FOLDING OF HORSE CYTOCHROME *C*: THERMODYNAMIC, KINETIC, AND STATISTICAL ANALYSIS

A Thesis
Submitted for the Degree of

DOCTOR OF PHILOSOPHY

by

N. PRAKASH PRABHU

SCHOOL OF CHEMISTRY
UNIVERSITY OF HYDERABAD
HYDERABAD 500 046
INDIA

January 2007

CONTENTS

School of Chemistry
University of Hyderabad
Central University P. O.,
Hyderabad 500 046
India

STATEMENT

I hereby declare that the work embodied in this dissertation is the result of the investigation carried out by me in the School of Chemistry, University of Hyderabad, Hyderabad, under the supervision of **Prof. Abani K. Bhuyan.**

N. PRAKASH PRABHU

January 2007

STATEMENT VERIFIED

(Prof. ABANI K. BHUYAN)

PROJECT SUPERVISOR

School of Chemistry
University of Hyderabad
Central University P. O.,
Hyderabad 500 046
India

CERTIFICATE

This is to certify that the work embodied in this thesis entiled **'Folding of Horse Cytochrome *c*: Thermodynamic, Kinetic, and Statistical Analysis'** has been carried out by Mr. *N. Prakash Prabhu*, under my supervision and the same has not been submitted elsewhere for a Degree.

Prof. ABANI K. BHUYAN
(THESIS SUPERVISOR)

Dean
School of Chemistry

ACKNOWLEDGEMENTS

காலத்தி னாற்செய்த நன்றி சிறிதெனினும்
ஞாலத்தின் மாணப் பெரிது. திருவள்ளுவர்

A timely help, though thing of little worth
The favor itself, in excellence transcends the world. - Thiruvalluvar.

I wish to express my deep sense of gratitude and profound thanks to **Prof. Abani K. Bhuyan** for all of his support, guidance, and encouragement over the years. It has been an honor to work with and learn from him.

I wish to express my sincere gratitude to Prof. G. R. Desiraju, Prof. E. D. Jemmis (Former Deans) and Prof. M. Periasamy (Dean, School of Chemistry) for providing all necessary facilities to work in this school. I extend my sincere thanks to all faculty members of the school for their help through out my course.

It has been a great pleasure to work with our group of people over the years: Rajesh, Yadaiah, Krishna, Nageswar, Harish and Shiva. I am extremely thankful to all of them for stimulating discussions, timely help, and morale support. I am very thankful to **all** members of our school: research scholars, M.Sc., students and non-teaching staffs. I like to thank the University authorities for providing necessary facilities. Research fellowships and financial support from DBT and CSIR, Government of India, are gratefully acknowledged.

I thank all my teachers who taught and impressed me at all stages of my career. In particular, Ms. Mariammal, Ms. Selvi, Mr. Subramanian, Profs' Somanathan, Eswaran, Thamarai Chelvan, Janarthanan, Dr. Ganesan, Dr. Srinivasan, Mr. Elangovan, Mr. Petchi Muthu and Ms. Jayalakshmi for their guidance and encouragement even in my failures.

I thank my listless number of friends who love and support me for what I am. Particularly Vivek, Sekar, Mani, Hari, Jeevan, PK, Aaru, Thirupathy, Ramesh, Selva, Ammu and Lakshmi. I must acknowledge all my friends in the University especially the 'Tamil Group' consists of Selva, Balaram, Arumugam, Hemu, Deva, Padhu, Madhu, Sundaram, Francis, Vairam, Senthil, Venky, Kannapiran, Gnanavel, Sankar, Damu, Chandrasekar, Ice-Murali, Bharathi, Saravanan, Sathish, Prathap, Sakthi, Vijayan, Arun, Murali Jr, Jegan, Comp-Senthil, Comp-Murali, Master-Paranthaman, Stat-Saravanan, Bhuvana, Anita, Manasa, Nisha, Rohini, Subha, Babu, Kumaran, Rathitharan, Vaasu, Purshoth, Thirumurugan, Abijith, Robin, Partha, Kutti-Vijayan, Sundar, Kannan, Suresh and all who made my environment so homely and pleasant in the campus. My special thanks to my seniors Drs. V. S. Senthilkumar, S. Sivakumar, K. Jayakumar, K. Senthilkumar, Ashok kumar, M. Srinivasan, S. Perumal, Philip Anthony, T.M. Mariappan, V. Anbazhagan, P. Sivakumar and Math-Srinivasan.

I acknowledge the timely helps of Mr. C. S. Murthy, Mr. Suresh, and the staffs of CIL, UoH. I thank all the staffs in CMSD, UoH for their kind help, especially Mr. Fazi and Mr. Lenin.

I am very grateful to my parents for their blessings, love and concern on my career and success. I thank my sisters Dhana and Ratna and brothers Prabhu and Sekar for their affection and care. I acknowledge my uncles Muruganantham and Arumugam and aunt Rajalakshmi for their blessings.

There is no word to thank people who dedicated their life to research. I salute them, without their intense work and inherent wisdom I could not indeed make this thesis.

- N. Prakash Prabhu.

Abbreviations

cyt *c*	-	cytochrome *c*
ferricyt *c*	-	oxidized form of cytochrome *c*
ferrocyt *c*	-	reduced form of cytochrome *c*
cyt-CO	-	carbonmonoxide-bound reduced cytochrome *c*
cyt-NO	-	nitricoxide-bound reduced cytochrome *c*
neu ferrocyt *c*	-	reduced cytochrome *c* at pH 7
alk ferrocyt *c*	-	reduced cytochrome *c* at pH 12.8
NCO	-	native state of carbonmonoxy ferrocytochrome *c*
UCO	-	unfolded state of carbonmonoxy ferrocytochrome *c*
NNO	-	native state of nitrosyl ferrocytchrome *c*
UNO	-	unfolded state of nitrosyl ferrocytochrome *c*
GdnHCl	-	guanidine hydrochloride
ΔG^{o}	-	unfolding free energy change
C_m	-	unfolding transition midpoint
m_{eq}	-	equilibrium surface exposure during unfolding
T_m	-	transition midpoint temperature
TOF	-	time of flight
CD	-	circular dichroism
NMR	-	nuclear paramagnetic resonance
EPR	-	electron paramagnetic resonance
HX	-	hydrogen exchange
MG	-	molten globule state
A – state	-	molten globule state in acid pH
B – state	-	molten globule state in base pH
ΔH_m	-	enthalpy change at transition midpoint
ΔS_m	-	entropy change at transition midpoint
ΔC_p	-	heat capacity change
r – value	-	correlation coefficient
p - value	-	probability of falsely rejecting the null hypothesis

† Symbols for different parameters are explained fully in the text.

Preface

எப்பொருள் யார்யார்வாய்க் கேட்பினும் அப்பொருள்
மெய்ப்பொருள் காண்ப தறிவு.

திருவள்ளுவர்

Though things diverse from different sages' words
To discern the truth in everything is Wisdom.

Thiruvalluvar

I am like a boy playing on the sea-shore, and diverting myself now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.

Isaac Newton.

Protein Folding in Classical and Landscape Perspectives

Protein must adopt a specific folded three-dimensional structure to be biologically active. 'How does the one-dimensional sequence of amino acids in the protein chain determine its three-dimensional folded conformation in space?' is an elusive question in molecular biology for the past four decades, instigated by Anfinsen's experiments[1, 2] and followed by Levinthal's paradox.[3]

The recent advances in experimental and theoretical methods have envisaged a '*new view*' of protein folding.[4-7] In a metaphor of landscapes, the *new view* accentuates folding as a parallel flow process of an ensemble of chain molecules with multiple folding routes, but emphasizes less on specific structures and pathways. An energy landscape is the free energy of each conformation as a function of the degrees of freedom. The vertical axis of a funnel represents the 'internal free energy' and the lateral axes represent the multidimensional conformational coordinates with many degrees of freedom of a protein chain (Figure 1).

On the other hand, the '*classical view*' is based on simple phenomenological kinetic models, such as

the on-pathway model:

$$U \rightleftharpoons I \rightleftharpoons N \qquad (1)$$

the sequential model:

$$U \rightleftharpoons I_1 \rightleftharpoons I_2 \rightleftharpoons \ldots \quad N \qquad (2)$$

and the off-pathway model:

$$I \rightleftharpoons U \rightleftharpoons N \qquad (3)$$

where U, I_n and N is unfolded, intermediate, and folded states, respectively.

While the folding process has been defined classically as a linear sequence of events, the *new view* explains folding as diffusion-like processes involving parallel multiple pathways at the microscopic level.

Misfolded Intermediates: The landscape model considers misfolded intermediates as protein configurations dominated by low-energy, kinetically trapped non-native interactions. Folding proceeds further by slow reconfiguration of such conformations, goes uphill in energy surface, and thus become more unfolded in an energetic sense before proceeding to the native state. Similarly, the classical pathway model also accommodates misfolded species as off-pathway intermediates which have to unfold completely before proceeding to the native state.

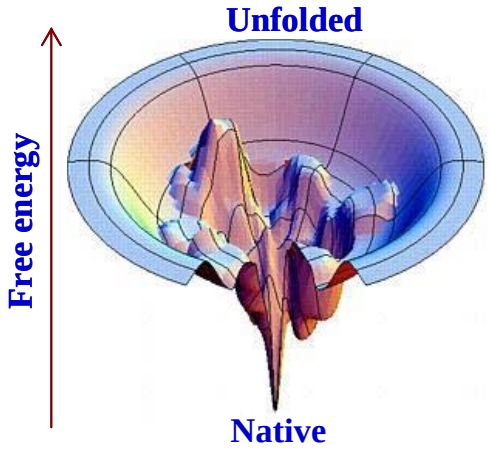


Figure 1. Energy Landscape metaphor of protein folding process.

Transition State: The *new view* characterizes the transition state as a broad ensemble without any specific structure. Small changes in solvent or temperature can vary the population of these molecular configurations. In classical sense, folding may proceed in two or more kinetic phases, with an initial fast collapse to a compact ensemble followed by slow reconfiguration into the native structure.

