distance predicted by the WLC at that force (Fig. 2F). Moreover, the unfolding and refolding rate constants determined from time-dependent extension traces (Fig. 3A, squares and S5) agree with those obtained from force distributions (Fig. 3A, circles).

Having characterized the folding energy profile of F6C/G89C, we used the constructs F6C/K73C and K19C/K73C (Fig. 1D) to investigate the effect of threading the polypeptide chain on folding rate constants and protein stability. Because force-extension curves of the constructs F6C/K73C and K19C/K73C did not reveal clear refolding transitions (Fig. 2A,C, blue traces), we determined refolding rate constants from the probability of refolding<sup>39</sup>,  $P_R$ , at varying refolding forces (2–6 pN) and increasing waiting times ( $\Delta t = 2.5, 5$  and  $10$  s) (Fig. 3C,E). The fit to the integrated Bell equation<sup>38</sup> to the values of  $P_R$  results in  $k_F^0$  values in the order of  $10^{-1} \text{ s}^{-1}$  (Fig. 3C,E and Table 1). Notably, the refolding rate constants for both constructs, which are untied using two different pulling geometries, are six orders of magnitude slower than the folding rate estimated for the construct with the knot in its unfolded state (F6C/G89C,  $k_F^0 \sim 1 \cdot 10^5 \text{ s}^{-1}$ ). These results indicate that the topology of the unfolded state is important for the folding energy barrier which increases significantly when the reaction starts from an unknotted unfolded state. In contrast, the unfolding rates  $k_U^0$  range between  $10^{-3} \text{ s}^{-1}$  for F6C/K73C and K19C/K73C to  $10^{-4} \text{ s}^{-1}$  for F6C/G89C (Fig. 3A,D,F, and Table 1). A similar trend is observed in the unfolding rate constants extrapolated at zero concentration of guanidinium hydrochloride (GdnHCl), ranging between  $10^{-2} \text{ s}^{-1}$  for F6C/K73C and K19C/K73C to  $10^{-3} \text{ s}^{-1}$  for F6C/G89C (Fig. S1C and Supplementary Table S1). It is notable that the relative difference in unfolding rate constants between constructs that untie the knot (F6C/K73C and K19C/K73C) vs. the construct that tightens up the knot (F6C/G89C) is the same irrespectively of the technique used. By using the values of  $k_F^0$  and  $k_U^0$  at zero force for F6C/K73C and K19C/K73C, we calculated a free energy difference  $\Delta G_{\text{F6C/K73C}}$  of  $2.7 \pm 1$  kcal/mol and  $\Delta G_{\text{K19C/K73C}} = 2.4 \pm 0.9$  kcal/mol (Table 1). These values are  $10 \pm 0.5$  kcal/mol lower than the thermodynamic stability estimated for the construct with a knot in its unfolded state determined from the quotient between the refolding and unfolding rate constants or directly by the application of the Crooks fluctuation theorem (F6C/G89C,  $\Delta G_{\text{F6C/G89C}} = 13 \pm 1$  kcal/mol, supplementary Fig. S6).



**Figure 2.** Force-extensions curves of the mechanical unfolding of MJ0366. Unfolding (red) and refolding (blue) trajectories for F6C/K73C (A), K19C/K73C (C) and F6C/G89C (E). Inset: Zoomed-in plot of the unfolding and refolding transitions. WLC analyses of plots of force vs. changes in extension upon unfolding for F6C/K73C (B), K19C/K73C (D) and F6C/G89C (F). Graphics were generated, using SigmaPlot version 10.0 (Systat; <https://systatsoftware.com/products/sigmaplot/>).

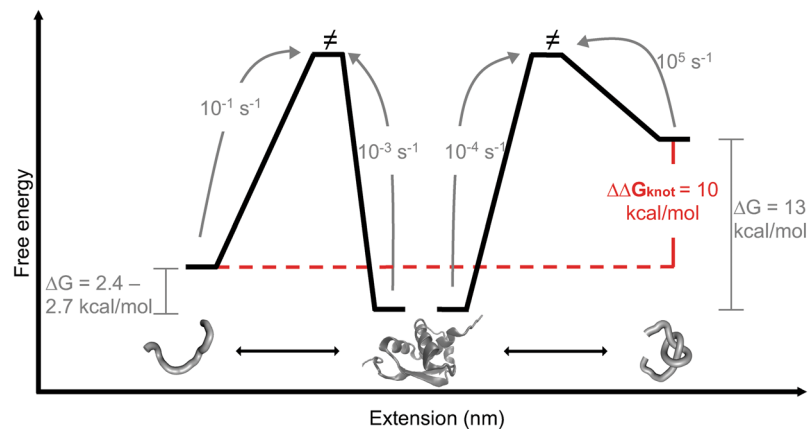
Figure 4 summarizes the thermodynamic and kinetic characterization of MJ0366. It should be noted that the energy of the native states, for the different constructs, is equivalent in the free energy profile of Fig. 4. This normalization is supported by two reasons; the unfolding kinetics is about  $10^{-3} - 10^{-4} \text{ s}^{-1}$  for the three constructs (Table 1) indicating that the free energy barrier for the mechanical unfolding of MJ0366 does not depend on the



