Received: 11 January 2024 | Revised: 6 May 2024 | Accepted: 9 May 2024

DOI: 10.1002/pro.5031

RESEARCH ARTICLE

THE PROTEIN SOCIETY WILEY

Targeting the protein folding transition state by mutation: Large scale (un)folding rate accelerations without altering native stability

Luis A. Campos[1,2] | **Victor Muñoz**[3,4]

[1]Instituto Madrileño de Estudios Avanzados en Nanociencia (IMDEA Nanociencia), Madrid, Spain

[2]Unidad de Nanobiotecnología Asociada al Centro Nacional de Biotecnología (CSIC), Madrid, Spain

[3]Department of Bioengineering, University of California, Merced, California, USA

[4]Center for Cellular and Biomolecular Machines, University of California, Merced, California, USA

Correspondence
Luis A. Campos, IMDEA Nanociencia, Faraday 9, 28049 Madrid, Spain.
Email: luis.campos@imdea.org

Victor Muñoz, Department of Bioengineering, University of California, 95343, Merced, CA, USA.
Email: vmunoz3@ucmerced.edu

Funding information
European Research Council, Grant/Award Number: ERC-2012-ADG-323059; National Science Foundation, Grant/Award Numbers: MCB-2112710, HRD-2112675

Review Editor: Aitziber L. Cortajarena

Abstract

Proteins are constantly undergoing folding and unfolding transitions, with rates that determine their homeostasis in vivo and modulate their biological function. The ability to optimize these rates without affecting overall native stability is hence highly desirable for protein engineering and design. The great challenge is, however, that mutations generally affect folding and unfolding rates with inversely complementary fractions of the net free energy change they inflict on the native state. Here we address this challenge by targeting the folding transition state (FTS) of chymotrypsin inhibitor 2 (CI2), a very slow and stable two-state folding protein with an FTS known to be refractory to change by mutation. We first discovered that the CI2's FTS is energetically taxed by the desolvation of several, highly conserved, charges that form a buried salt bridge network in the native structure. Based on these findings, we designed a CI2 variant that bears just four mutations and aims to selectively stabilize the FTS. This variant has >250-fold faster rates in both directions and hence identical native stability, demonstrating the success of our FTS-centric design strategy. With an optimized FTS, CI2 also becomes 250-fold more sensitive to proteolytic degradation by its natural substrate chymotrypsin, and completely loses its activity as inhibitor. These results indicate that CI2 has been selected through evolution to have a very unstable FTS in order to attain the kinetic stability needed to effectively function as protease inhibitor. Moreover, the CI2 case showcases that protein (un)folding rates can critically pivot around a few key residues-interactions, which can strongly modify the general effects of known structural factors such as domain size and fold topology. From a practical standpoint, our results suggest that future efforts should perhaps focus on identifying such critical residues-interactions in proteins as best strategy to significantly improve our ability to predict and engineer protein (un)folding rates.

KEYWORDS

folding rates, folding transition state, mutational analysis, natural selection, protein engineering, protein stability, unfolding rates

1 | INTRODUCTION

Proteins fold into defined three-dimensional structures, whether on their own (Dill & MacCallum, 2012) or helped by interactions with partners (Wright & Dyson, 2009). These native structures are usually minimally stable (Jaenicke, 1999), a trait that appears to have been selected through evolution according to recent ancestral protein reconstruction studies (Modi et al., 2021). As a result, in physiological conditions, proteins are in constant exchange between their native and unfolded states (Campos et al., 2020). Such conformational dynamics endows proteins with the ability to reshape on cue that is at the core of their biological functionality. The rates of folding and unfolding are also key elements of protein homeostasis by determining the time to activation of any newly synthesized protein, as well as its susceptibility to targeted degradation (Toyama & Hetzer, 2013). Developing an in-depth understanding of the rules that define protein (un)folding rates is thus of fundamental significance. In the era of deep-learning protein structure prediction (Jumper et al., 2021), a better understanding of the factors that determine these rates would provide a uniquely complementary tool for engineering optimized proteins towards biotechnological applications.

The kinetics of protein folding have been studied for many decades, both experimentally (Fersht, 1999) and theoretically (Onuchic et al., 1997). According to those studies, the folding and unfolding rates of two-state single-domain proteins are defined by the crossing of the folding transition state (FTS) in each direction. The FTS corresponds to the conformational sub-ensemble along the (un)folding reaction path with the highest free energy (Fersht, 1995), and hence to the top of the folding free energy barrier (Akmal & Muñoz, 2004). Kinetically, the FTS is manifested in a V-shaped dependence of the logarithm of the (un)folding relaxation rate as a function of denaturant concentration, also known as chevron plot (Figure 1a). The relative slopes for both limbs of the chevron indicate the FTS' position along the reaction (Zarrine-Afsar & Davidson, 2004), whereas their steepness is related to the overall height of the barrier (Naganathan et al., 2007). Experimental studies have shown that the folding and unfolding rates of single-domain proteins vary vastly, changing by at least six or eight orders of magnitude respectively (De Sancho & Muñoz, 2011). The main factor producing such wide ranges is thought to be the differences in native structure, which translate into FTS that are specific for each protein fold. In this context, structurally conservative changes in the aminoacid sequence affect the rates only through their effect on the stability of the native structure (Fersht et al., 1992). The FTS has intermediate properties relative to the native and unfolded states, both in terms of structure and sensitivity to denaturants. Hence the free energy

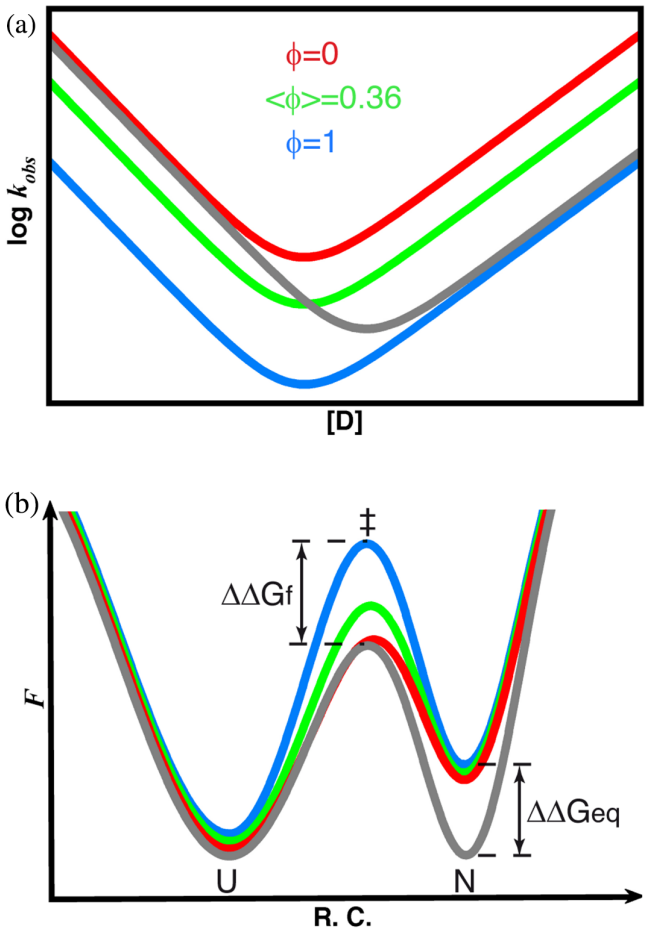


FIGURE 1 Schematic representation of the different scenarios for the (un)folding kinetic effects of mutations on proteins. (a) Generic chevron plots representing the logarithm of the relaxation rate ($k_{obs} = k_f + k_u$ for a two-state folder) of any given protein as a function of the concentration of chemical denaturant. The mutation produces a thermodynamic perturbation that results in a shift of the denaturation midpoint; in this case the shift is to the left indicating the destabilization of the native state. The red curve corresponds to mutations for which the whole thermodynamic perturbation is recovered in the unfolding rate, leading to the $\phi = 0$ scenario. The blue curve corresponds to mutations for which the entire thermodynamic perturbation is recovered in the folding rate, leading to the $\phi = 1$ scenario. The green curve corresponds to the average kinetic effect observed on over 900 mutations made in 52 different two-state folding proteins, which corresponds to $a < \phi > \sim 0.36$. About 95% of the mutations in that dataset are consistent with this scenario within the limits of the experimental uncertainty (Naganathan & Muñoz, 2010). (b) Free energy profiles for the three scenarios of A, showing the changes in stability induced by the mutation on the native state (N) and on the FTS (‡) relative to the unfolded state (U): $\Delta\Delta G_{eq}$ and $\Delta\Delta G_f$, respectively.

