doi:10.1016/j.jmb.2009.03.074 J. Mol. Biol. (2009) 389, 90–102

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Characterization of the Cofactor-Induced Folding Mechanism of a Zinc-Binding Peptide Using Computationally Designed Mutants

Jia Tang†, Seung-Gu Kang†, Jeffery G. Saven* and Feng Gai*

Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, USA

Received 7 January 2009; received in revised form 26 March 2009; accepted 31 March 2009
Available online 8 April 2009

Metals are the most commonly encountered protein cofactors, and they play important structural and functional roles in biology. In many cases, metal binding provides a major driving force for a polypeptide chain to fold. While there are many studies on the structure, stability, and function of metal-binding proteins, there are few studies focusing on understanding the kinetic mechanism of metal-induced folding. Herein, the Zn^{2+}-induced folding kinetics of a small zinc-binding protein are studied; the $CH1_1$ peptide is derived from the first cysteine/histidine-rich region (CH1 domain) of the protein interaction domains of the transcriptional coregulator CREB-binding protein. Computational design is used to introduce tryptophan and histidine mutations that are structurally consistent with $CH1_1$; these mutants are studied using stopped-flow tryptophan fluorescence experiments. The Zn^{2+}-induced $CH1_1$ folding kinetics are consistent with two parallel pathways, where the initial binding of Zn^{2+} occurs at two sites. However, the initially formed Zn^{2+}-bound complexes can proceed either directly to the folded state where zinc adopts a tetrahedral coordination or to an off-pathway misligated intermediate. While elimination of those ligands responsible for misligation simplifies the folding kinetics, it also leads to a decrease in the zinc binding constant. Therefore, these results suggest why these nonnative zinc ligands in the $CH1_1$ motif are conserved in several distantly related organisms and why the requirement for function can lead to kinetic frustration in folding. In addition, the loop closure rate of the $CH1_1$ peptide is determined based on the proposed model and temperature-dependent kinetic measurements.

Keywords: protein folding; computational design; stopped-flow fluorescence; zinc-binding peptides; folding intermediate

Edited by C. R. Matthews

Introduction

The study of the folding mechanism of single-domain proteins that lack cofactors has reached an advanced stage wherein the sequence of events along the course of refolding after denaturation begins to be understood.[1–10] In several cases, it has been even possible to describe in detail the energetics, position, and structure of the folding transition-state ensembles and/or intermediate states involved.[2,3,7–9] In contrast, there are far fewer studies on cofactor-induced folding kinetics, even though a large number of proteins require cofactors for proper folding and/or function.[11,12] While the effect of cofactor binding on structure and stability may differ for different proteins, many polypeptides can only attain a well-folded structure when associated with their respective cofactors. For example, cytochrome *c*, an important electron transfer protein, cannot fold without incorporation of its heme cofactor.[12] Similarly, many small metal-binding peptides, such as the zinc-finger motif,[13,14] also require their respective cofactors to maintain the integrity of their native fold. For such systems, cofactor binding thus provides not only stabilization but also a major driving force directing the folding of an otherwise unstruc-

*Corresponding authors. E-mail addresses: saven@sas.upenn.edu; gai@sas.upenn.edu.

† J.T. and S.-G.K. contributed equally to this work.

Abbreviation used: SCADS, statistical computationally assisted design strategy.

tured protein to a well-defined structure. Nonetheless, the prerequisite of a cofactor for folding may complicate the folding free-energy landscape and thus the kinetic mechanism[15] because cofactor binding often requires the formation of a precisely defined cofactor-binding environment in the native state. In addition, a protein may also transiently associate with a cofactor in a nonnative fashion, leading to kinetic frustration in folding. For example, cytochrome *c* exhibits rather complex folding kinetics arising from nonnative iron ligation during or prior to folding.[16–20] Herein, the folding of the designed variants of a small Zn^{2+}-binding protein is examined in an effort to probe the roles of native and nonnative interactions in the folding kinetics of metal-binding proteins.

Metal ions such as Zn^{2+}, Mg^{2+}, and Ca^{2+} are among the most commonly encountered cofactors. Metal binding not only plays many important structural and functional roles in cells[21] but also is increasingly used in *de novo* protein designs to generate or stabilize a specific fold.[22–32] For example, when coordinated with a metal ion, polypeptides composed of several heptad repeats are observed to undergo transition from a disordered conformation to an α-helical coiled-coil structure.[24,32] Despite the increase in the number of metal-binding peptides (or domains) designed and discovered in recent years, studies on metal-induced folding kinetics are scarce.[25,33–35] Moreover, small miniproteins have led to advances in our understanding of elementary events in protein folding and detailed comparisons between theory and experiment.[36–39] Similarly, it is useful to identify and develop small cofactor-containing miniproteins that lend themselves to kinetic folding studies. To provide further insight into understanding the mechanism of metal-mediated protein folding processes, herein we have carried out detailed stopped-flow studies of the Zn^{2+}-induced folding kinetics of computationally designed variants of a zinc-binding motif, the $CH1_1$ peptide.

The $CH1_1$ peptide (sequence ^{1}EVRACSLPHC–^{11}RTMKNVLNHM–^{21}THCQAGK) is a highly conserved motif found in the first cysteine/histidine-rich (Cys/His-rich) region (CH1 domain) of the protein interaction domains of the transcriptional coregulator CREB-binding protein.[40] NMR studies have revealed that the $CH1_1$ peptide binds a single Zn^{2+} through coordination with the side chains of C5, C10, H19, and C23 residues, and that the folded structure is composed of a 3_{10}-helix and a short α-helical stretch packing around the zinc (Fig. 1). The $CH1_1$ peptide has also been termed CHANCE finger (or a Cys/His-rich peptide exhibiting an unexpected conformational ensemble) as it constitutes a new protein fold.[40] Previous studies[40,41] have shown that this CHANCE fold is not only thermally stable (no substantial secondary structural changes occur up to 85 °C, as judged by CD) but also tolerant to both multiple alanine (Ala) mutations and changes in the sequence spacing of zinc-ligating ligands. In addition, it is found that the two nonnative ligands, especially His22, are conserved in the $CH1_1$ sequence

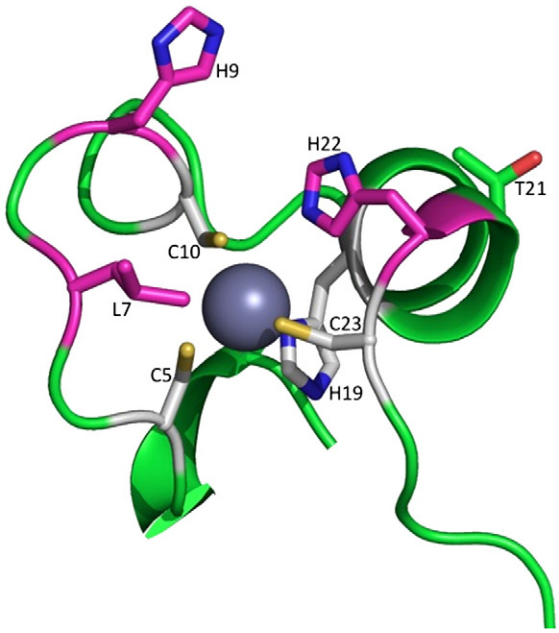


Fig. 1. The NMR structures of the $CH1_1$ Zn-binding domain (Protein Data Bank code 1liq).[40] Zn-coordinating residues (gray) and residues of the triple mutant L7W/H9V/H22I (magenta) are shown.

in several distantly related organisms,[40] suggesting that they may be functionally important, for example, by facilitating ligand binding via nonnative interactions. However, such interactions are expected to occur at the expense of introducing frustration into the folding kinetics. Thus, the advantages of studying the folding kinetics of the $CH1_1$ peptide are multifold: (a) the robustness of the peptide sequence renders multiple mutations feasible, a necessity for elucidating the folding mechanism; (b) the small size of the peptide also makes it feasible to achieve a mechanistic description of the folding process; (c) it allows one to examine the kinetic role of those nonnative ligands, thus providing a rationalization of their possible functional role; and (d) because of its small size, this miniprotein is amenable and should provide a rich model system for further computer simulation studies.

To study the zinc-induced folding process of the $CH1_1$ peptide, it is useful to introduce a probe that is sensitive to local structure and to mutate potential nonnative ligands H9 and H22 while retaining the parent structure. For the $CH1_1$ variants studied herein, a single tryptophan (Trp) residue was introduced to serve as a fluorescent probe of structure, since both cysteine (Cys) and histidine (His) side chains are efficient quenchers of Trp fluorescence.[42] Although the peptide is tolerant of multiple alanine mutations, mutations involving larger residues may not be compatible with folding to the native structure. Computational design methods facilitate the identification of viable mutations. Four structurally consistent mutants (L7W, T21W, L7W/H9V, and L7W/H9V/H22I) were designed using a

statistical computationally assisted design strategy (SCADS).[43,44] The folding process was initiated by mixing a Zn^{2+} solution with a $CH1_1$ variant solution using a stopped-flow technique, and the folding kinetics were monitored by following changes in Trp fluorescence intensity. The Zn^{2+}-induced folding kinetics of the $CH1_1$ peptide are complex and involve at least three kinetic phases. Based on extensive concentration and temperature-dependent studies, a model that is capable of describing the observed stopped-flow kinetics of these mutants is proposed. This model suggests that the binding of Zn^{2+} to the unfolded $CH1_1$ peptide involves parallel pathways, and that the initially formed Zn^{2+}-bound peptide can proceed either directly to the native state or to an off-pathway misligated state.