Molten Globule State: Landscape model defines molten globule as a collapsed, liquid drop-like state of a polymer with typically large conformational entropy and presumes its structure anywhere from negligible to a partial amount of native structure (Q~0.27).[6] Classical theory nevertheless associates molten globule with native-like chain topology in spite of its disordered tertiary structure, and places it on a probable on-pathway kinetic intermediate.

The contradictory explanations of folding scenario have generated a long debate in the folding community. Detailed experimental studies and valid analysis

of structural, thermodynamic, and kinetic properties of different proteins are necessity of the moment to investigate the enigma. In this regard, experiments on various forms of cytochrome *c* (Figure 2) and statistical analysis of small two-state proteins have been performed. The results corroborate the classical view rather than the funnel theory to explain the various aspects of protein folding.

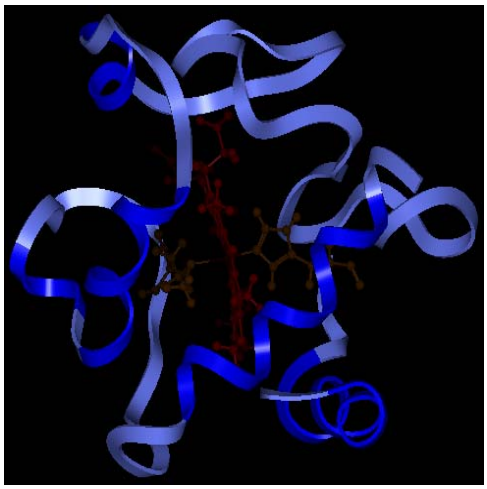
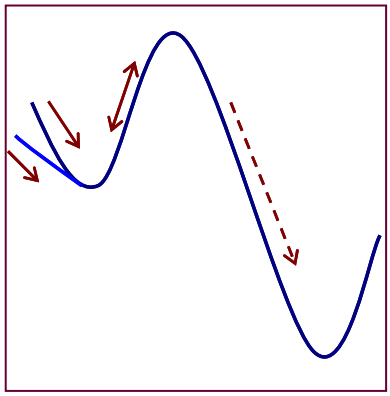


Figure 2. Ribbon diagram of horse Ferricytochrome *c*.

Chapter 1

Folding Barriers in Ferrocytochrome *c*: Support for a Classical Folding Pathway



1.1 Abstract

Proteins appear to pursue the classical principles of chemical kinetics. Test tube variants of ferrocytochrome *c* (ferrocyt *c*) whose thermodynamic stabilities are vastly different, refold to the same global minimum under a given final native condition, and they do so quickly at rates that do not reflect a strong dependence on the thermodynamic driving force. The transition-state ensemble is more unfolded-like, and the folding barrier offered is energetically sizable. The experiments involve neutral- (pH 7), alkaline- (pH 12.7), carbonmonoxy, and nitrosyl ferrocyt *c*, whose aqueous stabilities fall in the 3 – 18 kcal mol^{-1} range. But the large disparity in thermodynamic stability is not strongly reflected in their refolding rates. Burst relaxation signals in the refolding runs show the same quantitative dependence on GdnHCl, suggesting that the earliest relaxation or reconfiguration of the chains must be same and are independent of the initial equilibrium unfolded state. Analyses along the classical line indicate that the folding barrier is early in time and sizable in energy. These apparent inconsistencies raise some challenges to the predictions of the '*new view*' or funnel model of protein folding.

1.2 Introduction

One of the central problems in protein folding concerns the characteristics of inherent kinetic barriers that set the folding speed. Although certain properties of the rate-limiting transition state, including specific inter-residue interactions and dynamical order parameters, may be different for one protein from another, it is necessary to uncover the operative principles of the barrier-crossing event in order to understand the general mechanism of protein folding. Questions involving the placement of the transition state, the intrinsic or optional nature of the barrier, its height and the energetic stability, and the relationship of its energy to the refolding environment are directly relevant to address the folding problem.

Protein engineering and ϕ-value analysis has proved to be a powerful tool to probe the structure and dynamics of the transition state in a residue specific manner.[8-19] Complementary information has been provided by molecular dynamics simulation[20] and alternative interpretations of folding chevrons.[21, 22]

A different line of experiments, particularly suitable to folding kinetics of cyt *c* that relies on chemical modification of the protein and choosing appropriate solvent conditions to allow refolding, has also been useful to extract information about barriers that rate-limit folding.[23-29] The experiments involve removal of non-obligatory steps in folding that complicate interpretation of folding events in relation to the location and the organization of the transition state in cyt *c*. Non-native contacts between the heme and the polypeptide, prevalent in the unfolded state, lead to chain misfolding and obscure the relevant properties of the rate-limiting barrier.[24, 30] Persistent nonnative interactions whose dissolution rates are smaller than the protein-folding rate can act as fortuitous kinetic barriers. In essence, interpretation of data is simplified to a great deal when the optional events are suppressed.

The reduced state of horse cyt *c*, by virtue of ferrous heme reactivity, is intrinsically a two-state "fast folder".[30, 31] The absence of heme polypeptide misligation in the unfolded state removes interpretational ambiguity to a large extent. Heme–ligand replacements by chemical modification procedures that alter the protein structure minimally can be employed to study the folding barrier. An earlier study used a carbonmonoxy derivative of ferrocytochrome *c*, where the Fe^{2+}–M80 interaction is replaced by a Fe^{2+}–CO bond, to probe the folding barrier. The rationale for this replacement was the conjecture that interactions specific to achieving the native conformation of the M80 residue are the Fe^{2+}–M80 bond and the two key hydrogen-bonding interactions, one between the sulfur atom of M80, already engaged in iron bonding, and the side-chain oxygen of Y67, and the other

between the main-chain nitrogen of M80 and the hydroxyl oxygen of T78. The Fe^{2+}–M80 → Fe^{2+}–CO replacement did enhance the folding speed to some extent, but decelerated the unfolding rate by more than ten orders of magnitude.[28] Based on these observations, the Fe^{2+}–M80 bond and the M80-associated hydrogen bonding network were suggested to form at an early stage of folding.

This chapter describes equilibrium and kinetic folding studies with the carbonmonoxy and nitrosyl derivatives, and the alkaline form of ferrocytochrome *c*. The chemical modifications suppress the protein stability drastically at the cost of minimal structural alteration. The variants show the increase in unfolding rate, but make no change in the refolding rate. The invariant refolding rate thus is inconsistent with the earlier suggestion that Fe^{2+}–M80 bond and the M80-associated hydrogen-bonding cluster form at an early stage of folding.[28] The energetic stability of the transition state is indifferent to the formation of this structural cluster. In contrast to studies that have endeavored to establish the correlation between the slope of the folding funnel and the folding rate,[32-34] the results presented here show that the folding rate is independent of thermodynamic driving force, and the folding transition state is early in time and sizable in energy. The results provide substantial support for the classical folding pathway model, and are consistent, in particular, with the sequential folding mechanism already described for ferricyt *c*.[27,35,36] A folding model is presented that relies on a search in millisecond timescale for a transition state having native-like chain topology. The post-barrier folding is rapid because it is energetically downhill.

1.3 Materials and Methods

Cyt *c* was from Sigma (type VI). GdnHCl and sodium dithionite were obtained from USB and Merck, respectively. Pure and dry nitric oxide was prepared in the laboratory. Experiments at pH 7 were done in 0.1 M sodium phosphate buffer and at pH 13 were done in 10 mM CAPS buffer, containing 0.5–

3 mM freshly prepared sodium dithionite. Extreme care was taken to minimize air exposure of solutions.

1.3.1 ***Equilibrium Unfolding Measurements***:

Samples of ferrocytochrome *c*, prepared in the 0–7 M range of GdnHCl at neutral and alkaline pH, contained 10–15 μM protein. The solutions were deaerated and reduced under nitrogen with 2.3 mM sodium dithionite, and incubated in tightly capped quartz cuvettes or rubber-capped small glass tubes for 30 minutes. Tryptophan fluorescence excited at 280 nm (slit width 0.75 nm) was measured at 360 nm for neutral and at 370 nm for alkaline pH samples (slit width 1.25 nm) using a 1 cm square quartz cuvette in a photon-counting instrument. For equilibrium unfolding measurements of the nitrosylcytochrome *c* and carbonmonoxycytochrome *c*, the dithionite-reduced samples were saturated with NO and CO respectively, by passing a gentle stream of the dry gases through the solutions for one minute. The samples in sealed tubes were incubated for 45 minutes to ensure complete equilibration. Fluorescence emission measurements at 360 nm for cyt-NO and 370 nm for cyt-CO (excitation: 280 nm, 0.75 nm slit) employed a FluoroMax-3 instrument (Jobin-Yvon, Horiba). The obtained transition curves were least-squares fitted to a two-state model N ⇌ U using equation[46]

$$S_{obs} = \frac{C_f + m_f[D] + C_u + m_u[D]\exp\{(-\Delta G^\circ + m_g[D])RT\}}{1 + \exp\{(-\Delta G + m_g[D])/RT\}} \qquad (1)$$

where, S_{obs} is the observed signal, and C_f and C_u and m_f and m_u represent intercepts and slopes of native and unfolded baselines respectively, ΔG^o is free energy of unfolding. m_g (or the equilibrium *m*-value) is a parameter related to the change in surface area during the global unfolding of the protein, and [*D*] represents the concentration of the denaturant.

1.3.2 ***Measurements of Folding and Unfolding Kinetics***:

Cytochrome *c* (0.325 mM), initially unfolded in 6 M GdnHCl (pH 7), was reduced under nitrogen by the addition of a concentrated solution of sodium dithionite to a final concentration of 3.1 mM, and liganded with NO. Unfolded nitrosylcytochrome *c* thus obtained was loaded into a gas-tight syringe, and equilibrated at 10^{o}C for ten minutes. In the meanwhile, the refolding buffer solution containing the desired concentration of GdnHCl was reduced by the addition of dithionite to a final concentration of 3.1 mM, and filled in another gas-tight syringe. The protein and buffer solutions were then loaded into the stopped-flow mixing module. Folding was initiated by two-syringe mixing: 50 ml of the unfolded protein with 352 ml of the refolding buffer. The total flow rate was 8 ml/sec. After recording kinetics at a given concentration of the denaturant, a fresh protein solution and the refolding buffer of another concentration of the denaturant were loaded. The two-syringe procedure ensures minimal air oxidation of the protein. The final protein concentration in the folding mixture was 40 μM.