perturbations caused by mutations should partition between the folding and unfolding rates according to the degree of native structure that is already formed around the mutated area in the FTS (Fersht et al., 1992). Mutational effects can thus range between all of the free energy change being recapitulated in the folding rate, and all of it occurring on unfolding, shown in blue and red in Figure 1, respectively. This general scenario conforms with the vast majority of the existing mutational protein folding data (Fersht, 1999; Naganathan & Muñoz, 2010). In practice, mutational rate effects appear to be even further constrained since the vast majority of mutations used so far to analyze folding kinetics are consistent with ~2/3 of the free energy change occurring in unfolding and the remaining ~1/3 in folding (Naganathan & Muñoz, 2010) (green in Figure 1). The implication of this structure-centric scenario is that conservative mutations, whether arising from genetic drift or from protein engineering efforts, cannot target the FTS alone. In other words, conservative mutations shall affect both rates in opposite directions. This hypothesis has profound practical consequences because it sets fundamental limits on what is achievable in terms of engineering protein (un)folding rates.

However, how much the aminoacid sequence can alter the rates of protein (un)folding has not been systematically explored by experiment yet. Particularly, it remains unclear whether structurally conservative single-point mutations could be found that simultaneously accelerate, or decelerate, the folding and unfolding rates of a given protein. Here we address this issue by aiming to selectively stabilize the FTS of a slow two-state folding protein via mutation without affecting the overall stability of its native state. We use chymotrypsin inhibitor 2 (CI2) as model system for the study, a single-domain alpha/beta protein from barley seeds that is a paradigm for highly cooperative two-state folding behavior (Jackson & Fersht, 1991). CI2 has been thoroughly subjected to mutational analysis of protein folding (Itzhaki et al., 1995; Jackson et al., 1993; Jackson & Fersht, 1994; Lawrence et al., 2010). Hence, there are mutations available at nearly every possible position of the CI2 scaffold. The vast majority of these mutations are destabilizing and affect primarily the unfolding rate (Naganathan & Muñoz, 2010), pointing to a highly unstructured FTS that lies between the green and red curves of Figure 1. This makes CI2 a particularly challenging case for accelerating the folding rate. In addition, CI2 provides an opportunity to explore the role of functional constraints in defining the FTS of a natural protein fold. As a protease inhibitor, CI2 must be resilient to proteolytic degradation in order to effectively perform its inhibitory role. Achieving protease resilience may involve minimizing the half-life of the degradation-prone unfolded state. This can be achieved with a highly stable native structure and, more significantly, with a very slow unfolding rate. It has been in fact noted that the stability and kinetic barrier of CI2 are uncharacteristically high relative to other single-domain proteins (De Sancho et al., 2009). It is therefore quite possible that the properties of CI2's FTS, most particularly its insensitivity to mutations, have arisen from strong evolutionary pressure for functional reasons.

2 | RESULTS AND DISCUSSION

Our strategy for targeting the FTS started by identifying existing mutations on CI2 that seem to fall outside the range defined by the blue and red scenarios of Figure 1, even if minimally. Non-canonical effects on protein (un) folding rates have been studied using computer simulations (Ozkan et al., 2001). In experiments, confidently classifying mutational effects as non-canonical requires that the observed changes in rates are significantly larger than the experimental error in their determination, which is only true for the larger free energy perturbations (de los Rios et al., 2006; Sanchez & Kiefhaber, 2003). One issue is that the mutations used in folding studies are conservative by design, aiming at shaving a few localized native interactions through the truncation of aliphatic sidechains and/or replacements to alanine (Campos, 2022; Fersht et al., 1992). A close inspection of the existing data for CI2 confirmed that most reported mutations produce canonical rate effects close to scenario 1, that is without affecting the folding rate. The one exception came from studies that mutated the arginine at position 48 (R48) (Ladurner et al., 1998; Lawrence et al., 2010). One such study found that replacing R48 with other polar residues affected almost exclusively unfolding, as is usual for CI2; but mutating R48 to isoleucine or aromatic aminoacids anomalously produced large effects on the folding rate (Lawrence et al., 2010).

We thus focused on the CI2 region surrounding R48. This region contains not one, but two buried arginine residues (R46 and R48) that form a salt bridge network with the C-terminal carboxyl group, which is also buried and points inwards, and the partially solvent exposed E41 (Figure 2a). The salt bridge network has been deemed essential for locking the reactive loop into its binding-competent conformation (Jackson & Fersht, 1994). The two arginine residues that participate in the salt bridge network are indeed highly conserved in CI2 homologs, including all members of the potato inhibitor I family (Figure S1). A wider sequence alignment of 149 CI2 homologs showed that R48 is >70% conserved whereas R46 is found in every sequence, providing strong

evidence of the key functional role played by these residues. The C-terminal strand is also important to block CI2 from fold switching to a Rossman-like fold that is latently encoded in its sequence (Campos et al., 2019). We thus hypothesized that desolvation of the buried charges, which must occur before consolidating the salt bridge network detailed in Figure 2a, is a major contributor to the very high free energy of CI2's FTS. Such an effect could also help explain the insensitivity to mutations found for this protein. An earlier mutational study of the Arc repressor's folding coupled to dimerization provided some support for our hypothesis. That study showed that replacing two charged residues that form a salt-bridge buried in the dimer interface with sterically similar hydrophobic residues resulted on concerted rate accelerations in both the folding-dimerization and unfolding directions (Waldburger et al., 1996). We note, however, that the forward rate for the Arc repressor is determined by monomer–monomer collisions rather than by the intramolecular conformational dynamics that control folding of monomeric proteins such as CI2 (Muñoz, 2007).

We decided to explore whether charge burial does also control the FTS for CI2 by swapping the charge of either R48 or R46 with a hydrophobic moiety. Particularly, we mutated each arginine to isoleucine, which has a similar alkane chain that can be accommodated in either position without steric clashes (Figure 2b). The R48I mutation, which occurs in other CI2 homologs (Figure S1), recapitulated the result of Lawrence et al. (2010) Namely, R48I accelerated the folding rate by a factor of ~60 without much affecting unfolding (Figure 2c, Table 1). The net result for R48I is a ~10 kJ/mol stabilization of the native state, suggesting that the salt bridges between R48 and the C-terminal are not enough to compensate for the total charge desolvation penalty. This penalty includes the R48 terminal charge and possibly the C-terminal carboxylate, which may disengage from the salt bridge network to stay solvated once the R48 partner charge is missing. The fact that the kinetic effect is almost exclusive to the folding rate indicates that the R48 and C-terminal charges are already desolvated and interacting through the two salt bridges at the FTS. Hence, while statistically rare for CI2, R48I fits well to the canonical $\phi \approx 1$ scenario (blue in Figure 1).