Results

Identification of Trp-tag sites

Native fluorescence, especially that arising from Trp, is often used in protein folding and binding studies. However, the native sequence of the $CH1_1$ peptide does not contain a suitable fluorophore for monitoring folding or cofactor binding. Therefore, we have employed a statistical computational design method (i.e., SCADS) to identify Trp-containing mutants consistent with the known structure of $CH1_1$.[40] While the calculations suggest that several sites may be tolerant of Trp mutation, two single mutants L7W and T21W were targeted because, in both cases, the Trp residue is in proximity to two zinc ligands: C5/C10 for L7W and H19/C23 for T21W. The probabilities of these two Trp residues at their respective positions are shown in Fig. 2. As shown in the predicted model structures (Fig. 2), Trp7 in L7W is located close to the metal-binding pocket; thus, it is expected to be sensitive to Zn^{2+}–ligand coordination. On the other hand, while the side chain of Trp21 in T21W is highly exposed to solvent (Fig. 2), it is near the zinc-binding site, and the location (i.e., sandwiched between H19 and C23) makes it a good candidate for the observation of local environmental changes associated with transition from a heterogeneous ensemble of conformations in the unfolded state to a well-organized structure upon folding.

Identification of mutations for His9 and His22

To elucidate the roles of the two nonbinding His residues (i.e., H9 and H22) in $CH1_1$ folding, the site-specific amino acid probability profiles for sites 9 and 22 were calculated. At each position, 19 amino acids were permitted (i.e., Cys was excluded), and the modeled structure of the mutant L7W was used. As shown (Fig. 2), glutamic acid (Glu) shows a probability for site 9 that is slightly higher than those for several other amino acids at $T_{eff}=2$ kcal mol^{-1}. However, at lower energies ($T_{eff}=0.5$ kcal mol^{-1}), the aliphatic nonpolar amino acid valine (Val) is the most probable amino acid for this site due to stabilizing short-ranged van der Waals interactions.

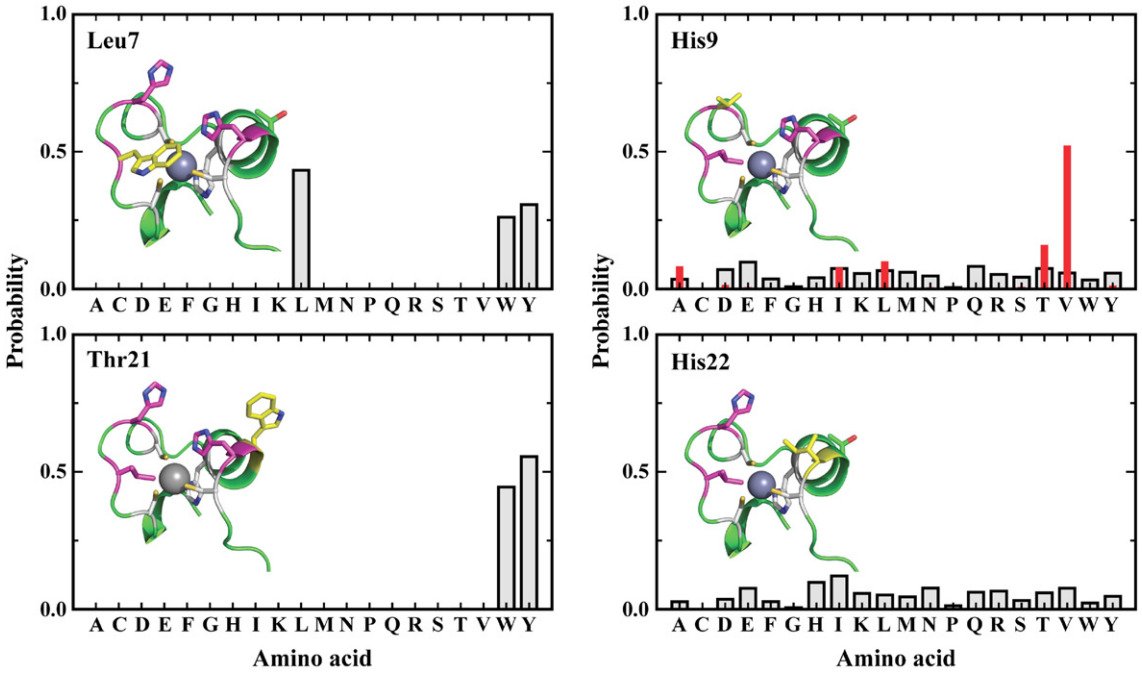


Fig. 2. Site-specific amino acid probability profiles for the introduction of Trp mutation on L7 and T21 (left) and His replacement for H9 and H22 (right). $T_{eff}=0.5$ kcal mol^{-1}. For H9, probabilities at two different sampling temperatures of 2.0 and 0.5 kcal mol^{-1} are presented with gray squares and red solid bars, respectively. A mutated residue appears in yellow. Zn-coordinating residues (gray) and residues of the triple mutant L7W/H9V/H22I (magenta) are shown.

With the goal of testing the function of metal-reactive His at position 9, it would be desirable to replace His with an amino acid that is neutral and unlikely to coordinate with the metal ion. Thus, Val was identified to replace His9, resulting in the double mutant L7W/H9V.

Further calculations were performed to replace His22 based on L7W/H9V. As shown (Fig. 2), isoleucine (Ile) and Val are highly probable amino acids in addition to His and have average conformational energies similar to that of His. However, taking into consideration that Ile has a higher helical propensity than Val for an α-helical motif,[45] Ile was selected to substitute His22, resulting in the triple mutant L7W/H9V/H22I.

CD studies

The CD spectra of $CH1_1$ mutants in aqueous solution, as that shown in Fig. S1 (Supporting Information) for apo L7W, show characteristics of unfolded peptides. Addition of $ZnCl_2$ to the peptide solution, however, leads to a distinct change in the CD spectrum: the development of bands around 208 and 222 nm, respectively, consistent with the formation of a partial helical structure (Supporting Information, Fig. S1). These results are in agreement with the study of Sharpe *et al.*, who have shown that, upon association with Zn^{2+}, the $CH1_1$ peptide folds into a structure that contains a short 3_{10}-helix and a short α-helical stretch.[40] Addition of 5 mM ethylenediaminetetraacetic acid into the peptide–zinc solution results in loss of the helical CD signal (data not shown) due to the chelation of Zn^{2+} by ethylenediaminetetraacetic acid,[46] further demonstrating the important structural role of Zn^{2+} in the folding of these $CH1_1$ mutants. In addition, one-dimensional proton NMR measurements also indicate that the zinc-bound peptide adopts a folded conformation (for an example, see Fig. S2 in Supporting Information).

Moreover, the CD spectra of the $CH1_1$ mutants considered here indicate that the helical content of each peptide is somewhat different. This is likely due to a combination of effects arising from the difference in the helix-forming propensities of different amino acids[45] and/or structural constraints induced by the presence or the absence of H9 and/or H22.[47] Indeed, Sharpe *et al.* have shown that the mean residual helicity of the $CH1_1$ peptide increases when 15 native residues are replaced with Ala.[40]

Fluorescence titrations

The Trp fluorescence intensity of these $CH1_1$ mutants increases upon association with Zn^{2+} and is therefore used to determine their apparent Zn^{2+}-binding constants. As an example, the fluorescence spectra of L7W obtained at various Zn^{2+} concentrations are shown (Fig. 3a). Interestingly, the fluorescence spectra of all mutants exhibit a common maximum centered at 354±1 nm that is independent of the concentration of Zn^{2+}, indicating that the Trp

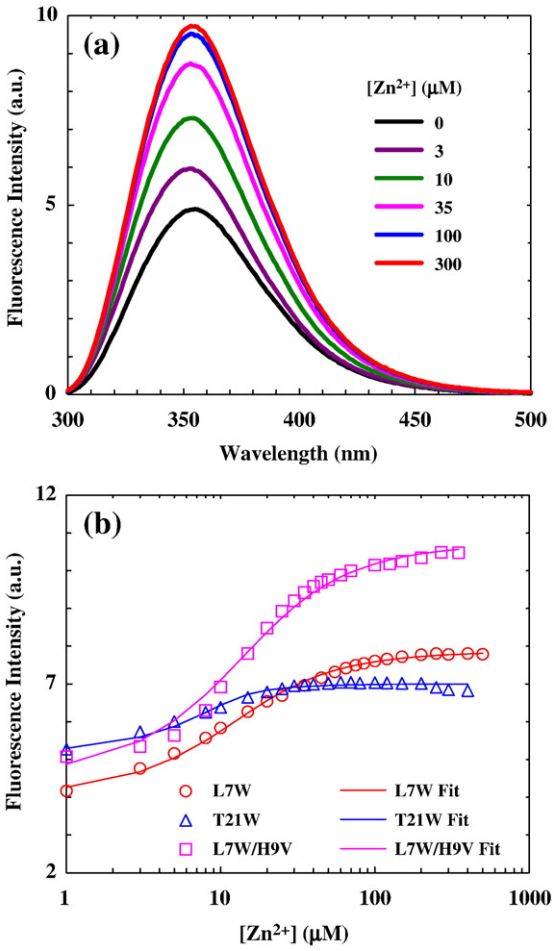


Fig. 3. (a) Representative Trp fluorescence spectra of L7W at various Zn^{2+} concentrations, as indicated. The peptide concentration was 10 μM (in 10 mM Tris buffer, pH 6.5). (b) Integrated area of the Trp fluorescence of $CH1_1$ mutants *versus* Zn^{2+} concentration. Lines are fits of these data to the binding model described in the text.

side chain is largely exposed to solvent in both folded and unfolded states. Thus, the increase in Trp fluorescence upon Zn^{2+} binding must arise from an alleviation of the intramolecular quenching of Trp fluorescence by various amino acids,[42,48–50] most likely Cys and His residues in the current case. The study of Chen and Barkley has shown that Cys and protonated His are among the strongest quenchers of Trp fluorescence existing in natural proteins.[42] As described above, the single Trp in all of these $CH1_1$ mutants is deliberately introduced through design at a position that is in close proximity to, at least, a pair of putative zinc-coordinating residues: C5/C10 or H19/C23. Therefore, it is expected that the Trp fluorescence of these peptides increases upon association with Zn^{2+} because coordination with a metal ion greatly reduces the quenching efficiency of these residues.[48]

As indicated (Fig. 3b), the total Trp fluorescence intensity (i.e., the integrated area of the fluorescence spectrum) of these $CH1_1$ mutants as a function of Zn^{2+} concentration can be described by a 1:1 binding

model. The apparent dissociation constants (K_d) determined from the best fit of these data to this model (i.e., Eqs. (1) and (4)) are K_d (L7W) = 3.8 ± 0.5 μM, K_d (T21W) = 0.52 ± 0.35 μM, and K_d (L7W/H9V) = 4.6 ± 0.8 μM. Because the triple mutant L7W/H9V/H22I was found to form aggregates at relatively high Zn^{2+} concentrations, we did not attempt to determine its K_d.