Unfolding experiments were performed following the same procedure of two-syringe mixing. Native nitrosylcytochrome *c* was prepared as follows. An unfolded protein solution (2.5 mM cytochrome *c*, 6 M GdnHCl), reduced and liganded with NO, was diluted tenfold with dithionite-containing native buffer to obtain the NO-bound folded protein in 0.6 M GdnHCl. This procedure was carried out manually in an inert atmosphere at room temperature before loading the protein sample into the stopped-flow syringe. The unfolding buffer containing a desired concentration of GdnHCl was reduced with dithionite, but contained no NO. The final protein concentration in unfolding experiments was ~30 μM. The unfolding kinetics were recorded in less than 20 minutes of the preparation of native nitrosylcytochrome *c*.

The kinetics of folding and unfolding were measured using an SFM 400 mixing module (Biologic) regulated at 10^{o}C by the use of an external water bath.

The 280 nm excitation source was obtained from a 150 W xenon lamp. Fluorescence emission by the mixed solution, contained in a 0.8 mm square flow cell, was measured using a 335 nm cut-off filter. The dead-time, determined experimentally by using the procedure suggested by the instrument manufacturer, was 3.2 ms for the high density mixer. Typically, 10–20 shots were averaged for noise reduction.

Similar two syringe mixing procedure was performed to follow the refolding and unfolding kinetics of carbonmonoxycyt *c* and alkaline cyt *c*.

1.3.3 ***Analysis of Kinetic Traces***:

Kinetic traces were fitted to single or double-exponential functions to obtain apparent rates, λ_i, the initial signal, S_o, which corresponds to the "zero-time" signal in the stopped-flow time window, the observed signal, S_{obs}, and the final equilibrium signal, S_{∞}, corresponding to the signal value at $t = t_{\infty}$. The S_o, S_{obs}, and S_{∞} signals were subjected to initial normalization by first subtracting the buffer fluorescence signals, and then dividing by the recorded signal of the unfolded protein at the highest GdnHCl concentration. In the unfolding set of measurements, the S_{∞} value of the kinetic trace at the highest concentration of GdnHCl employed was used to divide the fluorescence signals.

1.3.4 ***Chevron Analysis***:

The dependence of λ on GdnHCl concentration in the range 0.7–5.0 M was analyzed according to the two state model, N ⇌ U:

$$K = \frac{[U]}{[N]} = \frac{k_u}{k_f}$$

$\lambda_1 = k_u + k_f$

$\log k_u = \log k_u(H_2O) + m_u[\text{GdnHCl}]$

$\log k_f = \log k_f(H_2O) + m_f[\text{GdnHCl}]$ (2)

where, $k_u(H_2O)$ and $k_f(H_2O)$ are unfolding and refolding rate constants, respectively, in the absence of GdnHCl.

To account for the curvature in the unfolding limb of the rate-denaturant space the data were also analyzed by polynomial fits: [21]

$$\log k_f = \log k_f(H_2O) + m_{1,f}[\text{GdnHCl}] - m_{2,f}[\text{GdnHCl}]^2$$
$$\log k_u = \log k_u(H_2O) + m_{1,u}[\text{GdnHCl}] - m_{2,u}[\text{GdnHCl}]^2 \quad (3)$$

The fitted values of $k_f(H_2O)$ and $k_u(H_2O)$ along with the corresponding m values were compared with those calculated by the use of the equations:

$$\log k_f(H_2O) = \log k_u(H_2O) - \log K(H_2O)$$
$$\log k_u(H_2O) = \log k_f(H_2O) - \log K(H_2O) \quad (4)$$

where, $\log K(H_2O) = -\Delta G(H_2O) + m_g[\text{GdnHCl}]/2.3RT$

1.4 Results

1.4.1 *Preparation and Identity of Nitrosylferrocyt c*:

Nitric oxide has been known to bind to ferrocytochrome *c* by cleaving the Fe^{2+}-M80 bond even in aqueous solution at neutral pH.[37] Because the Fe^{2+}-M80 bond is disrupted in the unfolded state,[30] the binding affinity of NO for the ferrous heme is expected to be tighter under unfolding conditions. Indeed, the association equilibrium constant for nitrosylation of carboxymethylated cytochrome *c* in which the Fe^{2+}-M80 bond is disrupted by alkylation of the M80 side-chain, is larger than that for unmodified native ferrocyt *c* by at least two orders of magnitude.[37] Thus, it is easier to prepare UNO, and folding kinetics can be studied simply by diluting the protein solution with the refolding buffer. Since, the refolded product of UNO is NNO, the later can be used directly to measure unfolding kinetics. Here, NNO was prepared by using this protocol. It is then necessary to show the spectral and structural properties of NNO in order to establish that NO is persistently bound to the refolded product, and that it is no different from the native state of ferrocyt *c*.

Ligation of NO to the ferrous heme of folded, unfolded, and refolded cytochrome *c* can be demonstrated directly by their electron paramagnetic resonance (EPR) signatures (Figure 1). Since the frozen matrix of the GdnHCl-unfolded protein is not amenable to low-temperature measurements, ferrocyt *c* was nitrosylated at pH 13. Under this condition the protein is completely unfolded,[38] and can be refolded by simply diluting the solution to a final pH of 7. The X-band EPR spectra of UNO, NNO and refolded NNO clearly show the 'triplet of triplet' hyperfine structure, typical of nitrosylated cytochrome *c*.[41] It also confirms the fact that only one NO binds to the ferrous heme by replacing the M80 coordination. The

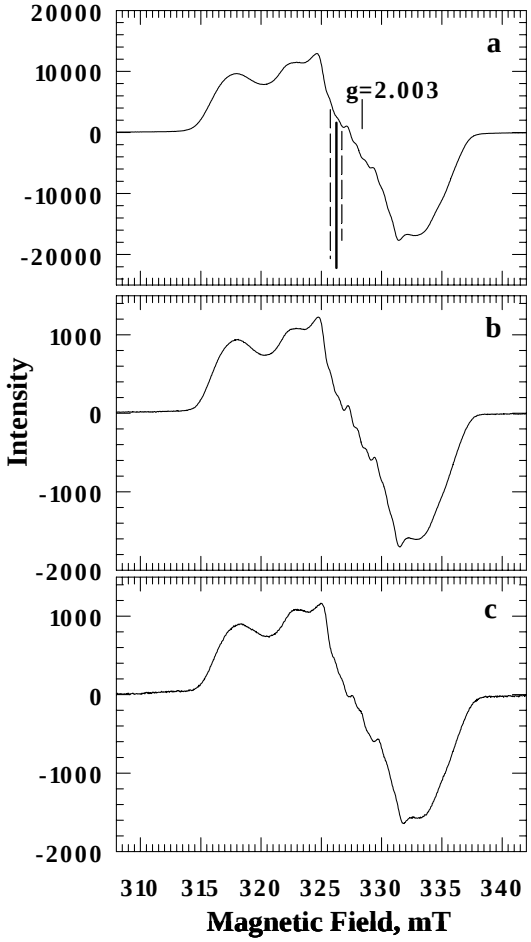


Figure 1. X-band EPR absorption at 110K of nitrosylated ferrocytochromes *c*. (a) UNO, unfolded at pH 13; (b) NNO, the native protein directly nitrosylated at pH 7; (c) NNO, refolded to pH7 from UNO, pH13. All spectra are virtually identical, characterized by the "triplet of triplet" hyperfine structure typical of NO-bound ferrocytochrome *c*. The heavy vertical line in (a) shows one of the three larger triplets, and the flanking broken lines show the hyperfine structure of this triplet.

other axial ligand is provided by H18. The spectra of NNO and refolded NNO are identical, suggesting that the refolded product achieved by the UNO → NNO reaction is structurally similar to the native state of the protein. The EPR spectra are dominated by the heme-NO interaction and not sensitive to tertiary structure changes.

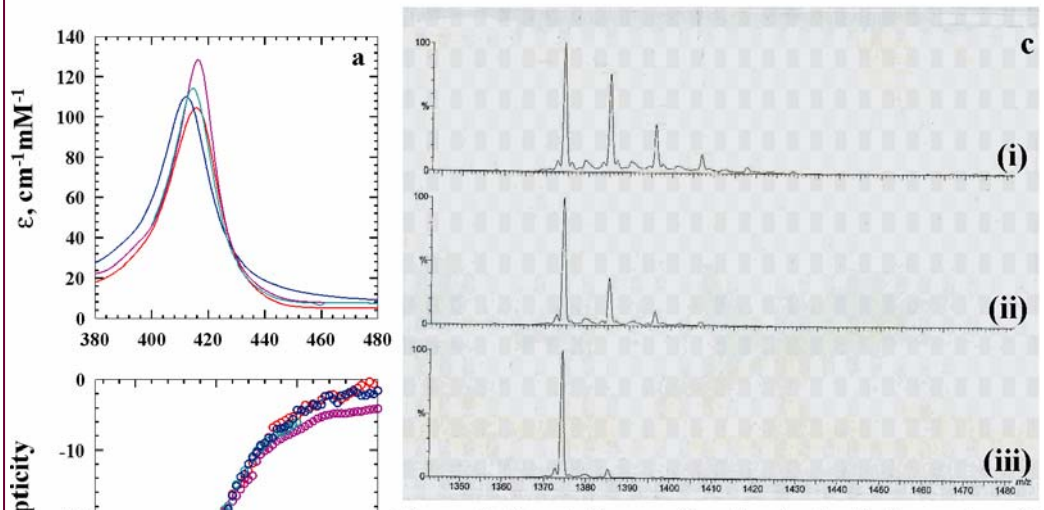


Figure 2. Spectral properties of variants of ferrocyt *c*. (a) Optical absorption spectra of Soret heme band. where, — represents ferrocyt *c* at neutral pH, — represents alkaline ferrocyt *c*, — represents NCO and — represents NNO. (b) Circular dichroism spectra at far UV – region. The colors represent the same as in panel a. (c) TOF mass spectra. (i) Carbonmonoxy ferrocyt *c* at pH 13 in 4M GdnHCl. (ii) Carbonmonoxy ferrocyt *c* at pH 13 without any GdnHCl. (iii) Native ferrocyt *c* at pH 6.5.