**Figure 3.** Folding and unfolding kinetics of MJ3066 constructs. (A) Force-dependent unfolding (red symbols) and refolding (blue symbols) rate constants for F6C/G89C. The semilogarithmic dependence of rate constants were determined either from the probabilities distribution of forces (circles) or from experiments performed at constant forces (squares). (B) Trajectories of extension vs. time for F6C/G89C (Top) and population distribution of folded and unfolded states at  $F_{1/2}$  (Bottom). Refolding probability ( $P_R$ ) as a function of force for F6C/K73C (C) and K19C/K73C (E). The waiting times for each curve are indicated in each figure. Force-dependent unfolding rate constants calculated for F6C/K73C (D) and K19C/K73C (C). Recovered kinetic parameters are shown in Table 1. In experiments performed at constant force (squares in A), error bars represent the fitting error calculated from the exponential fit to the dwell time distribution at each constant force. In experiments performed at constant velocity (circles in A, D and F), error bars correspond to the counting error associated with the frequency of each bin in the force distributions showed in Supplementary Fig. S4. Error bars in C and E correspond to the counting error of the total number of the observed unfolding events. This figure was generated using SigmaPlot version 10.0 (Systat; <https://systatsoftware.com/products/sigmaplot/>).

Constructs	Unfolding kinetics		Refolding kinetics		Stability	
	$k_U^0$ (s <sup>-1</sup> ) <sup>a,d</sup>	$\Delta x_U^\ddagger$ (nm)	$k_F^0$ (s <sup>-1</sup> )	$\Delta x_F^\ddagger$ (nm)	$\Delta G_{kin}$ (kcal/mol) <sup>b</sup>	$\Delta G_{crooks}$ (kcal/mol)
F6C/G89C	$4 \cdot 10^{-4}$ ( $3 \cdot 10^{-4}$ , $1 \cdot 10^{-3}$ )	$1.8 \pm 0.5$	$8 \cdot 10^5$ ( $4 \cdot 10^6$ , $3 \cdot 10^5$ ) <sup>a,d</sup>	$4.1 \pm 0.9$ <sup>a,d</sup>	$12.7 \pm 0.9$	$13 \pm 1$
F6C/K73C	$6 \cdot 10^{-3}$ ( $2 \cdot 10^{-3}$ , $2 \cdot 10^{-2}$ )	$0.6 \pm 0.1$	$0.6 \pm 0.2$ <sup>c</sup>	$2.2 \pm 0.2$ <sup>c</sup>	$2.7 \pm 1$	—
K19C/K73C	$2 \cdot 10^{-3}$ ( $7 \cdot 10^{-4}$ , $4 \cdot 10^{-3}$ )	$0.9 \pm 0.2$	$0.11 \pm 0.04$ <sup>c</sup>	$1.3 \pm 0.8$ <sup>c</sup>	$2.4 \pm 0.9$	—

**Table 1.** Kinetic and thermodynamic parameters. <sup>(a)</sup>Data obtained using Bell's kinetic model. <sup>(b)</sup> $\Delta G_{kin} = -RT \cdot \ln(k_U^0/k_F^0)$  at 23 °C. <sup>(c)</sup>Data obtained using refolding probability  $P_R = 1 - e^{-\Delta t k_F^0 e^{-F \Delta x_F^\ddagger / k_B T}}$ . <sup>(d)</sup>Values a two-sigma confidence interval. Parameters obtained by using a Bell fitting to the rate constant obtained by constant force experiments or from the force probability distribution.



**Figure 4.** Free energy landscape of MJ0366. The free energy landscape of MJ0366 is constructed by using the thermodynamic and kinetic parameters obtained from this study. When the knot was tightened up in the unfolded state, the calculated stability is ~13 kcal/mol. In contrast, the free energy difference decreases to 2.4–2.7 kcal/mol when refolding begins from the unknotted denatured state. The value of  $\Delta\Delta G_{knot}$  correspond to the energy cost to form a knot in the denatured state. The arrows represent the folding and unfolding rate constants determined from the denatured or native state respectively. The protein representations were produced with VMD version 1.9.4<sup>58</sup> (<https://www.ks.uiuc.edu/Research/vmd/>). The images of the unfolded states were generated using KnotPlot (<http://www.knotplot.com>).

pulling geometry and, chemical denaturation experiments indicate that all constructs have a similar stability, ranging from 5.2 to 6.0 kcal/mol (Supplementary Fig. S1B, Table S1) with similar unfolding rate constants extrapolated at zero molar of GdnHCl (Supplementary Fig. S1C and Table S1). Therefore, we surmise that the energy difference between F6C/G89C and either F6C/K73C or F19C/K73C is originated by the destabilization of the unfolded state, the state where the knot is tightened or untied in the setup of the optical tweezers. As shown in Fig. 4, the normalization of the native state's energy for the different protein constructs reveals that the  $3_1$  knot increases the energy of the unfolded state by  $10 \pm 0.5$  kcal/mol ( $\Delta\Delta G_{knot}$ ). Since  $\Delta\Delta G_{knot}$  is calculated between two unfolded states on the same protein, the energy cost is mainly entropic and comparable to the entropic cost of knotting a polypeptide. In other words, the value of  $\Delta\Delta G_{knot}$  likely represents the energy cost paid by the unfolded state to form a knot just by chance. This value is similar to the knotting cost calculated for the artificial knotted protein Arc-L1-Arc<sup>34</sup>.