The effect of R46I is quite different. This mutation accelerates both the folding and unfolding rates, by factors of ~6 and 45 respectively (Figure 2c, Table 1). The larger acceleration of unfolding lends a slight native destabilization of 5 kJ/mol. This result indicates that the energetic contribution from the salt bridges formed between R46 and the C-terminal carboxylate (Figure 2a) more than compensates the cost of desolvating the charges. On the other hand, the fact that R46I distinctly accelerates both rates constitutes an example of a distinctly non-canonical mutational effect in protein folding. In other words, R46I preferentially stabilizes the FTS relative to both the unfolded and native states. Structure–energy relationships shed some light onto the possible sources of this non-canonical behavior. For instance,

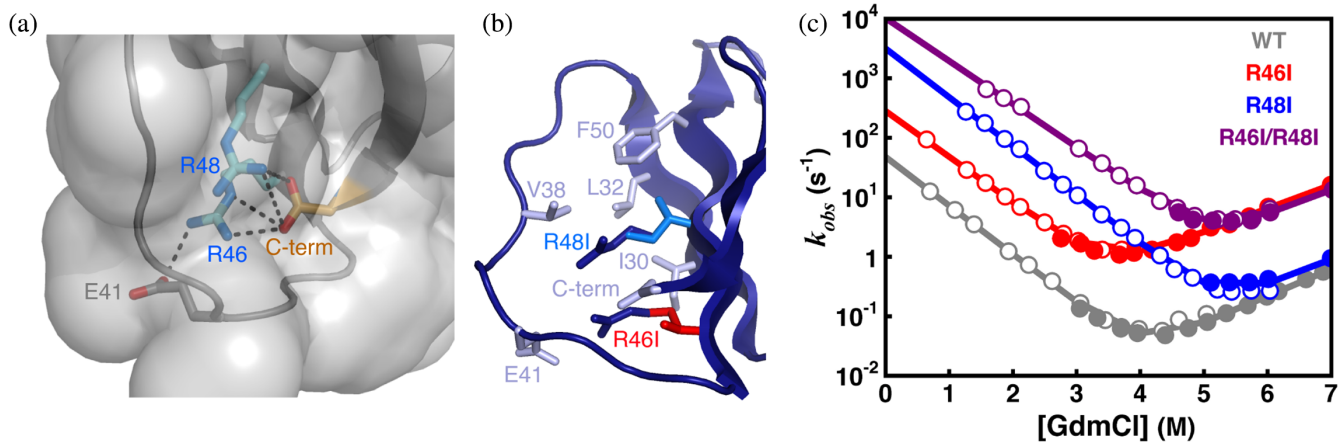


FIGURE 2 The role of the buried salt bridge network of CI2 on its (un)folding kinetics. (a) Detail of the electrostatic interactions between R46 and R48 with E41 and the C-terminal carboxylate shown in a surface representation of the CI2 native structure. The five depicted interactions constitute the buried salt bridge network that locks the functional loop into its active conformation. (b) Cartoon representation of the CI2 native structure showing the R46 and R48 side-chains together with their replacements to isoleucine. The arginine sidechains are shown in dark blue and the mutated isoleucine sidechains in red and bright blue, respectively. The side-chains of the residues that surround the two arginines in the CI2 native structure are shown in gray. (c) Experimental chevron plots for the R46I (red), R48I (blue), and R46I–R48I (purple) mutants of CI2 together with the wild-type (gray). Data points are shown as circles (filled for unfolding and open for refolding experiments) and best kinetic fits as continuous lines. The kinetic parameters from the fits are given in Table 1.

the ~4 kJ/mol stabilization of the FTS relative to the unfolded state likely reflects the penalty of desolvating the charge, as it occurs with R48I. In this regard, the smaller free energy gain for R46I (4 kJ/mol vs. 10 kJ/mol for R48I) is consistent with a lesser burial in the native structure (Figure 2a), which should make R46 potentially more accessible to water. The stabilization relative to the native state, on the other hand, suggests that the FTS allows for better packing of the hydrophobic residues around the mutated area than the native 3D structure does. This interpretation is consistent with previous reports on other proteins of mutations that enhanced hydrophobic packing and stabilized the FTS (Anil et al., 2005; Viguera et al., 2002; Zarrine-Afsar et al., 2008; Zhu et al., 2004).

To crosscheck this interpretation and determine whether the kinetic effects of both mutations could be effectively compounded, we performed the double mutation (R46I–R48I). We observed that the R46I–R48I mutant does indeed produce an additive effect in the kinetics (Figure 2c, Table 1). The double mutant maintains the ~20-fold faster unfolding rate of R46I together with a folding rate that roughly sums up the effects of both mutations. The folding rate extrapolated to water for R46I–R48I is ~200-fold faster. Hence, the results for the double-mutant further support a CI2 FTS that is energetically penalized by the desolvation of natively buried charges, including the terminal carboxylate. In contrast, the salt-bridge network seems to only consolidate after(before) crossing the FTS in the folding(unfolding) direction. The result of removing the buried salt bridge network is a highly stabilized FTS, as showcased by the acceleration of the rates in both directions.

However, R46I–R48I is significantly more stable than wild-type CI2, which makes it difficult to extricate FTS-specific effects from thermodynamic factors. Accordingly, we set on designing a CI2 variant that maximizes the FTS stability without altering unfolding thermodynamics so that we can exclusively target kinetic effects. For this design, we chose to eliminate the salt bridge network using the R46I mutation and swapping R48I for R48Y. The reasons for this swap were twofold: (1) The bulkier tyrosine, which is also found in other CI2 homologs (Figure S1), can further stabilize the FTS by increasing the local hydrophobic packing after the two charges are removed; (2) R48Y accelerates both folding and unfolding, producing a smaller overall change in native stability (Lawrence et al., 2010). As strategy for balancing the remaining thermodynamic bias, we looked for previously reported CI2 mutations that lend small native destabilizations by affecting only unfolding (red scenario of Figure 1). We chose V21A and K26A, which are in solvent exposed positions of the distant α-helix (Figure 3a) since they are known to produce minor destabilizations and accelerate unfolding (Itzhaki et al., 1995). Table S1 provides the rate changes expected for each of the four individual mutations.

The resultant engineered version of CI2 ($CI2_{FTS\text{-}opt}$) expressed at good levels, and could be purified for further characterization. Because $CI2_{FTS\text{-}opt}$ carries a total of four mutations, we first wanted to confirm its thermodynamic neutrality and explore any potential structural rearrangements that might be induced by mutation. Equilibrium chemical unfolding experiments showed that $CI2_{FTS\text{-}opt}$ has the same overall native stability than the wild-type, but in this case achieved by compensating a minor midpoint shift with a slightly lower sensitivity to chemical denaturant (Figure S2). The thermal stability of $CI2_{FTS\text{-}opt}$ is also consistent, as demonstrated in equilibrium thermal denaturation experiments (Figure S3). During the equilibrium characterization of $CI2_{FTS\text{-}opt}$ we noticed that the fluorescence emission of the single tryptophan, W5, in the native state differs from the wild-type's. In native wild-type CI2 the fluorescence of W5 is strongly quenched and blue-shifted relative to the unfolded state (dark blue vs. green in Figure S4). For $CI2_{FTS\text{-}opt}$ both the quenching and blue-shift are slightly reduced (orange in Figure S4). Further analysis showed that the changes in native fluorescence are incremental from single to double arginine substitutions and up to the four mutations of $CI2_{FTS\text{-}opt}$ (Figure S4). Overall, the changes in native fluorescence upon mutation are small relative to the effects caused by unfolding; but they indicate that the combined mutations, despite being thermodynamically neutral, induce some rearrangement around W5. To further assess this possibility, we used AlphaFold2 (Jumper

TABLE 1 (Un)folding kinetic parameters obtained from the chevron plots as a function of Guanidinium hydrochloride concentration of wild-type CI2 and all the mutations performed in this work.