Stopped-flow kinetics

The Zn^{2+}-induced folding kinetics of these $CH1_1$ mutants were investigated by a stopped-flow fluorescence technique.[51] Consistent with equilibrium measurements, rapid mixing of a Zn^{2+} solution with a peptide solution at pH 6.5 leads the Trp fluorescence to increase as a function of time. However, the stopped-flow kinetics thus obtained for these mutants are quite complex. As shown (Fig. 4a), the stopped-flow traces of L7W, T21W, and L7W/H9V are adequately described by a triexponential function with well-separated rate constants (Table 1). While such phenomenological kinetic analyses do not provide molecular insights into the Zn^{2+}-induced folding mechanism of the $CH1_1$ peptide, they do show that $CH1_1$ folding is not a simple two-state cooperative process. This is in marked contrast to results obtained for other Zn^{2+}-binding peptides,[35,52] which only show pseudo-first-order stopped-flow kinetics even under different experimental conditions (e.g., different zinc concentrations and pH values). In addition, concentration-dependent measurements show that the rate constant of the initial or fastest kinetic phase (i.e., k_{1obs} in Table 1) is linearly proportional to Zn^{2+} concentration (Fig. 4b), typical of a bimolecular reaction taking place under pseudo-first-order reaction conditions. Therefore, we attribute this kinetic phase to the initial binding of Zn^{2+} to the peptide. Moreover, the second-order Zn^{2+}-binding rate constant for L7W and T21W is determined to be about 1.8×10^6 M^{-1} s^{-1}, which is comparable to that observed for other zinc-binding proteins.[52–55] To further verify this assignment, additional stopped-

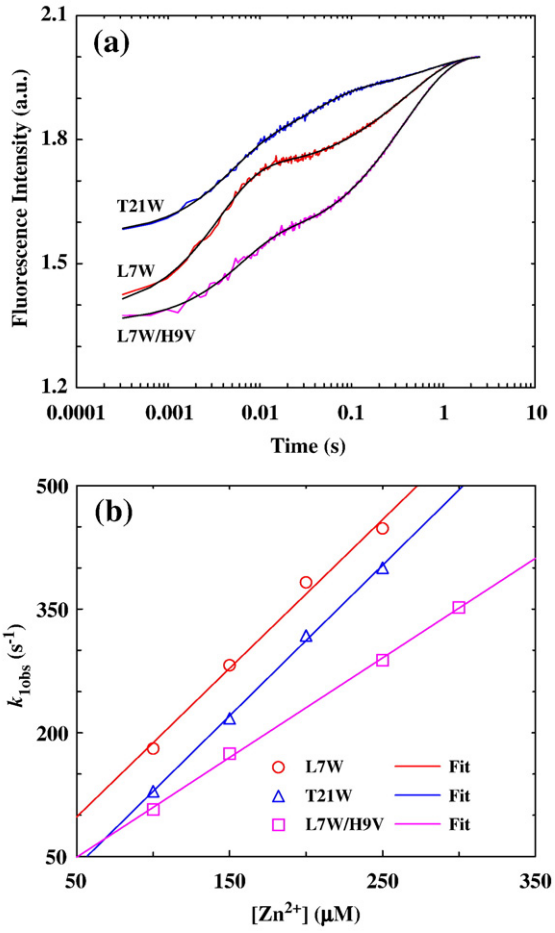


Fig. 4. (a) Representative stopped-flow kinetics of L7W, T21W, and L7W/H9V (these data have been offset for easy comparison) at 15 °C and pH 6.5, which were initiated by mixing equal volumes of a peptide solution and a $ZnCl_2$ solution. The final peptide and Zn^{2+} concentrations were 10 μM and 0.15 mM, respectively. Smooth lines represent the best fits to the following triexponential function: $F(t) = B - C[A_1\%\exp(-k_{1obs}t) + A_2\%\exp(-k_{2obs}t) + A_3\%\exp(-k_{3obs}t)]$, and the resultant best-fit parameters are listed in Table 1. (b) The observed rate constant of the first or fastest kinetic phase (i.e., k_{1obs}) *versus* Zn^{2+} concentration. Linear regressions to these data yield those straight lines with the following slopes: 1.8×10^6 M^{-1} s^{-1} (L7W), 1.8×10^6 M^{-1} s^{-1} (T21W), and 1.2×10^6 M^{-1} s^{-1} (L7W/H9V).

Table 1. Phenomenological rate constant (k_{iobs}) and relative amplitude (A_i%) of each kinetics phase obtained by fitting the stopped-flow kinetics presented in Figs. 4a and 5 to a triexponential function

	k_{1obs}	k_{2obs}	k_{3obs}	A_1%	A_2%	A_3%
L7W (pH 6.5)	282±29	11±1	2.0±0.1	56±5	10±2	34±2
T21 W (pH 6.5)	217±10	25±7	1.4±0.1	49±6	31±4	20±2
L7 W/H9V (pH 6.5)	174±10	8±1	2.2±0.6	31±1	11±2	58±3
T21W (pH 7.5)	956±50	39±3	1.3±0.2	54±6	30±4	16±2

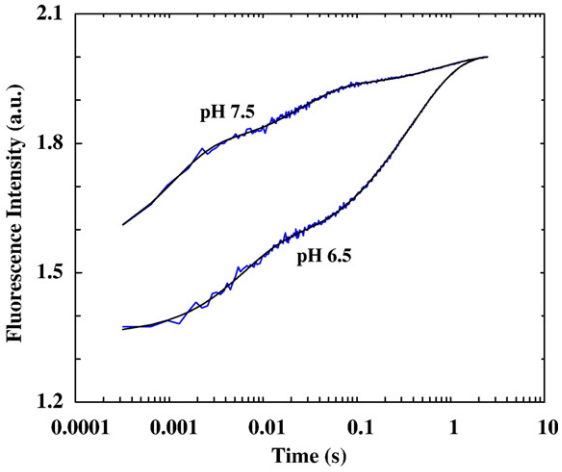


Fig. 5. Stopped-flow kinetics of T21W obtained at pH 6.5 and pH 7.5, as indicated. These data have been offset for easy comparison.

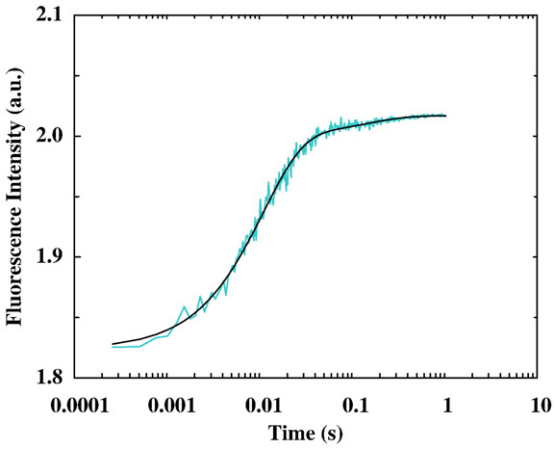


Fig. 6. Stopped-flow kinetics obtained by mixing equal volumes of a L7W/H9V/H22I peptide solution and a $ZnCl_2$ solution at 15 °C. The final peptide and Zn^{2+} concentrations were 10 μM and 0.15 mM, respectively. The smooth line is the best fit of these data to a biexponential function, and the resultant rate constants are $k_{1obs}=92\ s^{-1}$ and $k_{2obs}=6.8\ s^{-1}$.

flow mixing experiments were carried out at different pH conditions. As expected (Fig. 5 and Table 1), taking T21W as an example, the rate of the first stopped-flow kinetic phase increases approximately four times when the pH of the final mixed solution is increased from 6.5 to 7.5, due to an increase in the population of deprotonated ligands (especially deprotonated His). Interestingly, the Zn^{2+}-induced folding kinetics of L7W/H9V/H22I can be well described by a biexponential function (Fig. 6), suggesting that replacement of H9 and H22 with nonligating amino acids simplifies the folding process.

Finally, temperature-dependent measurements were carried out in the temperature range of 5–30 °C to further characterize the apparent activation energy (E_a) encountered by each kinetic phase. As shown (Fig. 7 and Table 2), all the rate constants show Arrhenius temperature dependence, with E_a ranging from 6 to 14 kcal mol^{-1}.

Table 2. Apparent activation energy of each stopped-flow kinetic phase

	L7W	T21W	L7W/H9V
E_a^1 (kcal mol^{-1})	6.2±0.5	7.1±1.7	8.4±1.5
E_a^2 (kcal mol^{-1})	14.1±1.2	10.2±0.5	12.8±0.3
E_a^3 (kcal mol^{-1})	9.3±0.6	12.9±0.3	12.0±1.5

Discussion

The $CH1_1$ peptide is a novel zinc-binding motif found in the first Cys/His-rich region of CREB-binding protein.[40] It is known that this peptide binds to zinc through a consensus CCHC motif (i.e., Cys-X_4-Cys-X_8-His-X_3-Cys, wherein Cys and His residues are the ligands for Zn^{2+}) and forms a thermostable well-defined tertiary structure that is extremely tolerant to sequence mutations at nonligating positions.[40,41] In other words, Zn^{2+} binding provides a major driving force for $CH1_1$ folding. Therefore, this peptide provides an ideal model system for understanding the mechanism of metal-induced protein folding processes.