1.4.2 ***Structural Characterization of Carbonmonoxy Ferrocyt c***: Under normal physiological conditions the axial ligand M80 bound to the heme group of ferrocyt *c* is not replaced by extrinsic CO. However, exposing the protein to either high concentration of denaturant or high pH augments the ligation of extrinsically added CO. The unfolded ligated protein can be successfully refolded into its

native form with the Fe^{2+}–CO bond intact.[28] Far-UV CD spectrum shows that the refolded NCO has the same amount of secondary structural contents as the native protein (Figure 2b).

1.4.3 *Structural Identity of Alkaline Ferrocyt c*:

Ferrocyt *c* at pH 12.7 (alk ferrocyt *c*) is substantially structured. The NMR spectrum does not show any major changes in intensity and chemical shift of resonances (data not shown), relative to those seen for native protein at pH 7 (neutral ferrocyt *c*). Soret heme absorption spectrum of the alk ferrocyt *c* does not evidence for much structural changes in the native form, but the unfolded-state spectrum does. The spectra indicate that in contrast to the presence of a small population of penta-coordinate heme in GdnHCl unfolded neutral ferrocyt *c*,[30] the unfolded population at alkaline pH is entirely hexa-coordinate. Circular dichroism studies carried out at ultraviolet region (222nm) also confirm that the alkaline and the native form of ferrocyt *c* have the same secondary structural content. Moreover, the TOF mass spectrum of the alkaline ferrocyt *c* resembles that of the native form, suggesting no aggregation or deamidation of the protein at alkaline pH (Figure 2c).

These facts suggest that the three different forms: NNO, NCO, and alk ferrocyt *c*, are structurally similar with the native form. For all the forms, the native state is fluorescence-silent, but the unfolded state is not (data not shown).

1.4.4 *Equilibrium Unfolding of Ferrocyt c and Its Variants*:

The steady-state fluorescence intensity of ferrocyt *c* and its variants were measured as a function of GdnHCl. Figure 3 presents the normalized unfolding transition curves. Two-state analysis using equation (1) yields the free energies of unfolding, ΔG^{o}, and the equilibrium *m*-values (Table 1). The unfolding transition curves for the variants are shifted to the left relative to the curve for the native

state, implying that the native state is the most stable form, and that there is a significant change in the stabilities of different variants of ferrocyt *c*.

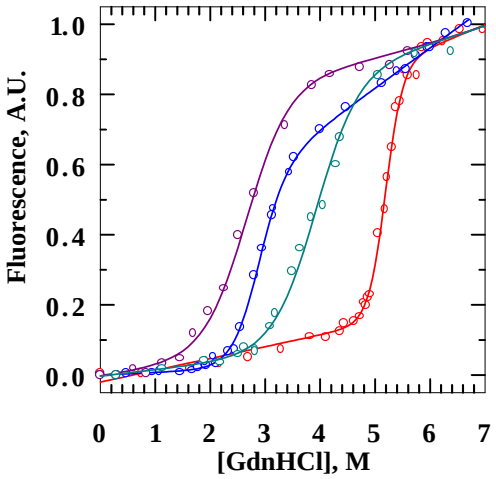


Figure 3. GdnHCl induced equilibrium unfolding of ferrocyt *c* at pH 7(o) and pH 13 (o), and carbonmonoxy ferrocyt *c* (o) and (o) nitrosylated ferrocyt *c*. The solid lines represent the fit to two-state model using equation 1. The fitted parameters are given in Table 1.

1.4.5 ***Folding and Unfolding Kinetics***:

1.4.5a. ***Kinetic Traces of Nitrosyl Ferrocyt c*:** Refolding to all concentrations of GdnHCl showed two kinetic phases: a fast major phase, and a slow minor phase. The amplitude of the minor phase of refolding averages to ~9% of the observed signal. This phase arises most likely from the fraction of oxidized protein that does not bind NO or from a small of non-monomeric aggregated product. This minor population refolds ~15 fold slower when native-like conditions (<1M GdnHCl) are approached. A significant fraction of the fluorescence signal expected for the folded protein is already recovered within the dead-time (~3.2 ms) of the instrument. In the presence of lowest concentration of GdnHCl (0.7 M), ~43% of the total expected signal is lost. This lost signal, also called missing amplitude or burst phase amplitude, has a strong dependence on the final concentration of GdnHCl in the refolding milieu (Figure 4b). However, the entire

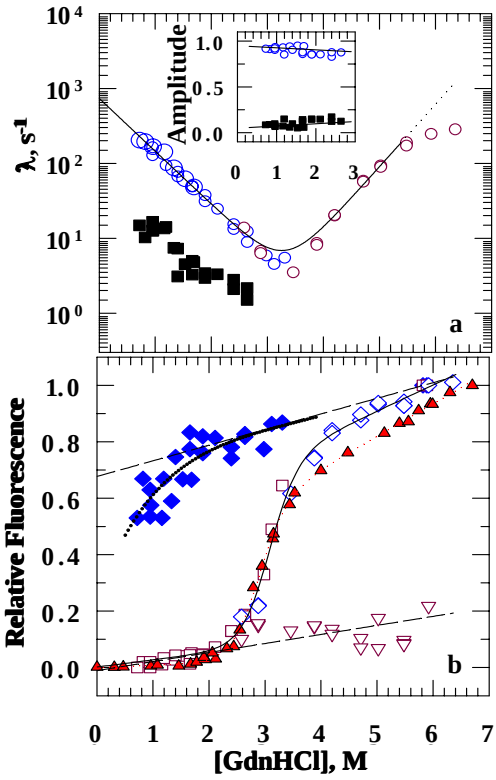


Figure 4. (a) Apparent rates of the major phase of folding (○) and unfolding (○), and the minor phase of folding (■) monitored by fluorescence. The continuous line represents two-state fit of data up to 5M GdnHCl using equation (2). The inset shows the 91:9 distributions of amplitudes of the major (○) and the minor (■) phase of folding. (b) Normalized signal values at time t = 0 (S_o)and t = ∞ (S_∞) in folding and unfolding kinetics of nirosylcyt c as a function of GdnHCl: S_o from refolding (♦) and unfolding (▽), and S_∞ from refolding (□) and unfolding (◊) experiments. The fluorescence monitored equilibrium unfolding (▲) is also replotted. The continuous line through the S_∞ data represents a two-state equilibrium fit.

unfolding course, regardless of final GdnHCl concentration, fits to a single exponential and, unlike refolding kinetics, there is no loss of the expected fluorescence signal within the dead-time of the spectrometer.

The chevron (Figure 4a) represents the denaturant dependence of the logarithm of the observed folding and unfolding rates at 10 °C. The relaxation minimum in the chevron plot (~3.2M GdnHCl) is reasonably close to the transition midpoint, C_m, for the equilibrium transition (N ⇌ U) of ferrocyt *c* in the presence of NO (Table 1). The folding limb of the chevron is linear within the limit of GdnHCl concentration employed in the refolding medium. The unfolding rate, although increases linearly with increasing concentration of the denaturant up to ~5M GdnHCl, deviates away from linearity under strongly unfolding conditions. The rollover in the unfolding limb has been suggested to originate from the

rate-limiting dissociation of M80 from the Fe^{2+}-M80 bond.[43] This suggestion is in conflict with the present observation that unfolding of nitrosylcytochrome *c* where the Fe^{2+}-M80 bond is eliminated, also displays a pronounced rollover. Iron-methionine deligation does speed unfolding, but does not mitigate the rollover. The continuous line through the chevron in figure 4a represents a two-state fit of folding kinetics by the use of equation (2). Rate data in the rollover region (>5M GdnHCl) were excluded from analysis. Values of $k_f(H_2O)$ and $k_u(H_2O)$ obtained from this analysis are listed in Table1.

1.4.5b. *Signal Amplitudes*: For all refolding reactions studied, a substantial fraction of the total expected fluorescence is lost in a burst phase that is completed within the dead-time of the stopped-flow. The burst phase analysis is based on the idea that GdnHCl dependence of the baseline signal (or $t \rightarrow \infty$ signal) of the burst kinetic phase, extracted by extrapolating the observed refolding curve to zero time (S_o), produces an unfolding transition curve of the burst folding intermediate.[47] It provides an approach to view the sub-millisecond regime when microsecond events have not been resolved in real time. Unfortunately, the procedure suffers from practical limitations of obtaining relatively accurate burst relaxation data. Nevertheless, figure 4b shows quantification of denaturant dependence of fluorescence signal amplitudes at different times of folding and unfolding. The fluorescence signal at both $t = 0$ and $t = \infty$ of refolding (S_o^f and S_∞^f respectively) have been normalized with reference to the fluorescence of the initial unfolded protein. Similarly, S_o^u and S_∞^u are, respectively, $t = 0$ and $t = \infty$ fluorescence of the unfolding kinetics normalized with respect to the $t = \infty$ signal of the unfolding in the presence of the highest concentration of GdnHCl employed. The denaturant dependence of both sets of values, S_∞^f and S_∞^u, coincides with the equilibrium unfolding transition except for the unfolded baseline region, due most likely to some slow relaxation of the UNO chain incomplete within the time window employed in the kinetic studies. The

continuous line is the iterated fit to S_{∞}^{f} and S_{∞}^{u} values for a two-state process.[46] The values obtained from the fit are 7.05 kcal mol^{-1} for the free energy of unfolding and 2.31 kcal mol^{-1} K^{-1} for the equilibrium *m*-value. The area between the S_{o}^{f} values and the unfolded baseline extrapolated

Table 1. Kinetic and thermodynamic parameters for folding and unfolding of various forms of ferrocyt *c*