Besides the thermodynamic stabilization of the denatured state by a knot, the kinetic analysis represented in Fig. 4 suggests two folding pathways for MJ0366; a slow-folding pathway involving the formation of a knot ( $k_F^{0, unknotted} = 0.1$  s<sup>-1</sup>, right side of Fig. 4) and a fast-folding pathway starting from a denatured state already knotted ( $k_F^{0, knotted} \sim 1 \cdot 10^5$  s<sup>-1</sup>, left side of Fig. 4). Therefore, although the protein can fold very fast from a knotted unfolded state, that fast pathway is unlikely to occur because the large thermodynamic bias towards the unknotted denatured state ( $\Delta\Delta G_{knot} = 7.7$  to  $9$  kcal/mol). The slow-folding rate  $k_F^{0, unknotted} = 0.1$  s<sup>-1</sup> seems unusually slow for a small single domain protein of 92 residues like MJ0366, while the fast-folding rate  $k_F^{0, knotted} \sim 10^5$  s<sup>-1</sup> is similar to the single-chain version of the repressor ARC, a small monomeric artificial homologue of MJ0366 devoid of a knot in its hydrophobic core<sup>40</sup>. ARC folds very fast at about  $1 \cdot 10^4$  s<sup>-1</sup><sup>34,40</sup>. Therefore, the formation of a trefoil knot during the folding of MJ0366 represent a large constrain that strongly increase the folding barrier ( $\Delta\Delta G_{knot}^{\ddagger}$ ) by 8.4–9.4 kcal/mol ( $\Delta\Delta G_{knot}^{\ddagger} = -RT \cdot \ln(k_F^{0, unknotted}/k_F^{0, knotted})$ ).

## Discussion

The large energy barrier increment associated with the threading in MJ0366 contrasts with the effect reported for the folding barrier of more complex proteins like YibK, YbeA and UCH-L1. It was observed in *in vitro* translation experiments that the folding kinetics of YibK and YbeA is reduced 0.7 to 10 times compared to the rates determined starting from their knotted unfolded state<sup>14,23</sup>. Similarly, mechanical unfolding experiments showed that the formation of a 5<sub>2</sub> knot in the structure of UCH-L1 decreases the folding rate constant at zero force by 10-fold<sup>24</sup>. In terms of free energy, the formation of a knot increases the value of  $\Delta\Delta G_{\text{knot}}^{\ddagger}$  between 0.7 and 2 kcal/mol. In comparison, the formation of a shallow knot in MJ0366 involves a large energy cost of  $\Delta\Delta G_{\text{knot}}^{\ddagger}$  of 8.4–9.4 kcal/mol. At first glance, this result seems intriguing because YibK, YbeA and UCH-L1 require the formation of deep and complex knots during folding, respectively, and therefore are expected to have similar or higher energy barriers when compared with a shallow 3<sub>1</sub> knotted protein like MJ0366. However, this comparison assumes that deep and complex knotted proteins have evolved similar strategies to overcome the intrinsic topological barrier of proteins with shallow and simpler knots.

Knotted proteins like YibK, YbeA and UCH-L1 display a broad repertoire of folding intermediates in chemical and mechanical denaturation studies<sup>14,15,18</sup>. The presence of intermediates suggests that the overall energy cost of forming a knot could be paid by a progressive stabilization of partially folded structures with the drawback of the formation of off-pathway intermediates, like in the case of mechanical and chemical perturbations in UCH-L1<sup>18,24</sup>. Additionally, since both methyltransferases need to thread a large extension of a chain and UCH-L1 forms a complex 5<sub>2</sub> knot, it is plausible that the contribution of non-native contacts effectively reduces their folding energy barriers. In support of this interpretation, folding simulations performed with a native centric potential revealed that the fraction of successful folding events is rather low for self-tying deeply knotted proteins, but it is significantly increased by the addition of specific non-native contributions<sup>7–9</sup>. Thus, it is possible that mutations on critical residues required for non-native interactions could increase the free energy barrier associated with the threading step, or alternatively increase the population of off-pathway and aggregating-prone intermediates<sup>18</sup>.

In contrast to observing several folding intermediates<sup>14–20,24</sup>, our data for MJ0366 are consistent, with a two-state folding mechanism. For example, for all constructs we always observed a single distribution of both unfolding forces (Supplementary Fig. S4A) and molecular extension (Fig. 3B) when the protein is refolded from the knotted denatured state. A similar two-state folding mechanism is deduced when refolding occurs from the unknotted denatured state. In this case, the unfolding transitions produce a single force and distance distributions irrespective of the waiting time used to refold the protein (Fig. 3C and E). The absence of detectable intermediates is consistent with the cooperative unfolding of each monomer of MJ0366 reported under equilibrium conditions<sup>35</sup>. Therefore, our experiments provide an opportunity to compare the thermodynamic stability with the values estimated from chemical denaturation experiments monitored with circular dichroism (Supplementary Fig. S1). When the kinetic and thermodynamic parameters derived from bulk and single molecule experiments are compared, there is a good agreement between the constructs designed to unite the knot in the denatured state of MJ0366 (F6C/K73C and K19C/K73C) than the construct designed to tie the knot in the denatured state of MJ0366 (K6C/K89C). For example, although the chemical stabilities for F6C/K73C and K19C/K73C differ by 2 kcal/mol with respect to their mechanical stabilities, this difference is about 8 kcal/mol for the construct designed to tie the knot (Table 1 and Supplementary Table S1). Considering that the stability calculated by chemical denaturation experiments performed under equilibrium conditions is not only similar among all constructs (Supplementary Table S1) but also similar to the stability calculated for the constructs designed to untie the knot with optical tweezers (F6C/K73C and K19C/K73C, Table 1), we surmise that chemical denaturation effectively unties the knot in the denatured state of MJ0366.