Since the native sequence of $CH1_1$ peptide does not contain a suitable fluorophore (e.g., Trp) that can be used to follow the kinetic process of binding or folding using fluorescence-based techniques, we have employed a statistical computational design method[43,44] to identify structurally consistent Trp-containing mutants of $CH1_1$. First, two single mutants, L7W and T21W, were identified as the best targets because, in both cases, the Trp residue is sandwiched between two zinc ligands (i.e., C5/C10 for L7W and H19/C23 for T21W). Thus, their fluorescence is likely sensitive to Zn^{2+} binding. Second, a double mutant, L7W/H9V, and a triple mutant, L7W/H9V/H22I, were designed to investigate the role of nonligating His residues (i.e., His9

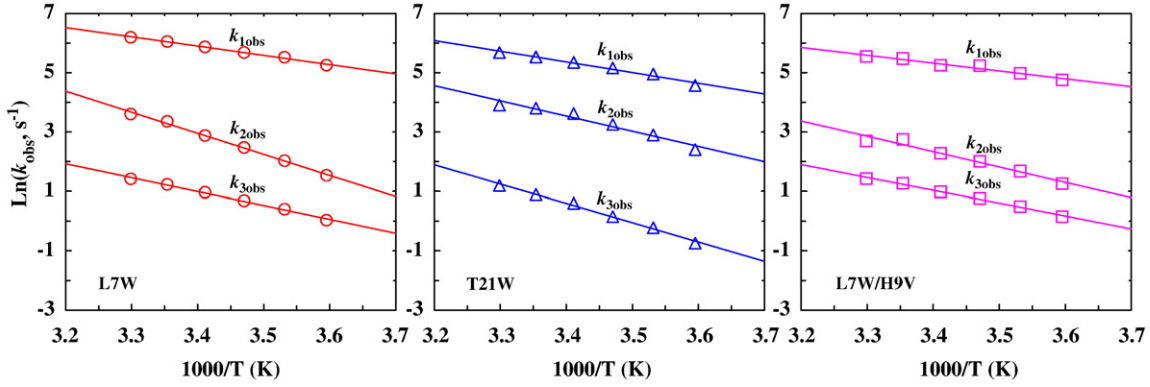


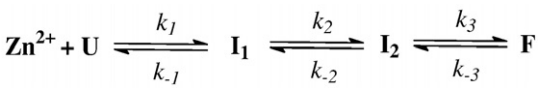
Fig. 7. Arrhenius plots of the rate constants obtained by fitting the stopped-flow kinetics of the $CH1_1$ mutants (as indicated) to the triexponential function discussed in Fig. 4. Lines are fits to the Arrhenius equation: $\ln(k_{iobs})=\ln(A)-E_a^i/RT$, where E_a^i is the apparent activation energy for kinetic phase i (i=1, 2, or 3), and its value is listed in Table 2.

and His22) in the kinetics of $CH1_1$ folding. Indeed, both equilibrium and stopped-flow fluorescence measurements show that, upon coordination with Zn^{2+}, the Trp fluorescence of these mutants increases significantly, indicative of the sensitivity of this fluorescent probe to metal-induced folding processes. Furthermore, the structural integrity of these mutants was verified by low-resolution methods (i.e., CD and one-dimensional NMR).

A simple sequential model is inadequate to describe the folding mechanism of the $CH1_1$ peptide

Compared to other metal-binding motifs,[34,35,54–56] these $CH1_1$ mutants exhibit rather complex and slow cofactor-induced folding kinetics. As shown (Fig. 4a), the stopped-flow kinetics of L7W consist of (at least) three exponential phases, indicative of a non-two-state folding scenario. While this result is not surprising, it is nevertheless interesting as many proteins of similar size that lack cofactors have been shown to fold in a two-state manner[37,57–59] and on a much faster timescale.[37,60–62] Therefore, the slower and more complex folding kinetics of L7W must arise from the fact that the folding of $CH1_1$ necessitates Zn^{2+} ligation in a tetrahedral geometry.

The fastest phase in the stopped-flow kinetics of L7W is characteristic of a bimolecular reaction, as its rate constant (i.e., k_{1obs} in Table 1) depends linearly on Zn^{2+} concentration (Fig. 4b). Therefore, we attribute this kinetic phase to the initial association of the peptide with zinc, wherein a Zn^{2+}–peptide complex ($\mathbf{I_1}$) is formed. Given the fact that the four native ligands in $CH1_1$ are segregated into two pairs—one (C5 and C10) close to the N-terminus and the other (C23 and H19) close to the C-terminus—it is likely that only two ligands in $\mathbf{I_1}$ are ligated. Indeed, Bombarda *et al.* have recently shown that the first step in the Zn^{2+}-induced folding kinetics of the distal CCHC finger motif of the human immunodeficiency virus-1 nucleocapsid protein corresponds to the formation of a bidentate state wherein only C36 and H44 are coordinated with the cofactor.[35] Therefore, the simplest interpretation of the triexponential stopped-flow fluorescence kinetics of L7W is that this initially formed Zn^{2+}–peptide complex $\mathbf{I_1}$ further evolves to produce the final folded state ($\mathbf{F}$) via a sequential mechanism involving an on-pathway folding intermediate ($\mathbf{I_2}$), as indicated in Scheme 1. In fact, zinc coordination has been proposed to proceed in an intrinsically multistep and sequential manner for several peptides.[35,54,63] While Scheme 1 can indeed fit the stopped-flow kinetics of L7W obtained at different Zn^{2+} concentrations, the microscopic rate



Scheme 1. A sequential folding mechanism involving an on-pathway folding intermediate.

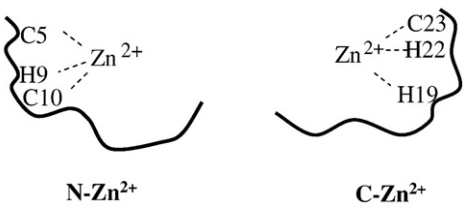
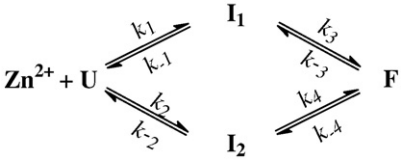


Fig. 8. Two putative Zn^{2+}–peptide complexes formed after the initial binding of zinc, where H9 and H22 are the two nonnative ligands that may also participate in binding.

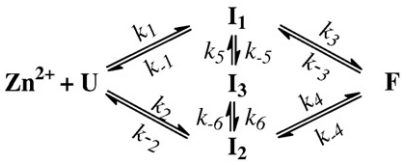
constants recovered from the best fit (Supporting Information, Table S1), however, give rise to an equilibrium constant that is not overwhelmingly favorable for the folded state. Instead, this model indicates that the folding intermediate states are significantly populated at equilibrium (e.g., $[\mathbf{U}]:[\mathbf{I_1}]:[\mathbf{I_2}]:[\mathbf{F}] = 1:9:9:14$ at 300 μM Zn^{2+}), inconsistent with previous[40] and current equilibrium results (Fig. 3b). This inconsistency suggests that a simple sequential model, such as Scheme 1, is inadequate to describe the Zn^{2+}-induced folding mechanism of the $CH1_1$ peptide.

$CH1_1$ folding involves parallel pathways

The inadequacy of Scheme 1 in describing the folding mechanism of the $CH1_1$ peptide prompted us to design and study a second mutant (i.e., T21W) wherein the Trp is placed near the second group of the Zn^{2+} ligands H19 and C23. The rationale is to explore whether Zn^{2+}-induced $CH1_1$ folding involves parallel pathways. In principle, the initial binding of Zn^{2+} to the peptide could occur at both N-terminal and C-terminal regions, leading to the formation of N-Zn^{2+} and C-Zn^{2+} (Fig. 8). Since Trp fluorescence quenching by various amino acids is strongly distance dependent, we expect both L7W and T21W to exhibit similar fluorescence intensity changes upon the initial binding of Zn^{2+} if both N-Zn^{2+} and C-Zn^{2+} are formed. Indeed, the relative amplitude of the first stopped-flow kinetic phase of T21W (Table 1), arising from the initial Zn^{2+}–peptide association, is 49%, similar to that (i.e., 56%) of L7W. Taken together, these results thus suggest that the Zn^{2+}-induced $CH1_1$ folding involves at least two parallel pathways, as shown in Scheme 2, which represents the simplest parallel folding mechanism wherein N-Zn^{2+} ($\mathbf{I_1}$) and C-Zn^{2+} ($\mathbf{I_2}$) are directly converted into the folded state. Similar



Scheme 2. A parallel folding mechanism involving two on-pathway folding intermediates.



Scheme 3. A parallel folding mechanism involving two on-pathway folding intermediates and a misligated state ($\mathbf{I_3}$).

to Scheme 1, this folding mechanism possesses a sufficient number of (fitting) parameters that can be adjusted to fit the observed stopped-flow kinetics of both L7W and T21W. However, the microscopic rate constants recovered from the best fits (Supporting Information, Table S2) suggest that this folding scheme is also unlikely to be correct. For example, the values of both k_3 and k_{-3} for T21W nearly approach 0, suggesting that T21W can only fold to its native state via $\mathbf{I_2}$, whereas in the case of L7W, k_3 is comparable to k_4. While it is possible that the folding mechanism of T21W is different from that of L7W, it seems more plausible that a more complicated kinetic scheme is required.