	$\log k_f(H_2O)$	$\log k_u(H_2O)$	$\Delta G(H_2O)$ (kcal mol^{-1})		C_m (M GdnHCl)	
			Kinetic	Equilibrium	Kinetic	Equilibrium
[a]Ferrocytochrome *c* (equations 3 and 4)	2.7	-15.3	23.2	19.3	4.95	5.1
Nitrosyl cytochrome *c*						
Two-state analysis (equation 2)	2.87	-2.22	6.6	6.8	3.15	3.0
Polynomial fit (equations 3 and 4)	2.79	-4.50	9.4	6.8	3.35	3.0
[b]Carbonmonoxy cytochrome *c*						
Two-state analysis (equation 2)	3.40	-2.14	7.2	11.65	3.95	3.95
Polynomial fit (equations 3 and 4)	2.96	-4.30	9.7	11.65	3.95	3.95
Alkaline ferrocytochrome *c*						
Folding – Unfolding						
pH 7 ↔ pH 7	2.80	-10.5	17.8	18.0	5.0	5.6
pH 12.7 ↔ pH 12.7	1.83	-0.27	2.8	3.0	2.3	2.6
Cross-pH refolding						
pH 12.7 → pH 7	2.53					
pH 7 → pH 12.7	1.95					
Cross-pH unfolding						
pH 12.7 → pH 7		-10.5				
pH 7 → pH 12.7		-0.24				

[a] Data from Ref 30.
[b] Data from Ref 28.

linearly to native-like conditions contains the unobservable fluorescence signal associated with burst phase folding kinetics. There is some scatter in the S_o^f data, and their GdnHCl dependence does not indicate a distinct denaturant-induced phase transition. The data in the range 0.7-5.0 M GdnHCl simulate the empirical relation, $S_o = a + bx + cx(exp(-dx))$, where a, b, c and d are constants, and x is the concentration of GdnHCl. The denaturant distribution of S_o^u shows (Figure 4b) that only a minor fraction the total expected unfolding amplitude is lost in the burst phase.

1.4.5c. *Kinetic Traces of Alkaline Ferrocyt c*: In general, kinetic traces were fitted to 1-exponential. In some cases 2-exponentials were needed to improve the overall quality of fits. Figure 5 shows GdnHCl dependence of the logarithm of the fast folding (unfolding) phases. In each case, the relaxation minimum closely matches the midpoint of the individual equilibrium unfolding transition. Under strongly refolding and unfolding conditions, the rates deviate away from linearity. This feature of folding chevron has been described in earlier studies of ferrocyt *c* conducted at 10 °C.[30] The refolding speed at pH 12.7 appears somewhat slower relative to that of pH 7.

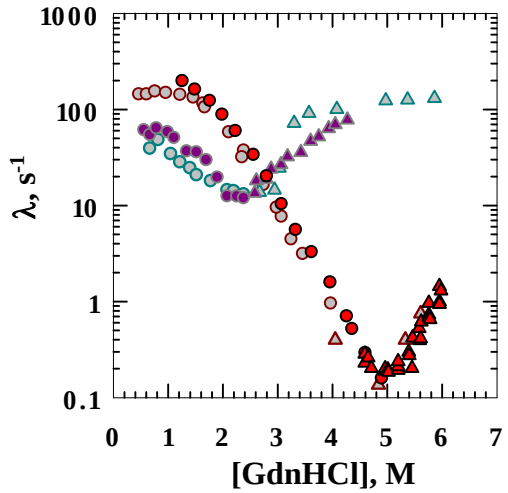


Figure 5. Rate-denaturant dependence for the reactions: pH 7→7 refolding (●) and unfolding (▲); pH 12.7→12.7 refolding (●) and unfolding (▲); pH 12.7→7 refolding (●) and unfolding (▲); pH 7→12.7 refolding (●) and unfolding (▲).

To examine how the structure, stability, and the configuration of the initial unfolded state affect kinetics of folding and unfolding, stopped-flow rates were measured by crossing from one pH to the other (i.e, pH 7→12.7 and pH 12.7→7) where the final condition was varied by adding different concentrations of GdnHCl. For clear comparison, the individual intra-pH data sets for neutral and alkaline ferrocyt *c* have been replotted. Remarkably, the GdnHCl dependences of folding-unfolding rates are almost identical for the pH 7→7 and pH 12.7→7 refolding, except for a small difference under strongly refolding conditions (<1 M GdnHCl) where the pH 12.7→7 refolding rate appears to slow down marginally. Likewise, denaturant dependences of folding rates measured by pH 12.7→12.7 and pH 7→12.7 refolding jumps are very similar. Folding and unfolding rates in water (k_f^o and k_u^o, respectively) were extracted by fitting the data to the quadratic relationship (Equation 3). Aqueous stabilities of neutral and alkaline ferrocyt *c* calculated from the fitted k_f^o and k_u^o values are in excellent agreement with those obtained from equilibrium unfolding (Table1).

1.4.5d. *Burst Phase Signals*: Figure 6 presents burst analysis for all the refolding and unfolding kinetic experiments. In pH 7→7 and pH 12.7→7 refolding reactions, the GdnHCl dependences of S_o values indicate (Figure 6a), within error, a distinct denaturant-induced phase transition. It may be interpreted as a sigmoidal curve which is missing the pre-transition, and thus the shape of the curve could simply reflect marginal stability of an intermediate. But, detailed equilibrium and kinetic studies[48] carried on cytochrome *c* confirm that the curves rather appear to simulate the GdnHCl dependence of fluorescence of the unfolded cyt *c*, and hence the burst phase may not entail an early folding intermediate. In contrast, the GdnHCl dependence of S_∞ signals of the observed refolding and unfolding traces produce a cooperative transition closely similar to the equilibrium unfolding curve for neutral ferrocyt *c*. In figure 6b, the same analysis

is presented for pH 7→12.7 and pH 12.7→12.7 refolding reactions. Again, the GdnHCl dependences of S_o signals are atypical of a cooperative transition, unlike the S_∞ signals whose denaturant titration reproduces the equilibrium unfolding transition of alkaline ferrocyt *c*. Thus, burst phase products are unlikely to contain any structures.

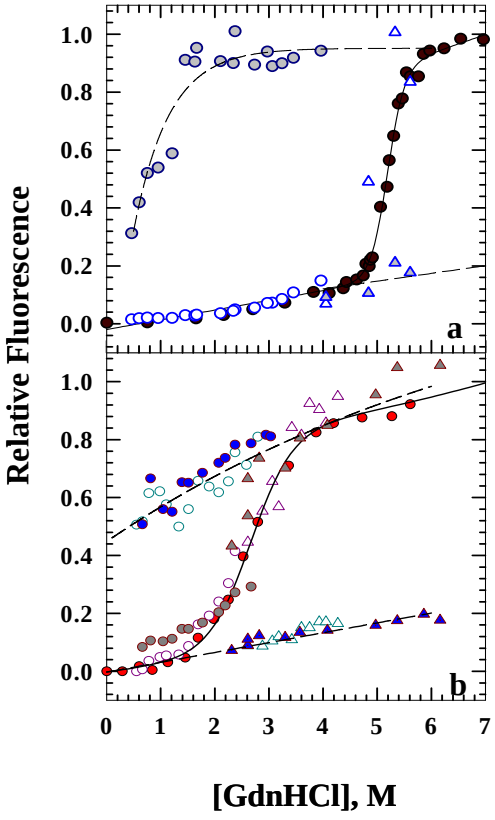


Figure 6. Normalized fluorescence signal values at time t = 0 and t = ∞ (S_o and S_∞ respectively) in folding and unfolding kinetics of different forms of ferrocyt *c*. (a) S_o from pH 7 ⇌ pH 7 refolding (●) and unfolding (▲); S_∞ from pH 7 ⇌ pH 7 refolding (○) and unfolding (△) reactions. Equilibrium unfolding transition (●) at pH 7 (b) S_o from pH 12.7 ⇌ 12.7 refolding (○) and unfolding (△); S_o from pH 7 → pH 12.7 refolding (●); S_∞ from pH 12.7 ⇌ pH 12.7 refolding (○) and unfolding (△); S_∞ from pH 7 → pH 12.7 refolding (●) reactions. Equilibrium unfolding transition (●) at pH 12.7. The dashed lines through S_o data in both panels are fits to the empirical relation $S_o = a + b \exp(cx)$, where *a*, *b* and *c* are constants and x is the concentration of GdnHCl. The solid lines through S_∞ data reproduce equilibrium stabilities under the final refolding-unfolding conditions.

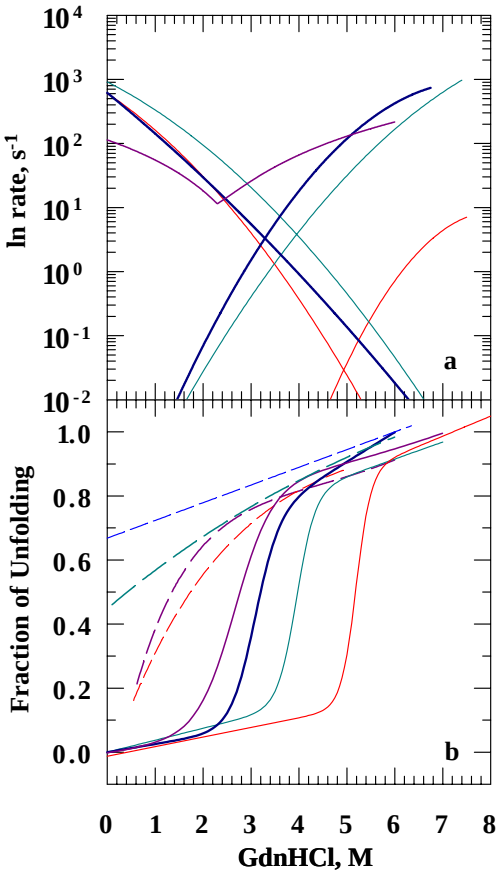


Figure 7. Comparison of the denaturant dependence of rates and amplitudes between ferrocytochrome *c* and its variants. (a) Two-state polynomial fits of folding chevrons. Note the vertical and horizontal shifts of the chevrons relative to each other. (b) GdnHCl dependence of S_o and S_∞ signals. The lines correspond to the test tube variants of ferrocyt *c* represented in Figure 3.