There are several molecular dynamic simulations that explain the mechanism of threading of MJ0366<sup>9,11,12,25,41–43</sup>. In general, the application of native centric potential support the formation of an on-pathway intermediate, guided by native interactions, needed to organize the structure for a further high energy step required to form the knot<sup>9,12</sup>. Experimentally, our single molecule experiments and those obtained by chemical denaturation<sup>35</sup> were unable to detect an intermediate, but its formation cannot be discarded since its structure is predicted to be unstable and very short-lived<sup>12</sup>. Nevertheless, a key aspect of these mechanisms is the ability of MJ0366 to form a knot without an explicit energetic contribution of non-native contacts, albeit at the expense of a large energy barrier<sup>9,12</sup>. This barrier, entropic in nature, arises from a limited number of accessible folding routes leading to a competition between threading and the packing of the polypeptide chain<sup>12</sup>. This feature could explain the large energy cost associated with the formation of a knot during the folding of MJ0366 ( $\Delta\Delta G_{\text{knot}}^{\ddagger} = 8.4–9.4$  kcal/mol) which is comparable to the entropic cost to form a knot in the unfolded state of MJ0366 ( $\Delta\Delta G_{\text{knot}} = 10 \pm 0.5$  kcal/mol). Structure-based molecular dynamic simulations on MJ0366 predict a peak in the free energy as a function of global parameter such as number of native contacts (Q). This peak correlates with threading of the polypeptide chain<sup>12</sup>. In agreement with the computational evidence, we observed experimentally a strong kinetic effect on the folding rate constants between unknotted and knotted unfolded proteins. Thus, if the contribution of non-native contacts for the folding of MJ0366 (for example by stabilization of the transition state) was significant, then it should be expected a lower value of  $\Delta\Delta G_{\text{knot}}^{\ddagger}$  with respect to the unguided formation of knot in the denatured state of MJ0366. Therefore, we postulate that the formation of a trefoil knot in MJ0366 is an unavoidable topological constrain, whose energy cost is transferred to the folding barrier without the assistance of non-native contacts.

Considering that a knot in the unfolded state of MJ0366 accelerates its folding by several orders of magnitude, chaperone encapsulation seems a plausible mechanism to increase its apparent folding rate constant<sup>25,43</sup>. Experiments with chaperonins like GroEL–GroES show a 20–40-fold acceleration of knotting of some proteins like YibK and YbeA relative to the folding initiated from their already knotted denatured state, respectively<sup>23</sup>. Although a molecular mechanism has not been yet determined, molecular dynamic simulations carried out in cylinders that mimic chaperonins suggest that a confined environment may enhance self-tying by increasing the probability of forming a knot randomly and by promoting alternative folding routes<sup>25,43</sup>. Nevertheless, simulations

performed with native centric potential show that a rapid packing or compaction of MJ0366 increase its folding barrier by increasing the formation of topological trap<sup>12</sup>. In this respect, a recent study performed with lattice and off-lattice models of knotted proteins, the authors explore the role of intermolecular hydrophobic interactions established early during the chaperonin cycle<sup>29</sup>. They find that strong hydrophobic interactions between protein and chaperonin cage counterbalance a (bad) side effect of steric confinement while moderate hydrophobic intermolecular interactions with the cage lead to a significant enhancement of knotting probability in relation to bulk conditions while simultaneously moderating the effect of steric confinement.

The process of co-translational folding also has been proposed to aid during the formation of knots in proteins<sup>10,27,41</sup>. However, by using a minimalist model of the ribosome exit tunnel, Chwastyk and Cieplak reported subtle effects on the folding mechanism of MJ0366<sup>41</sup>. Nevertheless, this model does not consider the potential formation of specific interactions between a nascent chain and the ribosome surface, which have been shown to be a critical aspect to explain the productive folding of a deeply knotted protein<sup>7,8</sup>. Specifically, Dabrowski-Tumanski *et al.* using coarse-grained molecular dynamics proposed that threading of a nascent chain could occur co-translationally through a twisted loop stabilized by specific interactions near of the ribosomal exit tunnel<sup>27</sup>. Since weak interactions between the ribosome and nascent chains has been described experimentally<sup>44</sup>, it is plausible that intermolecular interactions –rather than non-native interactions– could contribute to form a knot for proteins with intricate and complex topologies like MJ0366.