$CH1_1$ folding involves a misligated intermediate

Since the $CH1_1$ peptide contains two 'extra' zinc ligands (i.e., H9 and H22), it is therefore possible that its complex folding kinetics upon association with zinc arise from the formation of a misligated state involving either H9 or H22, or both. To test this possibility, we have further studied a double mutant, L7W/H9V, and a triple mutant, L7W/H9V/H22I. If misligation involving both H9 and H22 indeed occurs during $CH1_1$ folding, we would expect that only the stopped-flow kinetics of the triple mutant, wherein all nonnative zinc ligands are replaced by nonligating amino acids, will become simpler. Indeed, the Zn^{2+}-induced folding kinetics of L7W/H9V remain triexponential (Fig. 4a), whereas those of L7W/H9V/H22I can be well described by a biexponential function (Fig. 6). Taken together, these results hence support the idea that the Zn^{2+}-induced $CH1_1$ folding process involves misligation of Zn^{2+} to H9 and/or H22.

While other mechanisms are possible, Scheme 3 represents the simplest scenario wherein such a misligated state (i.e., $\mathbf{I_3}$) is populated. As shown (Supporting Information, Fig. S3), Scheme 3 not only adequately fits the stopped-flow kinetics of L7W and T21W obtained under different conditions but also yields physically meaningful fitting parameters (Table 3). For example, the values of k_1 and k_2 are comparable for both mutants, consistent with the aforementioned picture that the initial binding of Zn^{2+} to the peptide proceeds in two parallel pathways wherein N-Zn^{2+} and C-Zn^{2+} are formed. In addition, the overall bimolecular rate constant for zinc binding is determined to be $\sim 1.5\times10^6$ M^{-1} s^{-1} (i.e., the sum of k_1 and k_2 in Table 3), similar to that determined according to the Eigen–Wilkins mechanism (i.e., 5×10^6 M^{-1} s^{-1} for Zn^{2+} ligation).[54] Furthermore, both k_5 and k_6 are larger than k_{-5} and k_{-6}, indicating that $\mathbf{I_3}$ is only transiently populated (same as $\mathbf{I_1}$ and $\mathbf{I_2}$). More importantly, the overall equilibrium constant for each peptide [i.e., K_d (L7W) = 20.8 μM, K_d (T21W) = 3.2 μM, and K_d (L7W/H9V) = 30.3 μM], calculated based on the recovered microscopic rate constants, is comparable to that determined from the equilibrium titration experiment (Fig. 3b), which assumes a simpler binding model.

The relative fluorescence intensity determined from fitting for every species provides a more stringent test of the current model. Since the major Trp fluorescence quenchers in both L7W and T21W are Cys and His residues, it is possible to rationalize, albeit in a qualitative manner, the relative fluorescence intensities of the folded, unfolded, and intermediate states involved in a specific folding mechanism. For Scheme 3, it is expected that the fluorescence intensity of **F** is higher than those of **U**, $\mathbf{I_1}$, and $\mathbf{I_2}$

Table 3. Microscopic rate constants (k_i) and relative fluorescence intensities (Q) derived by globally fitting the stopped-flow kinetics obtained at different Zn^{2+} concentrations for each mutant to Scheme 3 or Scheme 2 (for L7W/H9V/H22I only)

	L7W	T21W	L7W/H9V	L7W/H9V/H22I
k_1 (M^{-1} s^{-1})	(5.8±0.5)×10^5	(6.2±1.2)×10^5	(2.0±0.4)×10^5	(2.4±0.6)×10^5
k_{-1} (s^{-1})	76±10	49±21	77±4	5±3
k_2 (M^{-1} s^{-1})	(8.6±0.4)×10^5	(8.5±1.6)×10^5	(6.9±0.7)×10^5	(3.5±0.2)×10^5
k_{-2} (s^{-1})	56±2	12±1	156±10	0.2±0.1
k_3 (s^{-1})	2.6±0.8	3.0±0.9	7.9±3	3.9±0.5
k_{-3} (s^{-1})	0.1±0.1	(3±1)×10^{-4}	0.6±0.2	1.8±0.4
k_4 (s^{-1})	1.4±0.1	1.3±0.4	0.1±0.1	2.1±0.9
k_{-4} (s^{-1})	0.7±0.3	0.3±0.1	0.02±0.01	1.2±0.5
k_5 (s^{-1})	22±8	9±2	7±0.3	—
k_{-5} (s^{-1})	0.2±0.1	7±3	0.5±0.3	—
k_6 (s^{-1})	4±2	7±2	~0	—
k_{-6} (s^{-1})	0.4±0.2	5±1	0.5±0.05	—
Q(**U**)	8.2±0.6	11.6±0.5	12±3	9.9±0.5
Q($\mathbf{I_1}$)	15±3	13±1	19±3	12.0±0.2
Q($\mathbf{I_2}$)	11±2	14.9±0.5	14±2	11.6±0.2
Q($\mathbf{I_3}$)	80±20	18±4	49±5	—
Q(**F**)	16±1	16±1	20±3	12.1±0.2

because more Trp fluorescence quenchers are bound to Zn^{2+} in the folded state (**F**). Similarly, if $\mathbf{I_1}$ and $\mathbf{I_2}$ represent N-Zn^{2+} and C-Zn^{2+} in Scheme 3, respectively, one would expect that, for L7W (T21W), the fluorescence intensity of $\mathbf{I_1}$ ($\mathbf{I_2}$) is larger than that of $\mathbf{I_2}$ ($\mathbf{I_1}$) because, in this case, the Trp fluorophore is located near the N-terminal (C-terminal) ligands. As indicated (Table 3), the relative fluorescence intensities of **F**, $\mathbf{I_1}$, and $\mathbf{I_2}$ recovered from the best fits of the stopped-flow data of L7W and T21W to Scheme 3 indeed meet these expectations. Moreover, for both mutants, the off-pathway intermediate $\mathbf{I_3}$ exhibits the largest relative fluorescence intensity, suggesting that it adopts a conformation wherein Trp fluorescence quenching by all potential quenchers (i.e., all Cys and His residues) is maximally reduced. Thus, this result corroborates the idea that the misligated state involves H9 and/or H22.

While the structure of the misligated state cannot be exclusively determined from the current study, a possible Zn^{2+} coordination is presented in Fig. 9, where Zn^{2+} not only coordinates with the native ligands but also is associated with one or both of the nonnative ligands (i.e., H9 and H22). Another possibility is that both H9 and H22 are in rapid exchange with one of the native ligands.[55,64] In addition, the more pronounced increase in the relative fluorescence intensity of $\mathbf{I_3}$ (compared to **F**) for L7W (and L7W/H9V) than that for T21W (Table 3) likely arises from a difference in the orientations of the indole rings of Trp7 and Trp21. As shown in the model structures provided by design calculations (Fig. 2), the side chain of Trp7 is oriented towards the zinc and is thus expected to be more sensitive than Trp21 to changes in the coordination number of Zn^{2+}.

Moreover, fitting the stopped-flow kinetics of L7W/H9V to Scheme 3 and those of L7W/H9V/H22I to Scheme 2 (where $\mathbf{I_3}$ is no longer needed because H9 and H22 are deleted) also yields a set of microscopic rate constants and relative fluorescence intensities (Table 3) that meet the aforementioned structural characteristics of $\mathbf{I_1}$, $\mathbf{I_2}$, and $\mathbf{I_3}$. For example, k_1 of L7W/H9V is smaller than that of L7W as a result of H9 deletion, and both k_1 and k_2 of L7W/H9V/H22I are decreased as compared to those of L7W as a result of H9 and H22 deletion. In addition, the resultant microscopic rate constants for L7W/H9V/H22I (i.e., k_3, k_{-3}, k_4, and k_{-4}) indicate that deleting nonnative ligands results in a less stable $CH1_1$ fold. This is consistent with the fact that both H9 and H22 are conserved in the $CH1_1$ sequence[40] to some extent and is also similar to that

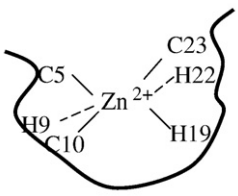


Fig. 9. A possible zinc coordination in the misligated state.

observed in a recent study[47] on the NZF-1 zinc-binding domain, which showed that an 'extra' histidine not only is important for maintaining the tertiary structure of the peptide but also can coordinate with the Zn^{2+} when the native histidine is absent.

Taken together, our results show that Scheme 3 is capable of providing a comprehensive description of the zinc-induced folding kinetics of four computationally designed mutants of the $CH1_1$ peptide, as monitored by Trp fluorescence. While earlier studies have shown that binding of Zn^{2+} to a peptide could proceed in a multistep and sequential manner[35,54,65] and that misligation of nearby nonnative ligands to Zn^{2+} could occur in the later stage of the folding process,[65] the current model highlights the importance of parallel binding-coupled folding pathways, as well as the nonnative ligands, for achieving a higher metal binding affinity. Furthermore, the finding that the two nonnative ligands (i.e., H9 and H22) stabilize the $CH1_1$ fold even though they complicate the folding kinetics provides strong evidence suggesting that, in certain cases, functional requirements for folding could produce significant local frustrations in the folding energy landscape, thus slowing folding. Similar conclusions regarding the divergent requirements for folding and function have also been reached in other studies.[66–69]