CI2 variant	k_f (H_2O)	m_f	k_u (H_2O)	m_u
Wild-type	50 ± 5	−0.83 ± 0.02	0.0006 ± 0.00015	0.43 ± 0.02
R46I	290 ± 23	−0.77 ± 0.02	0.027 ± 0.004	0.40 ± 0.01
R48I	3160 ± 350	−0.82 ± 0.02	0.0006 ± 0.0005	0.46 ± 0.05
R46I–R48I	10,100 ± 880	−0.72 ± 0.01	0.012 ± 0.005	0.43 ± 0.03
$CI2_{FTS\text{-}opt}$	13,500 ± 920	−0.72 ± 0.01	0.166 ± 0.028	0.39 ± 0.01

Note: k_f and k_u are expressed in s^{-1} and m_f and m_u in kJ/(mol M).

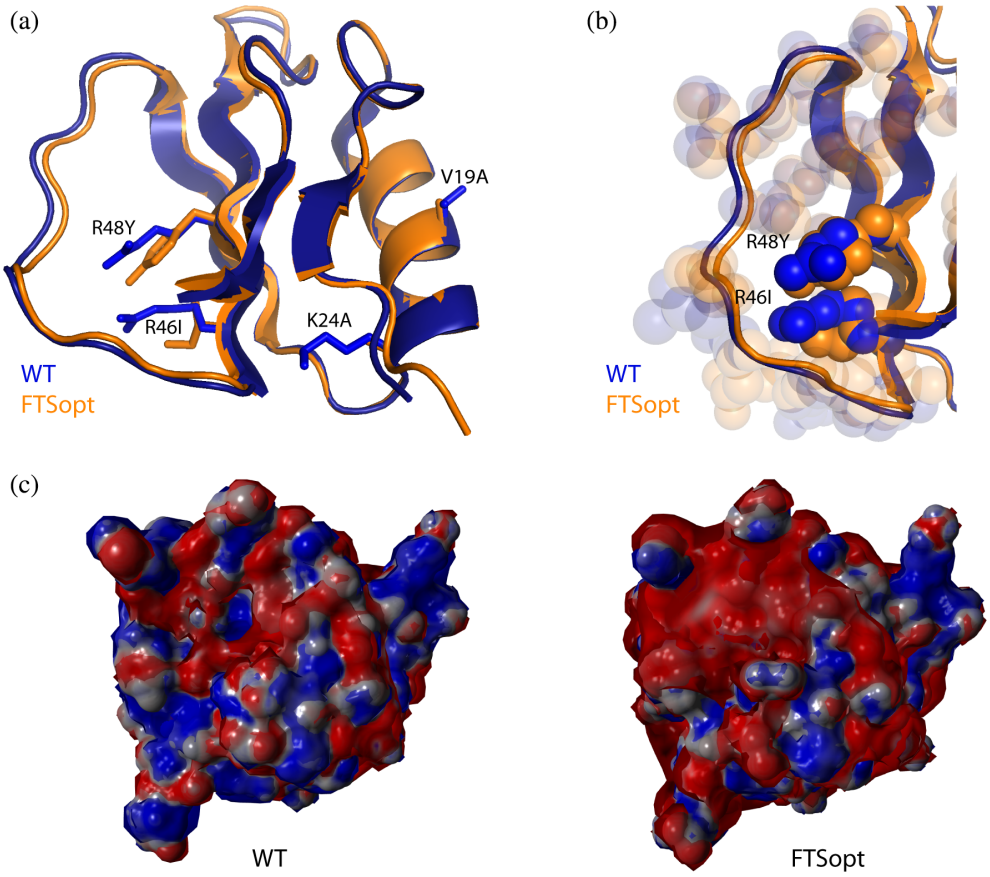
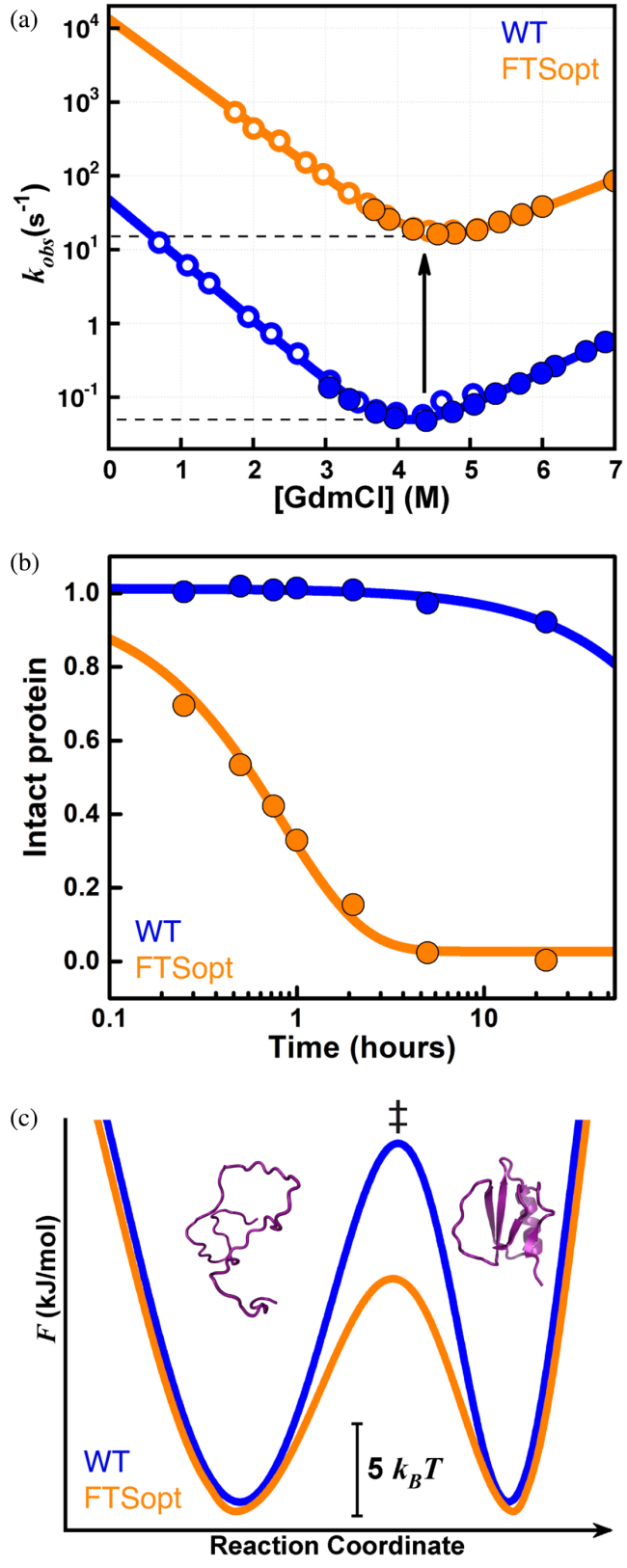


FIGURE 3 Structural characterization of $CI2_{TS\text{-}opt}$, the CI2 variant with mutations that selectively stabilize the FTS. (a) Cartoon representation of the $CI2_{TS\text{-}opt}$ native structure predicted by AlphaFold2 (orange) superimposed to the native structure of wild-type CI2 (dark blue). The side-chains of the four mutations made to produce the $CI2_{FTS\text{-}opt}$ variant: V19A, K21A, R46I, and R48Y, are also shown for both protein structures in stick representation following the same color scheme. (b) Detail of the packing of the functional loop region for both variants. The side-chains of all the relevant residues are shown in a space filling representation, with those at positions 46 and 48 in highlighted color for easy identification. The predicted structure for $CI2_{TS\text{-}opt}$ shows the functional loop slightly repacked around the more compact, hydrophobic pocket created by the I and Y at positions 46 and 48. (c) Representation of the electrostatic potential of wild-type CI2 (left) and $CI2_{TS\text{-}opt}$ (right) calculated applying the Poisson–Boltzmann treatment on the X-ray CI2 structure and the AlphaFold2 $CI2_{TS\text{-}opt}$ predicted structure, respectively.