## Methods

**MJ0366 constructs design.** Three different pulling geometries were used in this study, one to conserve the knot upon mechanical denaturation and two to untie it. To maintain the knot in the unfolded state, a construct was design by placing cysteines at the N-tail (F6C) and the C-tail (G89C). However, as MJ0366 has a cysteine at position 81, a C81A mutation was added as well. The other two pulling geometries were designed to untie the knot upon pulling based on the approach developed by Sulkowska *et al.*<sup>45</sup>. To untie the knot by the C-tail, one cysteine was placed at the N-tail (F6C) and the other one at the core of the knot (K73C). Here, to theoretically assure the untying of the protein the condition  $p_1 < 2p_2 - N$  must be met, where  $p_1$  is 6,  $p_2$  is 73 and  $N$  is the total number of residues. For the construct F6C/K73C, the condition to untie the knot upon stretching is accomplished ( $6 < 49$ ). To untie the knot by pulling from the core of the knot now  $p_1$  is 19 and  $p_2$  is 73. If there are no interactions that interrupt the unknitting of the polypeptide chain, the velocity of untying is the same in both ends. Therefore, each pulling site will slide 27 residues upon stretching. Since  $p_{1-27} = -8$  and  $p_2 + 27 = 100$ , theoretically the construct K19C/K73C will untie upon mechanical denaturation.

**Gene cloning, protein expression and purification.** The MJ0366 coding sequence (gene symbol MJ\_RS01930) was cloned into the pET21d (+) (Novagen) vector with a 6xHis tag at the C-terminal of the proteins. Proteins were overexpressed and purified as reported in Ramirez *et al.*<sup>46</sup>. To eliminate the affinity tag, the TEV protease cleavage site ENLYFQG was added between the C-terminal and the His tag. The dimeric state of the mutants was confirmed by size exclusion chromatography using a Superdex S75 column (GE Healthcare).

**Determination of MJ0366 stability and unfolding kinetics by chemical denaturation using circular dichroism spectroscopy.** The equilibrium denaturation and renaturation curves were determined in the presence of GdnHCl between 0 and 6 M. Protein samples were incubated at several concentrations of GdnHCl for 1 hr at 23 °C in 30 mM Tris-HCl pH 8, 6 mM Na<sub>2</sub>PO<sub>4</sub> pH 8, 500 mM NaCl, 1.5% glycerol and 1 mM DTT. Under these conditions the stability curves obtained with circular dichroism (CD) were superimpose. Far UV CD spectra were acquired in a Jasco-1500 dichrograph, using a protein concentration of 10 μM in a 1 mm cell. The folding stability curve was analyzed as indicated previously<sup>35</sup>, considering the unfolding of the monomer of MJ0366 through a two-state model of unfolding. For the unfolding kinetics experiments, the native protein at 100 μM was manually diluted to 10 μM in 30 mM Tris-HCl pH 8, 6 mM Na<sub>2</sub>PO<sub>4</sub> pH 8, 500 mM NaCl, 1.5% glycerol and 1 mM DTT 10 μM at several concentrations of GdnHCl. The change in secondary structure was measured at 222 nm. The observed kinetic traces were fitted to a single exponential decay using SigmaPlot version 10.0 (Systat) and, the dependence of the observed kinetic constants with GdnHCl was fitted to a linear regression to determine  $m$  values and the kinetic constant at 0 M of chaotropic.

**DNA handles attachment to proteins.** Proteins were concentrated to 3–5 mg/ml, incubated overnight with 50 mM DTT at 4 °C and purified by gel filtration in 50 mM Na<sub>2</sub>PO<sub>4</sub> pH 8 and 500 mM NaCl. The eluted protein was incubated with a 70-fold excess of 2,2'-Dithiodipyridine (DTDP) during 15 to 60 min. The excess of DTDP was removed by consecutively using two Micro Bio-Spin (Bio-Rad) previously equilibrated in 50 mM Na<sub>2</sub>PO<sub>4</sub> pH 8 and 500 mM NaCl. The DNA handle attachment was performed as described previously by Hao *et al.*<sup>47</sup>. Briefly, the activated proteins were incubated with 5'-Tiolmodified dsOligos of ~33 bp. The formation of the double labeled protein-ds-Oligos complex was evaluated by electrophoresis in a 15% Tris-Glycine polyacrylamide gel (Supplementary Fig. S7A) and purified by electroelution. The electric field for the electroelution was 3 V/cm for 1 hr, followed by one minute of 5 V/cm with the poles inverted. The purified protein was dialyzed against 50 mM Na<sub>2</sub>PO<sub>4</sub> pH 8 and 500 mM NaCl for 2 hrs. Next, the protein-dsOligos complex was ligated to dsDNA handles of 333 and 356 bp chemically modified with biotin or digoxigenin, respectively. The ligation was monitored by electrophoresis in a 6% TBE polyacrylamide gel (Supplementary Fig. S7B).

**Optical tweezers experiments.** Experiments were performed using a MiniTweezers device<sup>48</sup>. The protein samples were manipulated in presence of 50 mM Na<sub>2</sub>PO<sub>4</sub> pH 8 and 500 mM NaCl. The protein-DNA complexes were incubated with 3.1 μm anti-dig coated beads for 15 min at 20 °C and injected to the microfluidic chamber. As shown in Fig. 1C, a 3.1 μm bead was trapped by the laser beam and a second 2.1 μm bead coated with streptavidin was held by suction in a micropipette tip. To catch a molecule between the two beads, the 3.1 μm bead was moved



towards the 2.1  $\mu\text{m}$  bead, by moving the optical trap closer to the pipette tip. The presence of a single DNA-protein complex was evaluated by the determination of the overstretching distance of the dsDNA handles at  $\sim 65$  pN<sup>49,50</sup>, which corresponded to  $\sim 95\%$  of the theoretical total length expected for a DNA handle of 716 bp.