Estimating the loop closure rate

Temperature-dependent measurements indicate that all three stopped-flow kinetic phases exhibit Arrhenius temperature dependence, with an apparent activation energy ranging from 6 to 14 kcal mol^{-1} (Table 2), depending on the kinetic phase and mutant. While it is difficult to rationalize these apparent activation energies and also the subtle differences among them, they likely arise from one of the following events:[70] (a) dehydration of Zn $(H_2O)_6^{2+}$; (b) deprotonation of His and Cys ligands; (c) peptide chain motions required to bring the corresponding ligands to a suitable position for coordination; and (e) disruption of misligation for correct folding. Regardless of the origin of those energetic barriers, however, the value of E_a^3 obtained for each $CH1_1$ mutant could be used to estimate the loop closure rate for short peptides. According to Scheme 3, in order to form the final folded and closed conformation of $CH1_1$ from either N-Zn^{2+} ($\mathbf{I_1}$) or C-Zn^{2+} ($\mathbf{I_2}$), the opposite end of the peptide chain must be brought to a suitable distance to the peptide-bound Zn^{2+}, so that further coordination can take place. Therefore, using the apparent rate constant and the activation energy obtained for the third kinetic step, which approximately measures the total rate leading to the formation of the native and closed conformation of $CH1_1$ from both N-Zn^{2+} and C-Zn^{2+}, we can estimate the free diffusion rate of the $CH1_1$ peptide chain in aqueous solution. In other words, the prefactor in the Arrhenius equation $k_{3obs} = A\exp(-E_a^3/RT)$ provides an upper limit for the rate of the peptide chain diffusion. At 15 °C, this

rate was estimated to be 2.1×10^{7} s^{-1} for L7W, 6.6×10^{6} s^{-1} for L7W/H9V, and 7.0×10^{9} s^{-1} for T21W. Despite its simplicity, this simple analysis yields peptide chain diffusion rates that are in good agreement with those determined by measuring the end-to-end collision rate in short peptides (e.g., 7.1×10^{6} to 1.5×10^{8} s^{-1}),[71,72] providing further support for the proposed zinc-induced folding mechanism of $CH1_1$.

Conclusions

In summary, the Zn^{2+}-induced $CH1_1$ folding can be described by a parallel pathway wherein Zn^{2+} first binds to either the N-terminal group of ligands (C5/C10/H9) or the C-terminal group of ligands (H19/C23/H22). This initial binding event is followed by a rearrangement of the peptide chain, leading to the closure of the peptide loop and the concomitant formation of the native zinc tetrahedral coordination. During folding, however, a misligated Zn^{2+}–peptide complex, which may have a higher coordination number, is also transiently populated. While the existence of nonnative Zn^{2+} ligands is responsible for the formation of this off-pathway folding intermediate, these 'extra' ligands nevertheless facilitate the binding of Zn^{2+} to $CH1_1$ peptide and also enhance the stability of the final fold. In addition, the loop closure rate of this peptide is estimated to be in the range of 6.6×10^{6} to 7.0×10^{9} s^{-1}.

Materials and Methods

Materials and sample preparation

All peptides were synthesized with a standard Fmoc-based solid-phase method using a PS3 peptide synthesizer (Protein Technologies, Boston, MA) and purified by reverse-phase high-performance liquid chromatography. The identity of each peptide was further verified by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. All peptide solutions were prepared by directly dissolving lyophilized peptide solid in 10 mM deoxygenated Tris buffer (pH 6.5 or pH 7.5, depending on the experiment). To prevent cysteine oxidation, each peptide solution also contains a 10-fold excess of Tris(2-carboxyethyl) phosphine hydrochloride (Pierce Protein Research Products, Part of Thermo Fisher Scientific, Inc., Rockford, IL). The final peptide concentration of each solution was determined optically using Trp absorbance at 280 nm and a molar extinction coefficient of 5600 M^{-1} cm^{-1}. The Zn^{2+} solution was freshly prepared before each experiment by dissolving $ZnCl_2$ (Sigma, St. Louis, MO) in 10 mM deoxygenated Tris buffer (pH 6.5 or pH 7.5).

Equilibrium CD measurement

Far-UV CD spectra at 25 °C were collected with an AVIV 62DS spectropolarimeter (Aviv Associates, Lakewood, NJ) using a 1-mm quartz cell. For each case, the concentrations of the peptide and Zn^{2+} were ~30 μM and 1 mM, respectively.

Equilibrium fluorescence measurement

Trp fluorescence spectra were collected under a nitrogen atmosphere on a Fluorolog 3.10 spectrofluorometer (Jobin Yvon Horiba, Edison, NJ) using a 1-cm quartz sample holder. The excitation wavelength was 285 nm, and the spectral resolution was 2 nm (for both excitation and emission).

Equilibrium binding experiment

The Zn^{2+}-binding affinity of the $CH1_1$ mutants at 15 °C was determined by Trp fluorescence titration. The Zn^{2+}–peptide solution (10 μM peptide) for each measurement was prepared by adding an appropriate aliquot of a 4 mM $ZnCl_2$ stock solution to a peptide solution, which was allowed to equilibrate for 5 min before fluorescence measurement. The total fluorescence of each Zn^{2+}–peptide solution was taken as the sum of those of the free and Zn^{2+}-bound peptides, that is,

$$F_{\text{int}} = F_{\text{b}} \cdot [\text{P}]_{\text{b}} + F_{\text{f}} \cdot [\text{P}]_{\text{f}} \tag{1}$$

where F_{int} is the integrated area of the fluorescence spectrum; F_{f} and F_{b} represent the molar fluorescence intensities of the free and Zn^{2+}-bound peptides, respectively; and $[\text{P}]_{\text{f}}$ and $[\text{P}]_{\text{b}}$ denote the molar concentrations of the free and Zn^{2+}-bound peptides, respectively. For 1:1 binding, $[\text{P}]_{\text{f}}$ and $[\text{P}]_{\text{b}}$ follow the following relationship:

$$K_{\text{d}} = \frac{[\text{Zn}^{2+}][\text{P}]_{\text{f}}}{[\text{P}]_{\text{b}}} \tag{2}$$

where K_{d} is the dissociation constant of the Zn^{2+}–peptide complex. Since Zn^{2+} also binds to Tris, the Zn^{2+} concentration ($[Zn^{2+}]$) in the above equation should also satisfy the following relationship:

$$K_{\text{d}}' = \frac{[\text{Zn}^{2+}][\text{Tris}]_{\text{f}}}{[\text{Tris}]_{\text{b}}} \tag{3}$$

where K_{d}' is the dissociation constant of the Zn^{2+}–Tris complex (K_{d}' = 11.5 mM),[73] and $[\text{Tris}]_{\text{f}}$ and $[\text{Tris}]_{\text{b}}$ are the molar concentrations of the free and Zn^{2+}-bound Tris molecules, respectively. Since K_{d}' is relatively large and the total concentration of Tris (i.e., $[\text{Tris}]_{\text{T}} = [\text{Tris}]_{\text{f}} + [\text{Tris}]_{\text{b}}$ = 10 mM) is much larger than that of $[Zn^{2+}]$ (in the micromolar range), we have assumed that $[\text{Tris}]_{\text{f}} = [\text{Tris}]_{\text{T}}$ in the derivation of the following equation:

$$[\text{P}]_{\text{b}} = \left(\frac{([\text{Zn}^{2+}]_{\text{T}} + [\text{P}]_{\text{T}} + K_{\text{d}} + [\text{Tris}]_{\text{T}} K_{\text{d}}/K_{\text{d}}') - \sqrt{([\text{Zn}^{2+}]_{\text{T}} + [\text{P}]_{\text{T}} + K_{\text{d}} + [\text{Tris}]_{\text{T}} K_{\text{d}}/K_{\text{d}}')^2 - 4[\text{Zn}^{2+}]_{\text{T}}[\text{P}]_{\text{T}}}}{2} \right) \tag{4}$$

where $[\text{P}]_{\text{T}}$ is the total concentration of the peptide. For each experiment, the total concentrations of the peptide, Zn^{2+}, and Tris were known. Thus, Eqs. (1) and (4) allow one to determine the dissociation constant of the Zn^{2+}–peptide complex by fitting the equilibrium fluorescence binding curve (i.e., Fig. 3b) by varying K_{d}, F_{b}, and F_{f}.

Stopped-flow measurement

The zinc-induced $CH1_1$ folding kinetics were monitored via Trp fluorescence using an SFM-300 stopped-flow module (Bio-logic, Claix, France) equipped with home-

built optics, the detail of which has been described elsewhere.[51] In the current study, a microcuvette (μFC-08) with an optical pathlength of 0.8 mm was used, and the dead time of the system was ~600 μs, determined by *N*-acetyl-tryptophanamide fluorescence quenching by *N*-bromo-succinimide. The stopped-flow kinetics were initiated by mixing equal volumes of a peptide solution and a $ZnCl_2$ solution. The final peptide concentration was 10 μM, and the final Zn^{2+} concentration was 0.1–0.3 mM, depending on the experiment. The Trp fluorophore was excited at 285 nm, and the resultant fluorescence (for $\lambda_{em}>315$ nm) was measured by a photomultiplier tube. Sample temperature (5–30 °C) was controlled by a Thermo Neslab RTE-7 circulator (Thermo, Newington, NH). The stopped-flow traces shown in the figures correspond to an average of 6–10 shots. The stopped-flow kinetics thus obtained were fitted to either a multiexponential function using a nonlinear least squares method or a kinetic model (i.e., Scheme 1, Scheme 2, or Scheme 3). For the latter case, all individual kinetic steps were assumed to be of the first order with respect to each species, and the corresponding rate equations were solved numerically using the program Scientist (Micromath Research LC) to yield the best fit.