et al., 2021) to predict the 3D structure of $CI2_{FTS\text{-}opt}$. AlphaFold2 achieved an excellent sequence coverage for $CI2_{FTS\text{-}opt}$, predicting its residues 4–62 with >80% confidence, including the functional loop (Figure S5). The predicted $CI2_{FTS\text{-}opt}$ structure is shown in orange in Figure 3a, superimposed to the crystallographic structure of wild-type CI2 in dark blue. This comparison shows that the structure predicted for $CI2_{FTS\text{-}opt}$ is nearly invariant for most of the protein. The only obvious changes are a minimal displacement of the N-terminal tail (Figure 3a), and a slight compaction of the functional loop due to side-chain repacking (Figure 3b). The N-terminal displacement moves the ε-NH_2 of K2 slightly away from the W5 indole. This contact is presumably the main quencher of W5 in native CI2 (Chen & Barkley, 1998), potentially explaining the observed change in native fluorescence upon mutation. The minor repacking of the functional loop is also consistent with our structural interpretation of the (un)folding kinetic effects of the R46I and R48I mutations (see above). $CI2_{FTS\text{-}opt}$ also carries an increase in negative charge that results in a drop to 4.45 from the pI = 5.02 of the wild-type. Poisson–Boltzmann electrostatic calculations based on the wild-type structure and the $CI2_{FTS\text{-}opt}$ predicted structure highlight that the change in net charge does result in a more negative electrostatic potential for $CI2_{FTS\text{-}opt}$. However, the change in electrostatic potential is small in magnitude and appears localized to that specific protein region (Figure 3c).

We then moved to investigate the (un)folding kinetic behavior of $CI2_{FTS\text{-}opt}$ by measuring its chevron plot (orange in Figure 4a). Fitting of the $CI2_{FTS\text{-}opt}$ chevron plot rendered a native stability of 28.1 kJ/mol. This 'kinetic' stability is also identical to the wild-type's (blue



in Figure 4a), providing kinetic confirmation that the design is thermodynamically neutral. There are, on the other hand, vast differences between $CI2_{FTS\text{-}opt}$ and the wild-type in terms of rates. The chevron plot is shifted up by more than two orders of magnitude, indicating that $CI2_{FTS\text{-}opt}$ has much faster rates in both directions. The fitted $CI2_{FTS\text{-}opt}$ folding rate in water is 13,500 s^{-1}, corresponding to a 270-fold acceleration. The fitted unfolding rate in water for $CI2_{FTS\text{-}opt}$ is 280-fold faster than for the wild-type. These rate accelerations are rather extreme, but they are not additive since the sum of effects for the single mutations would amount to factors of 520 and 2540 for folding and unfolding, respectively (Table S1). The non-additivity of the mutational effects, particularly in the unfolding direction, suggests that some compensating effects are at play that reduce the overall native destabilization. One such effect could be the repacking of the hydrophobic pocket around the new tyrosine and isoleucine at positions 48 and 46 in the native state that is observed in the predicted structure (Figure 3b). Another difference of the $CI2_{FTS\text{-}opt}$ chevron is in the folding and unfolding slopes, both of which are reduced by 13% (Table 1). A symmetrically flatter chevron plot reflects a reduced sensitivity to denaturant for both rates, which has been shown to indicate a lower (un)folding free energy barrier (Naganathan et al., 2007). The flatter chevron of $CI2_{FTS\text{-}opt}$ is hence likely a consequence of the so much faster (un)folding kinetics of this mutant, and provides an independent test of the correlation previously found between these two properties (Naganathan et al., 2007). We note that a symmetrically flatter chevron results on an unfolding rate in water that is comparatively faster than is slower the folding rate. This is so because the unfolding rate in water is obtained from a long extrapolation from high denaturant concentrations. On the other hand, the denaturation midpoint for $CI2_{FTS\text{-}opt}$ is slightly shifted to the right, which compensates the flatter chevron to render the same 'kinetic' native stability of the wild-type in water, consistently with the thermodynamic unfolding characterization (Figure S2).

FIGURE 4 (Un)folding kinetics and susceptibility to proteolysis of $CI2_{TS\text{-}opt}$ versus wild-type CI2. (a) Experimental chevron plot of $CI2_{FTS\text{-}opt}$ (orange) versus the chevron of the wild-type CI2 (blue). Data points are shown as circles (filled for unfolding and open for refolding experiments) and best kinetic fits as continuous lines. The kinetic parameters for the fits are given in Table 1. (b) Resistance to proteolytic degradation by chymotrypsin of CI2 wild-type (blue) and $CI2_{FTS\text{-}opt}$ (orange). (c) Free energy profiles at the denaturation midpoint ($\Delta G_{eq} = 0$) of the wild-type (blue) and $CI2_{FTS\text{-}opt}$ (orange) calculated assuming a pre-exponential term of 200,000 s^{-1}. The profile shows the unfolded state (left), the FTS (‡) and the native state (right). The bar provides a reference for $5k_BT$.

The fact that $CI2_{FTS\text{-}opt}$, containing four mutations that are thermodynamically neutral, exhibits much accelerated kinetics provides us with an opportunity to investigate the effects of the (un)folding kinetics on the biological activity and homeostasis of CI2. We determined the protease inhibitory activity of CI2 and $CI2_{FTS\text{-}opt}$ by performing a standard chymotrypsin proteolysis assay in the presence of different amounts of each CI2 variant. Wild-type CI2 showed strong inhibition with $K_i = 1.6 \times 10^{-8}$ M. In contrast, the proteolysis assay showed no signs of chymotrypsin inhibition at any tested concentration of $CI2_{FTS\text{-}opt}$ (Figure S6). The loss of inhibitory activity of $CI2_{FTS\text{-}opt}$ is likely a consequence of the increased flexibility of the functional loop once the salt bridge network is removed. This interpretation suggests that the very high energy of CI2's FTS is not an intrinsic property of the mechanism to form the native fold, but a byproduct of the need to retain the loop locked into an active conformation for its function as protease inhibitor. The perfect conservation of R46 and very high conservation of R48 we found in 149 CI2 homolog sequences further supports this hypothesis. In terms of homeostasis, the much faster kinetics of $CI2_{FTS\text{-}opt}$ does not interfere with its ability to fold in vivo since the protein expresses well and does not form inclusion bodies. On the other hand, we found that $CI2_{FTS\text{-}opt}$ becomes highly sensitive to protease degradation (Figure 4b). Particularly, the half-life of $CI2_{FTS\text{-}opt}$ in the presence of chymotrypsin is only ~1 h, whereas wild-type CI2 lasts for days. Notably, the 1-h half-life of $CI2_{FTS\text{-}opt}$ corresponds to a ~250-fold reduction that is essentially equivalent to the overall acceleration of its (un)folding kinetics. Such correspondence highlights a quantitative relationship between the susceptibility to hydrolysis and the faster rates, particularly since their thermodynamic stability is identical (Figure S2). Our results on $CI2_{FTS\text{-}opt}$ are consistent with a previous CI2 study which showed progressively higher proteolytic susceptibility for mutations R48A, R46A, and the truncation of the C-terminal glycine, in that order (Radisky et al., 2005). Although this previous mutational study did not investigate the effects on folding kinetics or stability, our data for R46I, R48I, and R46I–R48I (Figure 2c) strongly suggest that proportional rate accelerations should also be expected when the buried charges are removed by other means. The combined data point to the stability of the FTS as the main factor determining the susceptibility to proteolytic degradation by CI2 variants, providing strong evidence that CI2 evolved to have a very high free energy FTS as a way to meet functional requirements.