**Determination of kinetic and thermodynamic folding parameters.** The force dependence of unfolding and refolding rate constants were extracted from the force distribution using the model developed by Dudko *et al.*<sup>36,37</sup>. As the dependence of force and the natural logarithm of rate constants was linear, we fitted this data using the model described by Bell<sup>38</sup>. As indicated in the Supplementary Fig. S4, the experimental distributions of unfolding and refolding forces were well predicted by the kinetics parameter calculated at zero force (Table 1). We determined the error of the unfolding rate constants, as the counting error of each bin in the force distributions (See Supplementary Fig. S4). This error was calculated as the inverse of the root square of the total events (frequency) contained in each bin ( $\text{Counting error} = 1/\sqrt{\text{frequency per bin}}$ ).

In the case of F6C/K73C and K19C/K73C the refolding transitions were not apparent from the force-extension trajectories. Therefore, the values of  $k_0$  and  $\Delta x^\ddagger$  were evaluated indirectly through the refolding probability ( $P_R$ ) calculated by the chance to observe an unfolding rip during the stretching cycle as indicated in the main text<sup>39</sup>. Data using three different waiting times (1, 5 and 10 s) was only collected with the mutant F6C/K73C, which allowed us to perform a global fit to the integrated form of Bell equation<sup>39</sup>. In the case of the mutant K19C/K73C data was obtained at only 10 s of waiting time.

The kinetic parameters of the F6C/G89C mutant, were also determined from experiments at constant force under equilibrium conditions. In this case, the temporal variation of the molecular fluctuation, were performed by using the electronic feedback system to keep the force constant during the change of the trap position<sup>51</sup>. The response time of the electronic feedback is 1 ms, which is faster than the molecular interconversion dwell times observed in the interval of forces explored which are in the order of seconds. The hopping traces were obtained for 11 molecules of F6C/G89C between 12.6 to 14.6 pN every 0.2 pN during of 2–5 min at each force (Fig. 3B). F6C/G89C hops only between two states; a compact (native) and extended (unfolded) state. The dwell times in each state were determined using the pCLAMP software version 10.6 (Molecular Devices)<sup>52</sup> and the unfolding and the refolding rate constants at different forces were determined from the exponential distributions of dwell time of the native and unfolded states respectively (Supplementary Fig. S5). The values of  $\Delta x^\ddagger$  and  $k_0$  were determined from the semi-logarithmic form of the Bell equation<sup>38</sup>.

The thermodynamic free energy of stability between the native and unfolded states were determined from the quotient between the refolding and unfolding rate constants ( $\Delta G = -RT \ln \left( \frac{k_0^U}{k_0^R} \right)$ ) or directly by the application of the Crooks fluctuation theorem<sup>53,54</sup> for the mutant F6C/G89C. The calculation of reversible work from constant velocity pulling experiments was determined as described previously by Bustamante *et al.*<sup>34</sup>.

**Calculating contour length.** The experimental values of contour length ( $L_c$ ) were determined from the distributions of ratios between the experimental extensions changes ( $\Delta x_{\text{exp}}$ ) with respect to the theoretical extension ( $\Delta x_{\text{theoretical}}$ ) expected to occur at identical forces as described previously<sup>34</sup>. The values of  $\Delta x_{\text{theoretical}}$  at several forces were determined by numerical integration of the WLC model<sup>55,56</sup>, which includes the distance between the cysteines in the native state (Fd). For the mutants F6C/G89C, F6C/K73C and K19C/K73C the value of  $L_{c,\text{theoretical}}$  are 30.3 nm; 24.4 nm and 19.8 nm respectively; and the value of Fd are 4.1 nm; 1.9 nm and 1.9 nm, respectively. The values of  $L_{c,\text{theoretical}}$  were calculated considering a persistence length of 0.65 nm and a  $L_c$  of 0.365 nm per residues<sup>57</sup>. The  $L_c$  at each particular rupture force was determined as  $L_{c,\text{theoretical}} \bullet ((\Delta x_{\text{exp}} + \text{Fd}) / \Delta x_{\text{theoretical}})$ . The average  $L_c$  was obtained by performing a Gaussian fit to the  $L_c$  distribution histogram for the mutants here reported (Fig. S2).

Received: 7 August 2019; Accepted: 12 May 2020;