Mutant design

Four structurally consistent mutants of $CH1_1$ were identified using SCADS, the details of which have been described elsewhere.[30,43,44,74] Briefly, the site-specific probabilities of amino acids at variable positions in a given structure are calculated, allowing identification of structurally consistent mutations. These probabilities are estimated by optimizing an effective sequence conformation entropy subject to constraints involving the total average energy of the structure, as measured using an atom-based energy function. The conformational energy is represented as an average over sequence and side-chain conformational (rotamer) states of the amino acids,[30,43,44] where all interatomic interactions are quantified using the AMBER force field[75] with a modified hydrogen-bonding energy[76] and a distance-dependent electrostatic dielectric constant ($\varepsilon=4r_{ij}$, where r_{ij} is the distance between atoms *i* and *j*). The nonbonding parameters for Zn^{2+} were obtained from Hoops *et al.* (Lennard–Jones parameters: $R*=1.1$ Å, $\varepsilon=0.012$ kcal mol^{-1}; partial charge: $q_{Zn^{2+}}=0.866$).[77] The protein backbone coordinates were constrained to those in the experimentally determined NMR structure (Protein Data Bank code 1liq).[40] Side-chain variability was addressed using a backbone-dependent rotamer library with a maximum of 81 rotamers per amino acid.[78]

The sequence design was divided into two components. The first involved identifying sites for introducing a Trp as an intrinsic fluorescent probe. The residue positions in the protein were scanned for possible mutations, allowing Trp, Tyr, and wild-type amino acid while fixing the remaining residues at their wild-type amino acids. For all calculations, the four metal-binding residues (Cys5, Cys10, His19, and Cys23) were fixed to their wild-type conformations, while other residues were allowed to take on rotamer states contained in the library. The second part of the design identified suitable mutations for the two nonligating histidines (His9 and His22) to inform the roles of these nonligating histidines in folding. All 19 natural amino acids, except Cys, were allowed for the site-directed mutagenesis of these two sites. The amino acid probabilities were obtained at medium ($T_{eff}=2.0$ kcal mol^{-1}) and low ($T_{eff}=0.5$ kcal mol^{-1}) effective temperatures. T_{eff} represents an effective temperature conjugate to the configurational energy of the sequences; lower temperatures correspond to a lower average energy of the sequence when in the target structure.[30,44] In addition, mutations between sites 9 and 22 were not correlated with each other, presumably due to the relatively large distance between them ($d_{C_\beta - C\beta}=14.1$ Å).

Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978 and GM-61267) and the National Science Foundation (DMR05-20020). Support for computational resources was provided, in part, by the National Science Foundation (CHE-0131132).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2009.03.074

References

1. Ferguson, N. & Fersht, A. R. (2003). Early events in protein folding. *Curr. Opin. Struct. Biol.* **13**, 75–81.
2. Fersht, A. R. & Sato, S. (2004). Ψ-value analysis and the nature of protein-folding transition states. *Proc. Natl Acad. Sci. USA*, **101**, 7976–7981.
3. Krishna, M. M., Lin, Y. & Englander, S. W. (2004). Protein misfolding: optional barriers, misfolded intermediates, and pathway heterogeneity. *J. Mol. Biol.* **343**, 1009–1095.
4. Oliveberg, M. & Wolynes, P. G. (2005). The experimental survey of protein-folding energy landscapes. *Q. Rev. Biophys.* **38**, 245–288.
5. Deechongkit, S., Nguyen, H., Jager, M., Powers, E. T., Gruebele, M. & Kelly, J. W. (2006). β-Sheet folding mechanisms from perturbation energetics. *Curr. Opin. Struct. Biol.* **16**, 94–101.
6. Naganathan, A. N., Doshi, U., Fung, A., Sadqi, M. & Munoz, V. (2006). Dynamics, energetics, and structure in protein folding. *Biochemistry*, **45**, 8466–8475.
7. Petrovich, M., Jonsson, A. L., Ferguson, N., Daggett, V. & Fersht, A. R. (2006). Ψ-Analysis at the experimental limits: mechanism of beta-hairpin formation. *J. Mol. Biol.* **360**, 865–881.
8. Sosnick, T. R., Krantz, B. A., Dothager, R. S. & Baxa, M. (2006). Characterizing the protein folding transition state using Φ analysis. *Chem. Rev.* **106**, 1862–1876.
9. Beck, D. A., White, G. W. & Daggett, V. (2007). Exploring the energy landscape of protein folding using replica-exchange and conventional molecular dynamics simulations. *J. Struct. Biol.* **157**, 514–523.
10. Dyer, R. B. (2007). Ultrafast and downhill protein folding. *Curr. Opin. Struct. Biol.* **17**, 38–47.
11. Gray, H. B. (2003). Biological inorganic chemistry at the beginning of the 21st century. *Proc. Natl Acad. Sci. USA*, **100**, 3563–3568.
12. Wilson, C. J., Apiyo, D. & Wittung-Stafshede, P. (2004). Role of cofactors in metalloprotein folding. *Q. Rev. Biophys.* **37**, 231–285.

13. Frankel, A. D., Berg, J. M. & Pabo, C. O. (1987). Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proc. Natl Acad. Sci. USA*, **84**, 4841–4845.
14. Krizek, B. A., Amann, B. T., Kilfoil, V. J., Merkle, D. L. & Berg, J. M. (1991). A consensus zinc finger peptide: design, high-affinity metal-binding, a pH-dependent structure, and a His to Cys sequence variant. *J. Am. Chem. Soc.* **113**, 4518–4523.
15. Zong, C. H., Wilson, C. J., Shen, T. Y., Wittung-Stafshede, P., Mayo, S. L. & Wolynes, P. G. (2007). Establishing the entatic state in folding metallated *Pseudomonas aeruginosa* azurin. *Proc. Natl Acad. Sci. USA*, **104**, 3159–3164.
16. Jones, C. M., Henry, E. R., Hu, Y., Chan, C. K., Luck, S. D., Bhuyan, A. *et al.* (1993). Fast events in protein-folding initiated by nanosecond laser photolysis. *Proc. Natl Acad. Sci. USA*, **90**, 11860–11864.
17. Shastry, M. C. R., Sauder, J. M. & Roder, H. (1998). Kinetic and structural analysis of submillisecond folding events in cytochrome *c*. *Acc. Chem. Res.* **31**, 717–725.
18. Englander, S. W., Sosnick, T. R., Mayne, L. C., Shtilerman, M., Qi, P. X. & Bai, Y. W. (1998). Fast and slow folding in cytochrome *c*. *Acc. Chem. Res.* **31**, 737–744.
19. Guidry, J. & Wittung-Stafshede, P. (2000). Cytochrome *c*(553), a small heme protein that lacks misligation in its unfolded state, folds with rapid two-state kinetics. *J. Mol. Biol.* **301**, 769–773.
20. Chen, E. F., Goldbeck, R. A. & Kliger, D. S. (2004). The earliest events in protein folding: a structural requirement for ultrafast folding in cytochrome *c*. *J. Am. Chem. Soc.* **126**, 11175–11181.
21. Bertini, I., Gray, H. B., Lippard, S. J. & Valentine, J. S. (1994). Bioinorganic Chemistry, 2nd edit University Science Books, Mill Valley, CA.
22. Cerasoli, E., Sharpe, B. K. & Woolfson, D. N. (2005). ZiCo: a peptide designed to switch folded state upon binding zinc. *J. Am. Chem. Soc.* **127**, 15008–15009.
23. DeGrado, W. F., Summa, C. M., Pavone, V., Nastri, F. & Lombardi, A. (1999). *De novo* design and structural characterization of proteins and metalloproteins. *Annu. Rev. Biochem.* **68**, 779–819.
24. Kohn, W. D., Kay, C. M., Sykes, B. D. & Hodges, R. S. (1998). Metal ion induced folding of a *de novo* designed coiled-coil peptide. *J. Am. Chem. Soc.* **120**, 1124–1132.
25. Kharenko, O. A. & Ogawa, M. Y. (2004). Metal-induced folding of a designed metalloprotein. *J. Inorg. Biochem.* **98**, 1971–1974.
26. Maglio, O., Nastri, F., Calhoun, J. R., Lahr, S., Wade, H., Pavone, V. *et al.* (2005). Artificial di-iron proteins: solution characterization of four helix bundles containing two distinct types of inter-helical loops. *J. Biol. Inorg. Chem.* **10**, 539–549.
27. Shiraishi, Y., Imanishi, M., Morisaki, T. & Sugiura, Y. (2005). Swapping of the β-hairpin region between Sp1 and GLI zinc fingers: significant role of the β-hairpin region in DNA binding properties of C_2H_2-type zinc finger peptides. *Biochemistry*, **44**, 2523–2538.
28. Iranzo, O., Ghosh, D. & Pecoraro, V. L. (2006). Assessing the integrity of designed homomeric parallel three-stranded coiled coils in the presence of metal ion. *Inorg. Chem.* **45**, 9959–9973.
29. Papworth, M., Kolasinska, P. & Minczuk, M. (2006). Designer zinc-finger proteins and their applications. *Gene*, **17**, 27–38.
30. Calhoun, J. R., Kono, H., Lahr, S., Wang, W., DeGrado, W. F. & Saven, J. G. (2003). Computational design and characterization of a monomeric helical dinuclear metalloprotein. *J. Mol. Biol.* **334**, 1101–1115.
31. Nanda, V., Rosenblatt, M. M., Osyczka, A., Kono, H., Getahun, Z., Dutton, P. L. *et al.* (2005). *De novo* design of a redox active minimal rubredoxin mimic. *J. Am. Chem. Soc.* **127**, 5804–5805.
32. Suzuki, K., Hiroaki, H., Kohda, D., Nakamura, H. & Tanaka, T. (1998). Metal ion induced self-assembly of a designed peptide into a triple-stranded α-helical bundle: a novel metal binding site in the hydrophobic core. *J. Am. Chem. Soc.* **120**, 13008–13015.
33. Farrer, B. T. & Pecoraro, V. L. (2003). Hg(II) binding to a weakly associated coiled coil nucleates an encoded metalloprotein fold: a kinetic analysis. *Proc. Natl Acad. Sci. USA*, **100**, 3760–3765.
34. Pozdnyakova, I. & Wittung-Stafshede, P. (2001). Biological relevance of metal binding before protein folding. *J. Am. Chem. Soc.* **123**, 10135–10136.
35. Bombarda, E., Grell, E., Roques, B. & Mely, Y. (2007). Molecular mechanism of the Zn^{2+}-induced folding of the distal CCHC finger motif of the HIV-1 nucleocapsid protein. *Biophys. J.* **93**, 208–217.
36. Du, D. G., Zhu, Y. J., Huang, C. Y. & Gai, F. (2004). Understanding the key factors that control the rate of beta-hairpin folding. *Proc. Natl Acad. Sci. USA*, **101**, 15915–15920.
37. Bunagan, M. R., Yang, X., Saven, J. G. & Gai, F. (2006). Ultrafast folding of a computationally designed Trp-cage mutant: Trp2-cage. *J. Phys. Chem. B*, **110**, 3759–3763.
38. Xu, Y., Purkayastha, P. & Gai, F. (2006). Nanosecond folding dynamics of a three-stranded β-sheet. *J. Am. Chem. Soc.* **128**, 15836–15842.
39. Sadqi, M., de Alba, E., Pérez-Jiménez, R., Sanchez-Ruiz, J. M. & Muñoz, V. (2009). A designed protein as experimental model of primordial folding. *Proc. Natl Acad. Sci. USA*, **106**, 4127–4132.
40. Sharpe, B. K., Matthews, J. M., Kwan, A. H., Newton, A., Gell, D. A., Crossley, M. & Mackay, J. P. (2002). A new zinc binding fold underlines the versatility of zinc binding modules in protein evolution. *Structure*, **10**, 639–648.
41. Sharpe, B. K., Liew, C. K., Kwan, A. H., Wilce, J. A., Crossley, M., Matthews, J. M. & Mackay, J. P. (2005). Assessment of the robustness of a serendipitous zinc binding fold: mutagenesis and protein grafting. *Structure*, **13**, 257–266.
42. Chen, Y. & Barkley, M. D. (1998). Toward understanding tryptophan fluorescence in proteins. *Biochemistry*, **37**, 9976–9982.
43. Zou, J. & Saven, J. G. (2000). Statistical theory of combinatorial libraries of folding proteins: energetic discrimination of a target structure. *J. Mol. Biol.* **296**, 281–294.
44. Kono, H. & Saven, J. G. (2001). Statistical theory for protein combinatorial libraries. Packing interactions, backbone flexibility, and the sequence variability of a main-chain structure. *J. Mol. Biol.* **306**, 607–628.
45. Oneil, K. T. & DeGrado, W. F. (1990). A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino-acids. *Science*, **250**, 646–651.
46. Nyborg, J. K. & Peersen, O. B. (2004). That zincing feeling: the effects of EDTA on the behaviour of zinc-binding transcriptional regulators. *Biochem. J.* **381**, e3–e4.
47. Berkovits-Cymet, H. J., Amann, B. T. & Berg, J. M. (2004). Solution structure of a CCHHC domain of neural zinc finger factor-1 and its implications for DNA binding. *Biochemistry*, **43**, 898–903.