The rate acceleration of $CI2_{FTS\text{-}opt}$ corresponds to a targeted stabilization of the FTS of 14.5 kJ/mol, or $\sim 5.8 k_B T$. This change in free energy is large, amounting to half of the total native stability of CI2. The question is how this stabilization compares with the height of CI2's natural (un)folding barrier. We can indirectly address this issue using an empirical estimate for the (un)folding rate pre-exponential (Campos et al., 2020). Placing the pre-exponential for CI2 conservatively at ~200,000 s^{-1} or 1/(5 μs) (Campos et al., 2020), we obtain an (un)folding barrier at the denaturation midpoint for the wild-type of ~39 kJ/mol or $\sim 15.7 k_B T$. The mutations present in $CI2_{FTS\text{-}opt}$ thus result on a sizable reduction of 37% in the barrier height at the denaturation midpoint. The estimated midpoint barrier for $CI2_{FTS\text{-}opt}$ is still high enough for the protein to exhibit pure two-state-like kinetics (Figure 4c). However, the estimation for the $CI2_{FTS\text{-}opt}$ folding barrier in water, or physiological conditions, becomes only 6.7 kJ/mol or $2.7 k_B T$, which is marginally above thermal energy. We note that these values are only estimates since the actual pre-exponential for CI2 is unknown. As an alternative, the flattening of the chevron plot can be used as a correlative approach for estimating overall (un)folding barriers, one which does not require knowledge of the pre-exponential term (Naganathan et al., 2007). Following that procedure, we find that a flattening of the chevron down to 87% together with the ~10 s^{-1} midpoint folding (or unfolding) rate of $CI2_{FTS\text{-}opt}$ corresponds to proteins that have marginal folding barriers between 2 and $3.5 k_B T$ in water (Naganathan et al., 2007). Therefore, both empirical estimates are consistently suggesting that $CI2_{FTS\text{-}opt}$ might be close to the downhill folding scenario in native conditions (Campos et al., 2020). The conjoined estimates indicate that, overall, the FTS for $CI2_{FTS\text{-}opt}$ in water has only 32% left of the free energy of the wild-type's, which is a large scale reduction in relative terms.

3 | CONCLUSIONS

The general understanding from many decades of work is that protein (un)folding rates are largely determined by domain size (De Sancho et al., 2009; Naganathan & Muñoz, 2005) and the topology of the native fold (Muñoz & Eaton, 1999; Plaxco et al., 1998). Mutations appear to only tune these structural factors through their effects in the stability of the native structure (Naganathan & Muñoz, 2010). It is also often considered that evolutionary pressure selects for proteins that can fold sufficiently fast to reach their functional states within the relevant biological timescales and avoid misfolding (Gelman & Gruebele, 2014). Herein we have managed to simultaneously accelerate by over 250-fold the folding and unfolding rates of the protease inhibitor CI2, a two-state folding protein with a high free energy FTS that has proven impervious to mutations. To the best of our knowledge, both, the ability to selectively target

the FTS by mutations and the magnitude of the rate acceleration, are unprecedented. Moreover, we achieved such vast kinetic acceleration by simply removing the terminal positive charges of two arginine residues, which form a buried salt bridge network in CI2's native structure.

The large FTS stabilization via elimination of the buried salt bridge network results on a CI2 variant that has the same thermodynamic stability, but loses entirely its inhibitory activity and becomes susceptible to quick degradation by chymotrypsin, its natural substrate. From these results we conclude that the high (un)folding free energy barrier naturally found in CI2 is essential for its biological function and homeostasis. In other words, CI2 appears to have been naturally selected through evolution to (un)fold very slowly. Previous work concluded that protein sequences are not selected for rapid folding per se, but for thermodynamic stability (Kim et al., 1998). CI2 is clearly in stark contrast, as its selective pressure appears to be mostly kinetic and going in the direction of slowing down (un)folding. Sequence analysis provides additional support to this idea. The 100% conservation of R46 in CI2 homologs (sup. inf.) suggests that the salt bridges between R46 and the C-terminal carboxylate, which we show are critical to destabilize CI2's FTS (Figure 2c), are indeed functionally essential. R48 is occasionally replaced by I or F in members of the CI2 family (Figure S1). In this regard it is noteworthy that, out of the 11 mutations attempted by Lawrence et al. (2010) in position 48, R48I, and R48F are the only ones that maintained a very slow unfolding rate. W is also found in position 48 in two family members (entries P82011 and P82381 in Figure S1). In CI2, the R48W mutation, in contrast to I and F, does result in much faster unfolding (Lawrence et al., 2010). However, P82011 and P82381 belong to the group of CI2 homologs that form a disulfide bond between the N-terminus and the functional loop's end (Campos et al., 2019). This disulfide bond should impede the opening of the protein, thus providing an alternative solution for attaining high kinetic stability. We thus reason that the folding properties we report here for CI2 respond to the need of protease inhibitors to achieve high kinetic stability in order to withstand degradation by its own substrates (Sanchez-Ruiz, 2010). The implication is that the high energy FTS of CI2 is a direct product of functional rather than structural constraints.

Our findings on CI2 also have fundamental implications for our understanding of the factors that determine protein (un)folding rates. We find that what determines the slow kinetics of CI2 is the burial of several charges prior to forming the salt bridge network. This result explains why the CI2 FTS is so refractory to scanning mutations (Naganathan & Muñoz, 2010), and why the (un)folding rates of this protein are much slower than expected for its size (De Sancho et al., 2009). More generally, the fact that only two, very specific, aminoacid changes in the CI2 sequence can induce such extreme rate accelerations sheds important light on a long standing puzzle in protein folding. The puzzle is as follows: after accounting for size and structural class, the (un)folding rates of single-domain proteins still fluctuate by 20-fold from the expected correlation line (De Sancho & Muñoz, 2011); but the rate effects caused by single-point mutations are strikingly uniform (Naganathan & Muñoz, 2010). These observations bear the questions of what factor is responsible for the remaining rate variability and how we can further improve rate predictions. The CI2 case suggests that each protein may have just a few select residues-interactions that define the basic energetics of its FTS over general structural considerations of size and fold topology, and which might be required to meet specific functional and homeostatic requirements. Such a protein architecture would provide resilience to mutational drift through evolution as well as specific targets for natural selection, as we find here for CI2. By the same token, the few critical residues-interactions could be easily missed in mutational scanning strategies, explaining the homogeneity of folding mutational data. In practical terms, this idea suggests that in order to achieve better predictions of (un)folding rates it might be important to shift the focus to searching for the residues and interactions that are key for each protein's FTS. Similarly to functional sites (Ferreiro et al., 2018), such FTS–centric interactions are likely to result in certain energetic frustration of the native structure, opening an avenue for their identification via computational methods (Parra et al., 2016).

4 | MATERIALS AND METHODS

4.1 | Cloning of CI2 variants

The sequence for the wild type CI2 used in this study is the following (numeration according to previous works on CI2):

MDLKTEWPELVGKSVEEAKKVILQDKPEAQIIVLPVGTIVTMEYRIDRVRLFVDKLDNIAEVPRVG

1 10 20 30 40 50 60

The individual positions mutated for the variants of this study are highlighted and boxed. Gene sequences were directly synthesized using gene optimization (Top Gene technologies, Inc.) or produced with the QuikChange II site-directed mutagenesis protocol (Agilent Technologies) using the wild type CI2 clone as template. Genes were cloned into the pBAT4 plasmid and transformed in BL21-Gold(DE3) *E. coli* competent cells (Agilent Technologies).