The closeness of S_o distributions shown for pH 7→7 and pH 12.7→7 and for pH 7→12.7 and pH 12.7→12.7 refolding reactions (Figure 6) provides a fundamental insight to the response of the unfolded chains when they are allowed to refold. The unfolded states of ferrocyt *c* at neutral and alkaline pH are unlikely to be similar at least structurally due mainly to differences in ionization and chain solvation. No matter how different they are in terms of chain configuration, chain solvation, or even stability, their fluorescence signals on being driven to refold under a given final solvent condition are nearly identical, suggesting that the earliest relaxation or reconfiguration of the chains must be the same and is independent of the initial equilibrium unfolded state. The GdnHCl dependences of S_∞ values obtained from pH 7→7 and pH 12.7→7

refolding reactions are the same within experimental error, indicating that the refolded native proteins under a given final conditions are structurally identical irrespective of the initial unfolded state. The same argument holds for the S_∞ data shown for pH 7→12.7 and pH 12.7→12.7 refolding reactions. These results provide a basis for phenomenological description of protein folding kinetics.

1.4.2e. *Comparison of All Forms of Ferrocyt c*: Earlier studies using ferrocyt *c*[30] and carbonmonoxycyt *c*[28], in which the Fe^{2+}-M80 replacement was introduced, have shown the role of the Fe^{2+}-M80 interaction in native-state stability and unfolding of cyt *c*. In the context, it will be useful to compare the major results of those studies with the present ones.

Figure 7a illustrates the polynomial-fitted chevrons of ferrocyt *c* (Fe^{2+}-M80), carbonmonoxycyt *c* ($Fe^{2+}CO$), nitrosylcyt *c* (Fe2+NO) and alkaline cyt *c*. In the order of apparent unfolding speed, alkalinecyt *c* > nitrosylcyt *c* > carbonmonoxycyt *c* > ferrocyt *c*. When extrapolated to the ordinate, $k_u(H_2O)$ values of nirosyl-, carbonmonoxy- and alkaline cyt *c* are very similar but are larger than the $k_u(H_2O)$ for ferrocytochrome *c* by at least ten orders of magnitude, indicating that the Fe^{2+}-M80 interaction has little role in controlling the folding speed. It is remarkable that the folding times of the three variants of ferrocyt *c*, in spite of being drastically different from each other in terms of stability, are nearly the same. Figure 7b compares the GdnHCl dependence of equilibrium fluorescence of the three proteins; values of $\Delta G(H_2O)$ are listed in Table 1. Also shown are the distribution of S_o^{fol}, the fluorescence at time t=0 of stopped-flow refolding. The S_o^{fol}-denaturant distributions for all three proteins are very similar within the experimental resolution achievable; suggesting that the immediate response of the unfolded polypeptides on being driven to refold is similar for all three proteins. However, none of the three distributions indicates a distinct denaturant-induced phase transition, suggesting that the burst phase products are unlikely to contain tertiary structure.

1.5 Discussion

In an attempt to delineate the character and the magnitude of the rate-limiting barrier in the folding of cytochrome *c*, the present work examines the stability and the folding speed of the nitric oxide derivative (nitrosyl cytochrome *c*) and the alkaline form of reduced cyt *c*. In the following, these results are compared and contrasted with those already available for the CO derivative (carbonmonoxycytochrome *c*) and the underivatized form of ferrocytochrome *c* under similar conditions to facilitate understanding the nature of the barrier and the folding mechanism of cytochrome *c*.

1.5.1 *Equilibrium Stability and Folding Speed*:

The folding equilibrium in each case is apparently two-state, involving only the native and the unfolded state without accumulation of structural intermediates to any detectable level. (Figures 3 and 7).[28, 30] The four proteins, even though structurally much similar, are vastly different in terms of equilibrium stability: ΔG (H_2O) values are 19.3, 6.8, 11.65 and 3.0 kcal mol^{-1} for ferrocyt *c*, nitrosylcyt *c*, carbonmonoxycyt *c*, and alkaline cyt *c*, respectively (Table1 and Figure 8). These stability differences, however, do not affect the folding rate of ferrocyt *c* to any considerable extent (Figures 5 and 7, and Table 1), a result that has major bearings on the protein folding mechanism. With regard to unfolding in water, the rates for nitrosylcytochrome *c*, carbonmonoxycytochrome *c*, and alkaline cytochrome *c* are faster than that of ferrocytochrome *c* by at least 10 orders of magnitude. These observations clearly show the enormous kinetic stability provided by the Fe^{2+}-M80 bond and the H-bond cluster involving the M80 to the cytochrome molecule.

1.5.2 *Burst Phase Signals*:

Another important element is the burst refolding kinetics. In refolding kinetics of all of the four proteins considered here, a part of the total expected

fluorescence signal is already lost in the submillisecond time: the stronger the refolding condition, the larger the signal loss (Figures 4b, 6 and 7b). The attribute to the burst phase is often referred to as controversial. A simple interpretation of the burst kinetics in the case of refolding of the oxidized form of cytochrome *c* is based on specific submillisecond chain collapse leading to the accumulation of a compact and partially α-helical intermediate, I_c.[49,50] It is also thought that burst phase signals may often arise from rapid non-specific structural rearrangements or simply contraction of the unfolded polypeptide in response to its transfer from the unfolding solvent to a strongly refolding solvent as is done in mixing experiment. It is indeed the case already justified by diluting non-folding fragments of cytochrome *c* from higher to lower concentrations of denaturant.[51,52] Experiments also showed that even intact cytochrome *c*, when transferred from one to another concentration of urea or GdnHCl under conditions where the protein is persistently unfolded, produces burst signals the denaturant dependence of which matches that observed in the real refolding experiment.[53] These results suggest that optical burst signals do not originate from a structured kinetic intermediate, but rather from a chain contraction event when the polypeptide is transferred to a milieu of different concentration of the denaturant. Direct submillisecond CD studies have also shown that very little secondary structure is formed in the submillisecond burst pahse.[54,55] It is therefore more likely that the product of the burst phase is just a collapsed unfolded state under native-like conditions rather than a structural intermediate. Results of the present study weigh in rather substantially on this side of the controversy about the character of the burst phase.

1.5.3 ***Barrier Height in Cytochrome c***:

By assuming that the apparent activation energies for folding and unfolding in water, E_a^f and E_a^u, respectively, are given by an expression of the type: $E_a = -RT \ln(k/A_o)$, where k is the rate coefficient $\{k_f(H_2O)$ and $k_u(H_2O)\}$, and A_o is the front factor, the heights of the activation barriers for folding and

unfolding in water can be estimated, although the appropriateness of the use of this thermally activated rate law in its general form is not quite clear.[56,57] $A_o \sim 2 \times 10^{10}$ s^{-1}, extracted from classical Arrhenius temperature dependence of refolding rates of ferrocyt *c*, was used in the previous study of carbonmonoxycyt *c*.[28] Recent theory[56,58,59] and strategic experiments[60,61] project it in between 1×10^6 and 1×10^8 s^{-1}. Using $A_o = 1 \times 10^8$ s^{-1}, suggested by Krieger *et al.*,[61] values of E_a^f and E_a^u obtained for the four different forms of reduced cytochrome *c* considered here are listed in Table 2.

Table 2. Barrier energies and surface burials in the transition states

	E_a^f (kcal mol^{-1})	E_a^u (kcal mol^{-1})	$\alpha_f^\#$
Neutral ferrocytochrome *c*	6.9	30.2	0.35
Nitrosylcytochrome *c*	6.6	16.1	0.37
Carbonmonoxycytochrome *c*	6.8	16.4	0.30
Alkaline ferrocytochrome *c*			
pH 7 ↔ pH 7	6.9		0.10
pH 12.7 ↔ pH 12.7	8.2	-	0.21
pH 12.7 → pH 7	7.3		0.05
pH 7 → pH 12.7	8.1		0.29

Interestingly, the folding energy barrier has the same height (E_a^f = 7.0 (± 0.5) kcal mol^{-1}) for all the four proteins. It also suggests that the Fe^{2+}-M80 bond and hydrogen bonding interactions involving M80 are not determinants of the folding barrier. These interactions are formed only after the passage over the barrier. This is consistent with the fact that the heme iron of cytochrome *c*, if bonded to non-native ligands in the initial unfolded state, finds its native-state M80 ligand in the final step of the sequential folding pathway.[62] On the other hand, the unfolding barrier height of 30.2 kcal mol^{-1} for ferrocyt *c* is reduced to 16.2 kcal mol^{-1} for the carbonmonoxycyt *c* and nitosylatedcyt *c*, suggesting that

the heme-M80 ligation and the three H-bonding interactions involving M80 somehow affect the height of the unfolding energy barrier. Evidently, this local structural cluster imparts massive kinetic stability to decelerate unfolding by more than ten orders of magnitude.

1.5.4 *Barrier Location*:

Experimentally, the location of the transition state is approximated by $\alpha = m^{\#}_{f\,(u)} / m_g$, where $m_f^{\#}$ and $m_u^{\#}$ are, respectively, kinetic *m*-values for refolding and unfolding.[63] Table 2 lists the α-values approximated from the chevrons presented in Figure 7. For all four variants of ferrocyt *c*, the α-value for refolding lies between 0.1 and 0.35, suggesting that only one-third or less than that of the surface that is buried in the native state becomes buried in the folding transition state. It follows that the rate-limiting step is early in folding, a derivation that lends significant support for the classical folding model of cytochrome *c*, where the rate-limiting step is the very fast step in a specified pathway involving cooperative folding units.[36]

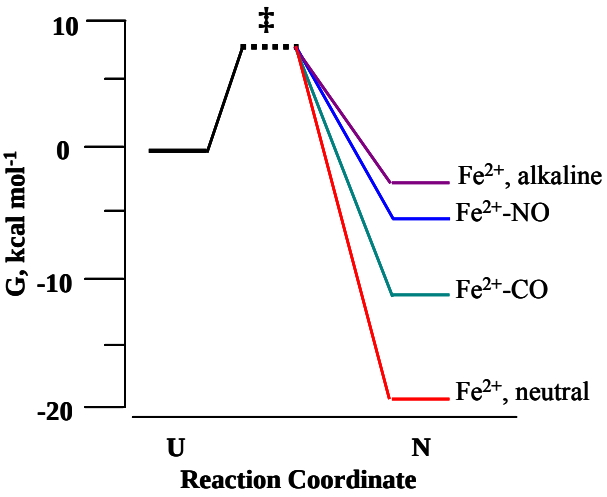


Figure 8. A free energy diagram illustrating the native, unfolded and transition states. Energy levels were estimated from Tables 1 and 2 and are scaled with reference to the energy of the U' state arbitrarily set to zero.