Published online: 12 June 2020

## References

- Jamroz, M. *et al.* KnotProt: a database of proteins with knots and slipknots. *Nucleic Acids Res.* **43**, D306–D314 (2015).
- King, N. P., Yeates, E. O. & Yeates, T. O. Identification of Rare Slipknots in Proteins and Their Implications for Stability and Folding. *J. Mol. Biol.* **373**, 153–166 (2007).
- Dabrowski-Tumanski, P. & Sulkowska, J. I. Topological knots and links in proteins. *Proc. Natl. Acad. Sci. USA* **114**, 3415–3420 (2017).
- Dabrowski-Tumanski, P. & Sulkowska, J. I. To Tie or Not to Tie? That Is the Question. *Polymers (Basel)*. **9**, (2017).
- Jackson, S. E., Suma, A. & Micheletti, C. How to fold intricately: using theory and experiments to unravel the properties of knotted proteins. *Curr. Opin. Struct. Biol.* **42**, 6–14 (2017).
- Faisca, P. F. N. Knotted proteins: A tangled tale of Structural Biology. *Comput. Struct. Biotechnol. J.* **13**, 459–468 (2015).
- Wallin, S., Zeldovich, K. B. & Shakhnovich, E. I. The Folding Mechanics of a Knotted Protein. *J. Mol. Biol.* **368**, 884–893 (2007).
- Skrbić, T., Micheletti, C. & Faccioli, P. The role of non-native interactions in the folding of knotted proteins. *PLoS Comput. Biol.* **8**, e1002504 (2012).
- Beccara, A. *et al.* Folding Pathways of a Knotted Protein with a Realistic Atomistic Force Field. *PLoS Comput. Biol.* **9**, e1003002 (2013).
- Chwastyk, M. & Cieplak, M. Cotranslational folding of deeply knotted proteins. *J. Phys. Condens. Matter* **27**, 1–15 (2015).
- Dabrowski-Tumanski, P., Jarmolinska, A. I. & Sulkowska, J. I. Prediction of the optimal set of contacts to fold the smallest knotted protein. *J. Phys. Condens. Matter* **27**, 354109 (2015).
- Noel, J. K., Sulkowska, J. I. & Onuchic, J. N. Slipknotting upon native-like loop formation in a trefoil knot protein. *Proc. Natl. Acad. Sci. USA* **107**, 15403–15408 (2010).
- Li, W., Terakawa, T., Wang, W. & Takada, S. Energy landscape and multiroute folding of topologically complex proteins adenylate kinase and zouf-knot. *Proc. Natl. Acad. Sci. USA* **109**, 17789–17794 (2012).