4.2 | Expression and purification of CI2 variants

LB medium with 100 mg/L ampicillin was used to grow a cell culture for all CI2 variants. Cell cultures were grown at 37°C to an OD_{600} of 1.2, at which time recombinant protein expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at 100 mg/L overnight at 30°C. Cells were harvested at 8000 *g* for 30 min, lysed by several freeze–thaw cycles and centrifuged down at 60,000 g for 60 min. Protein purification was performed by HPLC with an anionic exchange column (HiTrap Q HP from GE Healthcare Life Sciences) in 20 mM sodium borate buffer at pH 8.5 and a gradient of 0–1 M NaCl. Fractions containing the protein were confirmed by SDS-PAGE gel electrophoresis, pooled, dialyzed in the same buffer to remove salt and injected, after changing the pH to 2.5, onto a reverse phase column (Proto 300 C4 column from Western Analytical products) for a second purification step with an acetonitrile gradient. Fractions containing >99% pure CI2 variant protein were confirmed by SDS-PAGE, pooled, confirmed by mass spectrometry (MALDI-TOF), and freeze-dried.

4.3 | Equilibrium chemical denaturation experiments

Stock samples were prepared dissolving the lyophilized protein in 7 M GdmCl to eliminate potential non-specific aggregates induced by the freeze-dry procedure, and then dialyzed extensively against 20 mM sodium borate at pH 8.5. Protein solutions at different concentrations of GdmCl were prepared from stocks to a final protein concentration of 10 μM using the same buffer. Final GdmCl concentrations were determined from the refractive index of each sample measured by refractometry. Fluorescence spectra with excitation at 280 nm were collected at 298 K using a Jobin Yvon Fluorolog-3 spectrofluorometer from Horiba. The total fluorescence emission was calculated from the spectra and the data set for all the GdmCl concentration were fitted to the two-state equation:

$$F = \frac{(F_n + m_n[\mathrm{GdmCl}]) + (F_u + m_u[\mathrm{GdmCl}])e^{-(\Delta G_w - m[\mathrm{GdmCl}])/RT}}{(1 + e^{-(\Delta G_w - m[\mathrm{GdmCl}])/RT})}$$

where F_n and F_u are the fluorescence signals of the native and unfolded state at 0M GdmCl, respectively, m_n and m_u are the dependencies of the fluorescence signal with GdmCl concentration for the native and the unfolded state, ΔG_w is the folding free energy in water, m is the derivative of the folding free energy with respect to GdmCl concentration, R is the gas constant and T is the temperature in Kelvin. The fitting parameters were used to calculate the folded fraction at each GdmCl concentration.

4.4 | Equilibrium thermal denaturation experiments

Stock samples were prepared following the same procedure used for equilibrium chemical denaturation experiments. Protein solutions were prepared to a final protein concentration of 10 μM using 20 mM sodium borate at pH 8.5 as buffer. Fluorescence spectra with excitation at 280 nm were collected at 298 K using a Jobin Yvon Fluorolog-3 spectrofluorometer from Horiba. The total fluorescence emission was calculated from the spectra and the data set for all temperatures were fitted to the two-state equation:

$$F = \frac{(F_n + m_n T) + (F_u + m_u T)e^{-(\Delta H_{T_m}(1-(T/T_m))/RT)}}{(1 + e^{-(\Delta H_m(1-(T/T_m))/RT)})}$$

where F_n and F_u are the fluorescence of the native and unfolded state at $T = 300$ K, m_n and m_u are the dependences of the fluorescence signal with temperature for the native and the unfolded state, T_m is transition midpoint temperature in Kelvin, ΔH_m is the denaturation enthalpy at the transition midpoint, R is the gas constant

and T is the temperature in Kelvin. The fitting parameters were used to calculate the folded fraction for every temperature.

4.5 | Folding–unfolding kinetic experiments

Kinetic measurements of the folding–unfolding relaxation for all the CI2 variants were performed at 298 K on a SX20-LED Stopped Flow fluorometer from Applied Photophysics. The lyophilized protein was dissolved in 7 M GdmCl and diluted to a final protein concentration of 10 μM in either a low or a high final concentration of GdmCl. Unfolding experiments were performed by mixing the low GdmCl concentration stock with solutions of GdmCl at different concentrations (in 20 mM sodium borate buffer at pH 8.5) at a 1:10 ratio. Refolding experiments were performed the same way but with the high GdmCl protein stock. The fluorescence emission above 320 nm was collected with a long pass filter after excitation at 280 nm. Ten individual kinetic traces without signs of mixing artifacts were averaged for each experimental condition. The resulting averaged traces were fitted to a single exponential function (for unfolding) or to a single exponential plus a linear slope to account for the slow proline isomerization process of CI2 (for refolding) (Jackson & Fersht, 1991).

4.6 | Activity measurements

The protease inhibitor activity of CI2 variants was determined using a colorimetric assay based on the catalytic conversion of N-benzoyl-L-tyrosine ethyl ester (BTEE) to N-benzoyl-L-tyrosine by the enzyme chymotrypsin. Assays in the absence and presence of CI2 were performed following the protocol indicated by Sigma. Different concentrations of BTEE (from 10 to 1000 μM) were mixed with 53 mM $CaCl_2$, 200 nM of CI2 wild-type or $CI2_{FTS\text{-}opt}$ (or none) and 53 nM chymotrypsin to a final volume of 3 mL in 50 mM Tris buffer at pH 7.8. After addition of chymotrypsin, the sample was manually mixed and introduced in a Cary 100 bio UV–Visible spectrophotometer from Varian to follow the change in absorbance at 256 nm. The initial velocity of conversion is determined by the change in absorbance during the first seconds of the reaction fitted to a straight line. Initial velocities obtained for every different concentration of substrate were fitted to equation:

$$V_0 = \frac{V_{max}[S]}{K_m^{app} + [S]}$$

where V_0 is a vector with the measured initial velocities, V_{max} corresponds to the maximal velocity, [S] is a vector with the BTEE concentrations used in the assay and $K_m{}^{app}$ is the apparent Michaelis–Menten constant, which can be expressed as,

$$K_m^{app} = K_m\left(1 + \frac{[I]}{K_i}\right)$$

where K_m is the Michaelis–Menten constant, [I] is the concentration of inhibitor, and K_i is the dissociation constant for the inhibitor.

4.7 | Proteolysis assay

We used a protocol modified from the proteolysis assay we used for determining the inhibitor activity (see above). Briefly, we prepared CI2 variant samples at 40 μM by mixing with 133 mM $CaCl_2$, 10% methanol, and 1 μM chymotrypsin in 10 mM sodium borate buffer at pH 8.3 (final volume of 400 μL). The proteolysis reaction was started with the addition of chymotrypsin, and stopped at different times by adding 4 μL of 1M HCl. After stopping the reaction, samples for different times were loaded in a Superdex 200 Increase 10/300 GL size exclusion column (from GE Healthcare) run on a 1260 Infinity HPLC system (from Agilent Technologies). The absorbance for the peak maximum corresponding to the undigested CI2 variant was recorded to determine the amount of proteolysis as a function of time. Finally, the obtained data were fitted to a single exponential decay function.

4.8 | Sequence alignment for CI2 homologs

Twenty-six sequences corresponding to proteins previously included in the Potato Inhibitor I family were obtained from the Interpro webpage and aligned using ClustalW, followed by manual refinement. A larger alignment of 149 sequences was also obtained collecting homologous proteins obtained with BLAST and then aligned using ClustalW followed by manual refinement.

4.9 | Structure prediction for the TS_{opt} CI2 variant

The native structure of the CI2 variant containing the four mutations (TS_{opt}) was predicted with AlphaFold 2.0 using the ColabFold v1.5.5 server. The structure was

generated with the following parameters: pair-mode = unpaired-paired, num_recycles = 20, tol = 0.0, relax_max_iterations = 200, pairing_strategy = greedy. The calculation produced five different models from which we selected number four as the one with the highest rank in the lDDT-Cα score.

AUTHOR CONTRIBUTIONS

Luis A. Campos: Methodology; investigation; validation; writing – review and editing; writing – original draft; formal analysis; visualization; conceptualization. **Victor Muñoz:** Conceptualization; validation; formal analysis; project administration; funding acquisition; writing – review and editing; supervision.