1.5.5 *Conflicts with Theoretically Based Ideas:*

The conclusions reached here raise some challenges to the tenets of funnel-shaped energy landscape perspective or '*new view*' [4,5,64-66] that emanates

from microscopic theory and computer simulations of protein folding.[67-69] According to the new view, the folding time correlates approximately with the energy of the native state, [67,70] and folding is rate-limited by the size and slope of the folding funnel. But the experimental observations presented here show that there is no effect on the folding rate of ferrocyt *c* in spite of a 17 kcal mol^{-1} change in stability (Figures 7 and 8, and Table 1). This is so in spite of the fact that cytochrome *c* folds in an apparent two-state manner. The new view suggests that the folding transition-state region appears rather late at a time when a substantial fraction of native contacts are already formed (Q ~0.6).[64,68] But the results here show that the rate-limiting barrier is early (Table 2). The present study demonstrates a sizable barrier for refolding reaction, whereas the new view envisages no apparent barrier or perhaps a tiny barrier (~ $3k_BT \approx 1.687$ kcal mol^{-1}at 10 ^oC[64]). More specifically, smooth-funneling witnesses no specific bottle neck conformation. The folding rate is intrinsically limited by the steepness and size of the funnel.

These observations are not consistent with the belief that folding speed depends on the global thermodynamic driving force. The results, in fact, fit exactly into the picture of a simple classical free energy profile with folding rate determined by the energy difference between the unfolded state and the transition state.

1.5.6 ***A Classical Folding Model of Cytochrome c***:

The details of thermodynamic stability, refolding rapidity and the nature of the transition state derived here along with the basic results from cytochrome literature allow one to assemble a folding pathway, the key events and plausible interpretations of which are summarized below.

1.5.6a. ***Kinetically Constrained Ultrafast Chain Relaxation*****:** As soon as a refolding condition is created, the unfolded chains (U) initially at low-energy equilibrium states are driven instantaneously to new higher energy level(s).[71] The excited chains, now in non-equilibrium states, respond to the new solvent condition where only few, if any, denaturant molecules can affect the nascent polypeptides (U'). The exact nature of this ultrafast relaxation is not clear. From the extensive equilibrium and kinetic studies carried on cytochrome *c* in different denaturants, [31] it appears that the initial ultrafast relaxation may involve random collapse or chain contraction, desolvation and/or redistribution of backbone dihedral φ-ψ angles.[24,51,72] It is missed out in millisecond stopped-flow measurements and is termed burst folding phase (Figure 6). The burst relaxation may escape experimental detection if the probe employed is silent to the U→U' relaxation.[73]

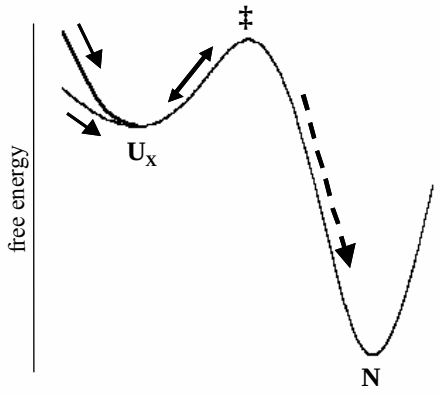


Figure 9. Schematic of the classical folding profile (U→U'→‡→N). High energy chains produced instantaneously with a refolding jump relax to the nascent state, U', where they are energetically degenerate. A long search for an appropriate set of tertiary interactions that can support a native-like chain topology ensues. The transition-state is achieved when the correct combination of interactions develops. Folding is then rapid and energetically downhill, and is not strongly dependent on the slopes and depths of energy surface.

Figure 9 schematizes folding of two configurationally different unfolded states of the same amino acid sequence to a given refolding condition. Instantaneous with the refolding jump is the deformation of initial unfolded energy levels. Under the same refolding condition, the 'zero-time' non-equilibrium states of different forms of ferrocyt *c*, which are not yet relaxed, may or may not be different. In any

case, at the end of the ensuing relaxations the nascent states (U') derived from neutral-unfolded and alkali-unfolded chains (U) are identical and energetically degenerate, as though a coherence has been established. This follows from near-concurrence of denaturant distributions of burst signals for all the variants of ferrocyt *c* (Figure 6). Implicit to the model is the deterministic nature of the U'→U process. Structural or configurational differences in the initial unfolded states do not alter the refolding reaction; it is the amino acid sequence that determines the nature of the relaxation and the configuration and/or redistribution of dihedral angles in U'. This is a kinetically constrained process. ΔG^{o} values determined from equilibrium and kinetic experiments (Table 1) appear to indicate energetic similitude of U and U' states. This follows from the result that aqueous stability of different forms ferrocyt *c*, measured from equilibrium experiments appears indifferent to that estimated from kinetic experiments. Equilibrium data provide ΔG^{o} for the two-state U ⇌ N reaction, whereas ΔG^{o} calculated by using the rate constants reflects the energy difference between U' and N, both under final refolding condition, since in millisecond kinetic experiments all information about the energetics of the initial U state is lost in the dead-time of mixing. The U states may still differ in energy to some extent, but such differences are masked by error in data.

1.5.6b. ***Search for Transition-State Topology and Barrier Crossing*****:** The ability to assign the sub-millisecond event to a non-specific process of chain collapse simplifies the folding problem of cyt *c*. Because folding-unfolding kinetics of ferrocyt *c* exhibits two-state kinetics,[28,30,31] only the U'→N process, where U' and N are separated by a barrier, needs to be considered, and the earlier observations augmented by the present ones indicate that the U'→N reaction time is ~2 ms at 10 ^{o}C (Table 1). This suggests that the folding reaction is effectively a slow process relative to the fast timescales appropriate for the formation of local

structural units or short secondary structural elements in isolation. The sluggishness of the reaction must be due to the 'long time' U' takes to find a supportive transition state. Since there is no post-barrier kinetic intermediate for two-state proteins such as cyt *c*, further folding will be rapid and downhill once a stable transition state is achieved. In addition, the early rate-limiting barrier in the refolding ($\alpha \leq 0.35$, Table 1) suggests a less-organized transition state in terms of native-like surface burial. Such a transition state could possibly be supported by a few correctly formed native-like tertiary contacts that can provide a native-state chain topology. This view, also held earlier by Sosnick *et al*.,[24] draws support from proposals that folding rates and mechanisms often appear to be largely determined by the native-state topology or contact order,[74,75] and that the topology has a bearing in defining the structure of the rate-limiting transition state.[76,77] The possibility of search-finding a set of appropriate tertiary contacts within the folding time of a millisecond or longer appears conceivable from recent advances in theory and experiments of the rate of inter-residue contact formation in chains.[58,78,79] with local conformational transitions.[80]

1.5.6c. ***Post-Barrier Downhill Folding*:** Achieving the transition state is analogous to the assumption of the classical (transition state) theory, that systems that have surmounted the barrier in the direction of products do not turn back. Folding kinetics of ferrocyt *c* is apparently two-state, and do not indicate accumulation of structural intermediate. The post-transition reaction coordinate shows no minima appreciably low in energy and the folding is rapid and downhill. However, earlier results of real-time NMR HX experiments[81] and equilibrium NMR HX data[62] have demonstrated the existence of short-lived intermediates. They could populate the kinetic pathway when bad contacts frustrate folding in the post-transition stage, as in the case with the refolding of oxidized cytochrome *c*.[23,82]

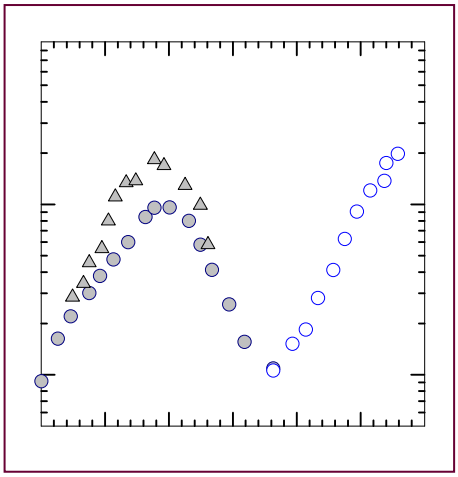
The stages of this pathway and the placements of different elements are based on the results obtained with the reduced form of cyt *c*. The model is similar, except for some details, to the search-nucleation mechanism[27,29] that emanated from a series of elegant experiments with the oxidized form of cyt *c*.[24,51] It turns out that the refolding kinetics of the oxidized and the reduced forms of cyt *c* are almost identical when the former is allowed to refold under conditions where the persistent ligation of non-native histidine residues to the ferric heme is eliminated. The need for the development of a correct set of tertiary interactions to achieve the transition state is a similarity of the proposed mechanism to the nucleation-condensation model.[83] In the latter, these interactions are required to stabilize the nucleus composed of adjacent residues. More studies will be needed to examine the existence of nucleation site in cytochrome *c*.