14. Mallam, A. L. & Jackson, S. E. A Comparison of the Folding of Two Knotted Proteins: YbeA and YibK. *J. Mol. Biol.* **366**, 650–665 (2007).
15. Mallam, A. L. & Jackson, S. E. Probing Nature's Knots: The Folding Pathway of a Knotted Homodimeric Protein. *J. Mol. Biol.* **359**, 1420–1436 (2006).
16. King, N. P., Jacobitz, A. W., Sawaya, M. R., Goldschmidt, L. & Yeates, T. O. Structure and folding of a designed knotted protein. *Proc. Natl. Acad. Sci. USA* **107**, 20732–20737 (2010).
17. Wang, L., Chen, S.-Y. & Hsu, S.-T. D. Folding analysis of the most complex Stevedore's protein knot. *Sci. Rep.* **6**, 31514 (2016).
18. Zhang, H. & Jackson, S. E. Characterization of the Folding of a 52-Knotted Protein Using Engineered Single-Tryptophan Variants. *Biophys. J.* **111**, 2587–2599 (2016).
19. Lou, S.-C. *et al.* The Knotted Protein UCH-L1 Exhibits Partially Unfolded Forms under Native Conditions that Share Common Structural Features with Its Kinetic Folding Intermediates. *J. Mol. Biol.* **428**, 2507–2520 (2016).
20. Andersson, F. I., Pina, D. G., Mallam, A. L., Blaser, G. & Jackson, S. E. Untangling the folding mechanism of the 52-knotted protein UCH-L3. *FEBS J.* **276**, 2625–2635 (2009).
21. Mallam, A. L., Rogers, J. M. & Jackson, S. E. Experimental detection of knotted conformations in denatured proteins. *Proc. Natl. Acad. Sci. USA* **107**, 8189–8194 (2010).
22. Wang, P., Yang, L., Liu, P., Gao, Y. Q. & Zhao, X. S. Single-molecule detection reveals knot sliding in TrmD denaturation. *Chem. - A Eur. J.* **19**, 5909–5916 (2013).
23. Mallam, A. L. & Jackson, S. E. Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. *Nat. Chem. Biol.* **8**, 147–153 (2011).
24. Ziegler, F. *et al.* Knotting and unknotting of a protein in single molecule experiments. *Proc. Natl. Acad. Sci. USA* **113**, 7533–7538 (2016).
25. Niewiczzerzal, S. & Sulkowska, J. I. Knotting and unknotting proteins in the chaperonin cage: Effects of the excluded volume. *PLoS One* **12**, 1–23 (2017).
26. Zhao, Y., Dabrowski-Tumanski, P., Niewiczzerzal, S. & Sulkowska, J. I. The exclusive effects of chaperonin on the behavior of proteins with 52 knot. *PLoS Comput. Biol.* **14**, e1005970 (2018).
27. Dabrowski-Tumanski, P., Piejko, M., Niewiczzerzal, S., Stasiak, A. & Sulkowska, J. I. Protein Knotting by Active Threading of Nascent Polypeptide Chain Exiting from the Ribosome Exit Channel. *J. Phys. Chem. B* **122**, 11616–11625 (2018).
28. Lim, N. C. H. & Jackson, S. E. Mechanistic Insights into the Folding of Knotted Proteins *In Vitro* and *In Vivo*. *J. Mol. Biol.* **427**, 248–258 (2015).
29. Especial, J., Nunes, A., Rey, A. & Faisca, P. F. N. Hydrophobic confinement modulates thermal stability and assists knotting in the folding of tangled proteins. *Phys. Chem. Chem. Phys.* **21**, 11764–11775 (2019).
30. Thiruselvam, V. *et al.* 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**, 264–269 (2017).
31. Bustamante, C., Marko, J., Siggia, E. & Smith, S. Entropic elasticity of lambda-phage DNA. *Science*. **265**, 1599–1600 (1994).
32. Dzubiella, J. Sequence-Specific Size, Structure, and Stability of Tight Protein Knots. *Biophys. J.* **96**, 831–839 (2009).
33. He, C., Lamour, G., Xiao, A., Gsponer, J. & Li, H. Mechanically Tightening a Protein Slipknot into a Trefoil Knot. *J. Am. Chem. Soc.* **136**, 11946–11955 (2014).
34. Bustamante, A. *et al.* The energy cost of polypeptide knot formation and its folding consequences. *Nat. Commun.* **8**, 1581 (2017).
35. Wang, L., Chen, S.-Y. & Hsu, S.-T. D. Unraveling the Folding Mechanism of the Smallest Knotted Protein, MJ0366. *J. Phys. Chem. B* **119**, 4359–4370 (2015).
36. Dudko, O. K., Hummer, G. & Szabo, A. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci. USA* **105**, 15755–15760 (2008).
37. Dudko, O. K., Hummer, G. & Szabo, A. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.* **96**, 108101 (2006).
38. Bell, G. I., Dembo, M. & Bongrand, P. Cell adhesion. Competition between nonspecific repulsion and specific bonding. *Biophys. J.* **45**, 1051–64 (1984).
39. Min, D., Jefferson, R. E., Bowie, J. U. & Yoon, T.-Y. Mapping the energy landscape for second-stage folding of a single membrane protein. *Nat. Chem. Biol.* **11**, 981–7 (2015).
40. Robinson, C. R. & Sauer, R. T. Equilibrium Stability and Sub-Millisecond Refolding of a Designed Single-Chain Arc Repressor. *Biochemistry* **35**, 13878–13884 (1996).
41. Chwastyk, M. & Cieplak, M. Multiple folding pathways of proteins with shallow knots and co-translational folding. *J. Chem. Phys.* **143**, 045101 (2015).
42. Noel, J. K., Onuchic, J. N. & Sulkowska, J. I. Knotting a Protein in Explicit Solvent. *J. Phys. Chem. Lett.* **4**, 3570–3573 (2013).
43. Soler, M. A., Rey, A. & Faisca, P. F. N. Steric confinement and enhanced local flexibility assist knotting in simple models of protein folding. *Phys. Chem. Chem. Phys.* **18**, 26391–26403 (2016).
44. Goldman, D. H. *et al.* Mechanical force releases nascent chain-mediated ribosome arrest *in vitro* and *in vivo*. *Science*. **348**, 457–460 (2015).
45. Sulkowska, J. I., Sulkowski, P., Szymczak, P. & Cieplak, M. Untying Knots in Proteins. *J. Am. Chem. Soc.* **132**, 13954–13956 (2010).
46. Ramirez, M. P. *et al.* Single molecule force spectroscopy reveals the effect of BiP chaperone on protein folding. *Protein Sci.* **26**, (2017).
47. Hao, Y., Canavan, C., Taylor, S. S. & Maillard, R. A. Integrated Method to Attach DNA Handles and Functionally Select Proteins to Study Folding and Protein-Ligand Interactions with Optical Tweezers. *Sci. Rep.* **7**, 10843 (2017).
48. Smith, S. B., Cui, Y. & Bustamante, C. Optical-trap force transducer that operates by direct measurement of light momentum. *Methods Enzymol.* **361**, 134–162 (2003).
49. Smith, S. B., Cui, Y. & Bustamante, C. Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules. *Science*. **271**, 795–799 (1996).
50. Cluzel, P. *et al.* DNA: An Extensible Molecule. *Science*. **271**, 792–794 (1996).
51. Li, P. T. X., Collin, D., Smith, S. B., Bustamante, C. & Tinoco, I. Probing the Mechanical Folding Kinetics of TAR RNA by Hopping, Force-Jump, and Force-Ramp Methods. *Biophys. J.* **90**, 250–260 (2006).
52. Sherman-Gold, R. The axon guide for electrophysiology & biophysics laboratory techniques. *Axon Instruments* <https://doi.org/10.1136/bjo.2009.169052> (1993).
53. Crooks, G. E. Path-ensemble averages in systems driven far from equilibrium. *Phys. Rev. E* **61**, 2361–2366 (2000).
54. Crooks, G. E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* **60**, 2721–2726 (1999).
55. Keller Mayer, M. S. Folding-Unfolding Transitions in Single Titin Molecules Characterized with Laser Tweezers. *Science*. **276**, 1112–1116 (1997).
56. Rief, M. Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM. *Science*. **276**, 1109–1112 (1997).
57. Dietz, H. & Rief, M. Protein structure by mechanical triangulation. *Proc. Natl. Acad. Sci. USA* **103**, 1244–1247 (2006).
58. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

## Acknowledgements

We thank to Dr. Timothy H. Warren, Dr. María Fe Lanfranco and Dr. Diego Ferreiro for helpful discussions. This work was supported by Fondecyt 1151274 and 1191153 (MB), Fondecuip EQM140151. Maira Rivera was supported by the Conicyt Doctoral fellowship 21130254. Rodrigo Maillard was supported by the NSF (MCB 1715572) and NIH (1R15GM135866).

## Author contributions

M.B., R.M. and M.R. designed research and analyzed data; M.R. and Y.H. performed research; M.B., R.M. and M.R. wrote the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-66258-5>.

**Correspondence** and requests for materials should be addressed to R.A.M. or M.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020