ACKNOWLEDGMENTS

The research described in this article was supported by grants: ERC-2012-ADG-323059 (European Research Council), MCB-2112710 (National Science Foundation) and HRD-2112675 (National Science Foundation).

ORCID

Victor Muñoz https://orcid.org/0000-0002-5683-1482

REFERENCES

Akmal A, Muñoz V. The nature of the free energy barriers to two-state folding. Proteins. 2004;57(1):142–52.

Anil B, Sato S, Cho JH, Raleigh DP. Fine structure analysis of a protein folding transition state; distinguishing between hydrophobic stabilization and specific packing. J Mol Biol. 2005;354(3):693–705.

Campos LA. Mutational analysis of protein folding transition states: phi values. Methods Mol Biol. 2022;376:3–30.

Campos LA, Sharma R, Alvira S, Ruiz FM, Ibarra-Molero B, Sadqi M, et al. Engineering protein assemblies with allosteric control via monomer fold-switching. Nat Commun. 2019;10(1):5703.

Campos LA, Sadqi M, Muñoz V. Lessons about protein folding and binding from archetypal folds. Acc Chem Res. 2020;53(10):2180–8.

Chen Y, Barkley MD. Toward understanding tryptophan fluorescence in proteins. Biochemistry. 1998;37(28):9976–82.

de los Rios MA, Muralidhara BK, Wildes D, Sosnick TR, Marqusee S, Wittung-Stafshede P, et al. On the precision of experimentally determined protein folding rates and phi-values. Protein Sci. 2006;15(3):553–63.

De Sancho D, Muñoz V. Integrated prediction of protein folding and unfolding rates from only size and structural class. Phys Chem Chem Phys. 2011;13(38):17030–43.

De Sancho D, Doshi U, Muñoz V. Protein folding rates and stability: how much is there beyond size? J Am Chem Soc. 2009;131(6):2074–5.

Dill KA, MacCallum JL. The protein-folding problem, 50 years on. Science. 2012;338(6110):1042–6.

Ferreiro DU, Komives EA, Wolynes PG. Frustration, function and folding. Curr Opin Struct Biol. 2018;48:68–73.

Fersht A. Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding. New York: Macmillan; 1999.

Fersht AR. Characterizing transition states in protein folding: an essential step in the puzzle. Curr Opin Struct Biol. 1995;5(1):79–84.

Fersht AR, Matouschek A, Serrano L. The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. J Mol Biol. 1992;224(3):771–82.

Gelman H, Gruebele M. Fast protein folding kinetics. Q Rev Biophys. 2014;47(2):95–142.

Itzhaki LS, Otzen DE, Fersht AR. The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. J Mol Biol. 1995;254(2):260–88.

Jackson SE, Fersht AR. Folding of chymotrypsin inhibitor 2.1. Evidence for a two-state transition. Biochemistry. 1991;30(43):10428–35.

Jackson SE, Fersht AR. Contribution of residues in the reactive site loop of chymotrypsin inhibitor 2 to protein stability and activity. Biochemistry. 1994;33(46):13880–7.

Jackson SE, elMasry N, Fersht AR. Structure of the hydrophobic core in the transition state for folding of chymotrypsin inhibitor 2: a critical test of the protein engineering method of analysis. Biochemistry. 1993;32(42):11270–8.

Jaenicke R. Stability and folding of domain proteins. Prog Biophys Mol Biol. 1999;71(2):155–241.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with alphafold. Nature. 2021;596(7873):583–9.

Kim DE, Gu H, Baker D. The sequences of small proteins are not extensively optimized for rapid folding by natural selection. Proc Natl Acad Sci U S A. 1998;95(9):4982–6.

Ladurner AG, Itzhaki LS, Daggett V, Fersht AR. Synergy between simulation and experiment in describing the energy landscape of protein folding. Proc Natl Acad Sci U S A. 1998;95(15):8473–8.

Lawrence C, Kuge J, Ahmad K, Plaxco KW. Investigation of an anomalously accelerating substitution in the folding of a prototypical two-state protein. J Mol Biol. 2010;403(3):446–58.

Modi T, Campitelli P, Kazan IC, Ozkan SB. Protein folding stability and binding interactions through the lens of evolution: a dynamical perspective. Curr Opin Struct Biol. 2021;66:207–15.

Muñoz V. Conformational dynamics and ensembles in protein folding. Annu Rev Biophys Biomol Struct. 2007;36:395–412.

Muñoz V, Eaton WA. A simple model for calculating the kinetics of protein folding from three-dimensional structures. Proc Natl Acad Sci U S A. 1999;96(20):11311–6.

Naganathan AN, Muñoz V. Scaling of folding times with protein size. J Am Chem Soc. 2005;127(2):480–1.

Naganathan AN, Muñoz V. Insights into protein folding mechanisms from large scale analysis of mutational effects. Proc Natl Acad Sci U S A. 2010;107(19):8611–6.

Naganathan AN, Doshi U, Muñoz V. Protein folding kinetics: barrier effects in chemical and thermal denaturation experiments. J Am Chem Soc. 2007;129(17):5673–82.

Onuchic JN, Luthey-Schulten Z, Wolynes PG. Theory of protein folding: the energy landscape perspective. Annu Rev Phys Chem. 1997;48(1):545–600.

Ozkan SB, Bahar I, Dill KA. Transition states and the meaning of φ-values in protein folding kinetics. Nat Struct Biol. 2001;8(9):765–9.

Parra RG, Schafer NP, Radusky LG, Tsai M-Y, Guzovsky AB, Wolynes PG, et al. Protein frustratometer 2: a tool to localize energetic frustration in protein molecules, now with electrostatics. Nucleic Acids Res. 2016;44(W1):W356–60.

Plaxco KW, Simons KT, Baker D. Contact order, transition state placement and the refolding rates of single domain proteins. J Mol Biol. 1998;277(4):985–94.

Radisky ES, Lu C-JK, Kwan G, Koshland DE. Role of the intramolecular hydrogen bond network in the inhibitory power of chymotrypsin inhibitor 2. Biochemistry. 2005;44(18):6823–30.

Sanchez IE, Kiefhaber T. Origin of unusual phi-values in protein folding: evidence against specific nucleation sites. J Mol Biol. 2003;334(5):1077–85.

Sanchez-Ruiz JM. Protein kinetic stability. Biophys Chem. 2010; 148(1–3):1–15.

Toyama BH, Hetzer MW. Protein homeostasis: live long, won't prosper. Nat Rev Mol Cell Biol. 2013;14(1):55–61.

Viguera AR, Vega C, Serrano L. Unspecific hydrophobic stabilization of folding transition states. Proc Natl Acad Sci U S A. 2002; 99(8):5349–54.

Waldburger CD, Jonsson T, Sauer RT. Barriers to protein folding: formation of buried polar interactions is a slow step in acquisition of structure. Proc Natl Acad Sci U S A. 1996;93(7):2629–34.

Wright PE, Dyson HJ. Linking folding and binding. Curr Opin Struct Biol. 2009;19(1):31–8.

Zarrine-Afsar A, Davidson AR. The analysis of protein folding kinetic data produced in protein engineering experiments. Methods. 2004;34(1):41–50.

Zarrine-Afsar A, Wallin S, Neculai AM, Neudecker P, Howell PL, Davidson AR, et al. Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding. Proc Natl Acad Sci U S A. 2008;105(29):9999–10004.

Zhu Y, Fu X, Wang T, Tamura A, Takada S, Saven JG, et al. Guiding the search for a protein's maximum rate of folding. Chem Phys. 2004;307(2):99–109.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Campos LA, Muñoz V. Targeting the protein folding transition state by mutation: Large scale (un)folding rate accelerations without altering native stability. Protein Science. 2024;33(7):e5031. https://doi.org/10.1002/pro.5031