48. Mély, Y., De Rocquigny, H., Piemont, E., Demene, H., Jullian, N., Fournie-Zaluski, M. C. *et al.* (1993). Influence of the N-terminal and C-terminal chains on the zinc-binding and conformational properties of the central zinc-finger structure of Moloney murine leukemia-virus nucleocapsid protein: a steady-state and time-resolved fluorescence study. *Biochim. Biophys. Acta*, **1161**, 6–18.
49. Mély, Y., De Rocquigny, H., Morellet, N., Roques, B. P. & Gerad, D. (1996). Zinc binding to the HIV-1 nucleocapsid protein: a thermodynamic investigation by fluorescence spectroscopy. *Biochemistry*, **35**, 5175–5182.
50. Chen, Y., Liu, B., Yu, H. T. & Barkley, M. D. (1996). The peptide bond quenches indole fluorescence. *J. Am. Chem. Soc.* **118**, 9271–9278.
51. Tucker, M. J., Tang, J. & Gai, F. (2006). Probing the kinetics of membrane-mediated helix folding. *J. Phys. Chem. B*, **110**, 8105–8109.
52. Buchsbaum, J. C. & Berg, J. M. (2000). Kinetics of metal binding by a zinc finger peptide. *Inorg. Chim. Acta*, **297**, 217–219.
53. Keeble, A. H., Hemmings, A. M., James, R., Moore, G. R. & Kleanthous, C. (2002). Multistep binding of transition metals to the H-N-H endonuclease toxin colicin E9. *Biochemistry*, **41**, 10234–10244.
54. Bombarda, E., Roques, B. P., Mély, Y. & Grell, E. (2005). Mechanism of zinc coordination by point-mutated structures of the distal CCHC binding motif of the HIV-1NCp7 proteins. *Biochemistry*, **44**, 7315–7325.
55. Heinz, U., Kiefer, M., Tholey, A. & Adolph, H. W. (2005). On the competition for available zinc. *J. Biol. Chem.* **280**, 3197–3207.
56. Dutta, S. J., Liu, J. & Mitra, B. (2005). Kinetic analysis of metal binding to the amino-terminal domain of ZntA by monitoring metal–thiolate charge-transfer complexes. *Biochemistry*, **44**, 14268–14274.
57. Bunagan, M. R., Cristian, L., DeGrado, W. F. & Gai, F. (2006). Truncation of a cross-linked GCN4-p1 coiled-coil leads to ultrafast folding. *Biochemistry*, **45**, 10981–10986.
58. Du, D. G. & Gai, F. (2006). Understanding the folding mechanism of an α-helical hairpin. *Biochemistry*, **45**, 13131–13139.
59. Qiu, L. L., Pabit, S. A., Roitberg, A. E. & Hagen, S. J. (2002). Smaller and faster: the 20-residue Trp-cage protein folds in 4 μs. *J. Am. Chem. Soc.* **124**, 12952–12953.
60. Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H., Woodruff, W. H. & Dyer, R. B. (1996). Fast events in protein folding: helix melting and formation in a small peptide. *Biochemistry*, **35**, 691–697.
61. Du, D. G., Tucker, M. J. & Gai, F. (2006). Understanding the mechanism of β-hairpin folding via ϕ-value analysis. *Biochemistry*, **45**, 2668–2678.
62. Wang, T., Zhu, Y. J., Getahun, Z., Du, D. G., Huang, C. -Y., DeGrado, W. F. & Gai, F. (2004). Length dependent helix-coil transition kinetics of nine alanine-based peptides. *J. Phys. Chem. B*, **108**, 15301–15310.
63. Bombarda, E., Morellet, N., Cherradi, H., Spiess, B., Bouaziz, S., Grell, E. *et al.* (2001). Determination of the pK_a of the four Zn^{2+}-coordinating residues of the distal finger motif of the HIV-1 nucleocapsid protein: consequences on the binding of Zn^{2+}. *J. Mol. Biol.* **310**, 659–672.
64. Siemann, S., Badiei, H. R., Karanassios, V., Viswanatha, T. & Dmitrienko, G. I. (2006). ^{68}Zn isotope exchange experiments reveal an unusual kinetic lability of the metal ions in the di-zinc form of IMP-1 metallo-β-lactamase. *Chem. Commun.*, 532–534.
65. Li, W. F., Zhang, J., Wang, J. & Wang, W. (2007). Metal-coupled folding of Cys_2His_2 zinc-finger. *J. Am. Chem. Soc.* **130**, 892–900.
66. Jager, M., Zhang, Y., Bieschke, J., Nguyen, H., Dendle, M., Bowman, M. E. *et al.* (2006). Structure–function–folding relationship in a WW domain. *Proc. Natl Acad. Sci. USA*, **103**, 10648–10653.
67. Ferreiro, D. U., Hegler, J. A., Komives, E. A. & Wolynes, P. G. (2007). Localizing frustration in native proteins and protein assemblies. *Proc. Natl Acad. Sci. USA*, **104**, 19819–19824.
68. Hills, R. D. & Brooks, C. L. (2008). Coevolution of function and the folding landscape: correlation with density of native contact. *Biophys. J.* **95**, L57–L59.
69. Ivarsson, Y., Travaglini-Allocatelli, C., Brunori, M. & Gianni, S. (2008). Folding and misfolding in a naturally occurring circularly permuted PDZ domain. *J. Biol. Chem.* **283**, 8954–8960.
70. Blasie, C. A. & Berg, J. M. (2002). Structure-based thermodynamic analysis of a coupled metal binding–protein folding reaction involving a zinc finger peptide. *Biochemistry*, **41**, 15068–15073.
71. Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2000). Measuring the rate of intramolecular contact formation in polypeptides. *Proc. Natl Acad. Sci. USA*, **97**, 7220–7225.
72. Roccatano, D., Sahoo, H., Zacharias, M. & Nau, W. M. (2007). Temperature dependence of looping rates in a short peptide. *J. Phys. Chem. B*, **111**, 2639–2646.
73. Magyara, J. S. & Godwin, H. A. (2003). Spectropotentiometric analysis of metal binding to structural zinc-binding sites: accounting quantitatively for pH and metal ion buffering effect. *Anal. Biochem.* **320**, 39–54.
74. Zhu, Y. J., Fu, X. R., Wang, T., Tamura, A., Takada, S., Saven, J. G. & Gai, F. (2004). Guiding the search for a protein's maximum rate of folding. *Chem. Phys.* **307**, 99–109.
75. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G. *et al.* (1984). A new force-field for molecular mechanical simulation of nucleic-acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784.
76. Kono, H. & Doi, J. (1996). A new method for sidechain conformation prediction using a Hopfield network and reproduced rotamers. *J. Comput. Chem.* **17**, 1667–1683.
77. Hoops, S. C., Anderson, K. W. & Merz, K. M. (1991). Force-field design for metalloproteins. *J. Am. Chem. Soc.* **113**, 8262–8270.
78. Dunbrack, R. L. & Cohen, F. E. (1997). Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Sci.* **6**, 1661–1681.