1.6 Conclusion

The folding rate of ferrocytochrome *c* is independent of thermodynamic driving force, and the folding transition state is early in time, and substantial in energy. When allowed to refold, the unfolded chain undergoes a submillisecond non-specific chain collapse (U$\rightarrow$U'). The ensuing U' $\rightarrow$ N reaction is apparently two-state, and is slow because of the time it consumes in the search for a few appropriate native-like tertiary contacts that could support a native-state chain topology. This process rate-limits folding. Acquisition of native structure thereafter is rapid and downhill without accumulation of structural intermediates. The Fe^{2+}-M80 bond and the hydrogen bonding interactions involving M80 are not determinants of the folding barrier. They possibly provide sufficient kinetic stability to protect ferrocytochrome *c* from unfolding during the large-amplitude conformational fluctuations required for its electron donating performance.[84]

Chapter 2

Dead-end Intermediate in Ferrocytochrome *c* Folding



2.1 Abstract

Formation of unproductive or misfolded off-pathway structures during the folding of single-domain protein has been a subject of debate for a long time. While the refolding of horse ferrocytochrome *c* from the GdnHCl-unfolded state at pH 7 is simple and two-state, the alkali-unfolded protein liganded with carbonmonoxide exhibits extreme misfolding. The misfolded product typifies a dead-end intermediate, and unfolds to the initial nascent state in order to have a chance to fold correctly to the native state. The rate of unfolding of the abortive misfold limits the overall rate of protein folding. The misfolding possibly results from an acute failure of the polypeptide chain to achieve the supportive transition state by stochastic search relevant for successful folding. Such misfolding may be analogous to the failure of a sizable set of proteins in the intracellular milieu to fold to the functionally active native state.

2.2 Introduction

As explicable by the principles of chemical kinetics, the nascent polypeptide chain has to pass through a relatively high-energy transition region to achieve the functionally active native state of a protein. Recent advances in theory and experiments with small single-domain proteins appear to indicate that the transition state for folding is topologically similar to the native state.[74,75,85] The polypeptide chain finds the transition state presumably by a time-consuming search for certain crucial intra-polypeptide interactions that can support the required native-like chain topology.[24,27,85] The stochastic nature of the search mechanism, however, does not guarantee the full achievement of the relevant transition state. A failure could result in a wrong transition state with undesirable chain topology, the passage over which may then lead to partial or complete misfolding depending on the extent and location of the nonnative interactions. Instances of formation of some nonnative interactions in certain segments, and

hence transient trapping of folding structures in local energy minima, are known.[88-90] In the case of acute misfolding, where overwhelming majority of nonnative interactions are acquired, the misfolded state must unfold to the initial state of the nascent chain in order to have a chance to refold on the right track. Classical kinetics place such an intermediate at the dead-end of the folding pathway schematized by $I \rightleftharpoons U' \rightleftharpoons N$, where I, U′, and N, all three considered under the final refolding condition, refer to the off-pathway intermediate, the nascent state, and the native state, respectively. The misfolded state is prone to aggregate or interact with other molecules, which seems to be the case for some proteins within the cell.[91] One of the functions of cellular chaperones is to shield and protect the nonnative polypeptides from misfolding and subsequent undesirable processes.[92,93] Folding experiments *in vitro* have, however, not reported misfolding so severe that the consequent product is aborted altogether.

This chapter provides a direct demonstration of misfolding of equine ferrocytochrome *c*, by observing denaturant dependence of the refolding rate constant of the alkali-unfolded protein at a final pH of 7. Auxiliary kinetic experiments, analyses, and kinetic modeling place the misfolded species at the dead end of the folding pathway. NMR mapping of amide hydrogen exchange kinetics shows that the misfolded species refolds to the correct native state via the unfolded state without engaging any intermediate structure.

2.3 Materials and Methods

Cyt *c* was purchased from Sigma (Type VI). GdnHCl and sodium dithionite were obtained from USB and Merck, respectively. Pure and dry carbon monoxide gas was prepared in the laboratory. All experiments were done in strictly anaerobic atmosphere at 22 ^{o}C using 0.1M sodium phosphate buffer for pH 7 conditions, and NaOH (with or without 1mM CAPS) for pH 12.7

conditions. Solutions contained 0.5-3mM freshly prepared sodium dithionite, and experiments were completed within 2 hours of exposing the protein to high pH.

2.3.1 ***Equilibrium Unfolding:***

Cyt *c* solutions containing 1 mM CO, prepared in the 0-7 M range of GdnHCl or in the 7-13 range of pH were deaearted and reduced under nitrogen with 1 mM sodium dithionite, and incubated in tightly capped quartz cuvettes or rubber-capped glass tubes for ~30 minutes. The samples were excited at 280 nm and the fluorescence emission spectra were taken in a FluoroMax-3 instrument (Jobin-Yvon, Horiba). High-pH fluorescence changes were corrected by using NATA (N-acetyl tryptophan amide) fluorescence. The denaturant unfolding data for were analyzed using standard two-state equation for equilibrium unfolding (Equation 1 in Chapter 1).[46] The pH titration data were fitted to the equation:

$$y = \frac{c_u + c_f[10^{n(pH-C_m)}]}{1+10^{n(pH-C_m)}} \qquad (1)$$

where, C_u and C_f are normalized fluorescence signals for the unfolded and refolded state, respectively, n the number of OH^- titrated, and C_m the pH-midpoint for the transition.

2.3.2 ***Stopped-flow Kinetic Measurements:***

Cyt *c* initially dissolved in pH 13 buffer in the presence or absence of 4 M GdnHCl was reduced with sodium dithionite to a final concentration of ~3 mM. The solution was then saturated with CO under 1 atm of the gas. The resulting carbonmonoxycyt *c* is unfolded (UCO) under the conditions used. Refolding was initiated by eight-fold dilution of the protein solution with the refolding buffer. The final protein concentration in the refolding mixture was in 5-50 μM range. For unfolding experiments, the native state of carbonmonoxycyt *c* (NCO) was prepared by manual dilution of the alkali-unfolded protein with the refolding buffer in 1:40 ratio. NCO was then unfolded by eight-fold dilution into the

unfolding buffer (0.1 M phosphate, 3 mM sodium dithionite, pH 7) containing different concentrations of GdnHCl. The final protein concentration in these experiments was ~4 μM.

In interrupted refolding experiments, the alkali-unfolded protein solution (pH 13, with or without 4 M GdnHCl) was diluted six-fold into the refolding buffer. Folding was allowed to proceed for variable time, t_i, before the solution was transferred to the appropriate unfolding medium. The amplitude of the unfolding reaction was analyzed as a function of t_i. A Bio-Logic SFM 400 instrument was used for these experiments.

2.3.3 ***Carbonmonoxide Binding to Ferrocyt c****:*

Cyt *c* (~2 μM) reduced with sodium dithionite in 10 mM CAPS buffer at pH 12.8 was kept in an air tight cuvette. The protein was titrated with saturated CO solution. The CO binding was measured using the absorbance difference between liganded and unliganded ferrocyt *c* at 414 nm. The CO concentration was evaluated from the binding curve of myoglobin. Binding constant for ferrocyt *c* was derived from the plot of absorbance difference between free and CO-bound cyt *c* against free CO concentration using,

$$\Delta A = \frac{1}{1+10^{n([CO]_{free}-K)}} \qquad (2)$$

where ΔA is the difference in absorbance, n is the number of binding sites and K is apparent dissociation constant.

2.3.4 ***Preparation of Flash-frozen Intermediate and EPR:***

Cyt *c* was dissolved in 45% glycerol at pH 13, reduced with sodium dithionite, and liganded with nitric oxide. The refolding buffer, consisting of 45% glycerol buffered with 50 mM phosphate at pH 6.6, was placed in quartz EPR tube, and cooled to -11^{o}C. The NO-liganded unfolded protein was rapidly diluted ten-fold into the pre-cooled refolding buffer, and transferred immediately into a bath of liquid nitrogen. The final protein concentration was ~125 μM. The control

sample was prepared by direct nitrosylation of the native protein at pH 7. Spectra were recorded at -150^{o}C using a JEOL (X-band) spectrometer with a microwave frequency of 9.18 GHz.

2.3.5 *Computational Method*:

For numeric calculation of the microscopic rate constants, k_j (j = IU', U'I, U'N and NU') in the scheme I ⇌ U' ⇌ N, a 3×3 matrix was set up from the three coupled linear differential equations for I, U' and N. The observable rate constants, λ_i (i = 1,2), computed by diagnolizing the rate matrix, are functions of k_js, the logarithm of which are assumed to have linear dependence on GdnHCl concentration.[86] The equilibrium unfolding transitions were calculated from the eigenvectors by assigning expected values of fluorescence to I, U and N in the range of 0-1.

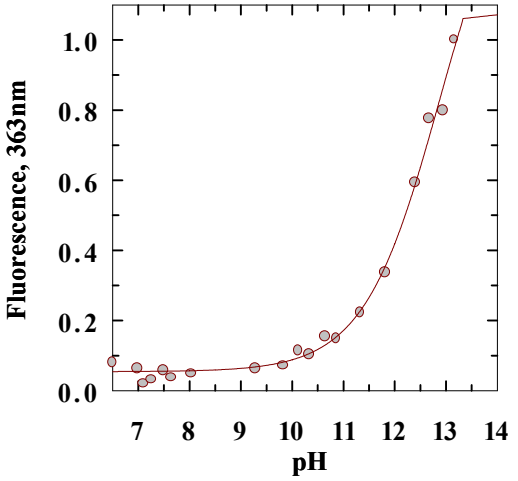


Figure 1. Alkaline pH induced unfolding of ferrocyt *c* in the presence of CO at 22 ºC. The solid line is fit to equation 1. The pH-midpoint for the titration is ~12.5.

2.4 Results

2.4.1 *Alkaline Carbonmonoxyferrocyt c*:

Ferrocyt *c* at neutral pH does not bind CO, but the denatured protein in either high concentration of GdnHCl or alkaline pH does. The protein unfolds in alkaline aqueous medium when carbonmonoxide is allowed to bind to its ferrous heme iron (Figure 1). As the pH value increases, the intrinsic fluorescence of tryptophan diminishes. The obtained fluorescence values for cyt

c were corrected for the pH effect using standard NATA fluorescence change. The pH-midpoint for the unfolding transition is ≈12.5. Unfolding of carbonmonoxycyt *c* (cyt-CO) is complete even in the aqueous medium, because the addition of the denaturant GdnHCl to the protein solution at pH 13 or higher does not produce any considerable change in the fluorescence emission by the lone tryptophan (W59).

2.4.2 *Folding and Unfolding Kinetics*:

2.4.2a. *The Reverse Denaturant Effect*: For a quantitative kinetic study, unfolded carbonmonoxycyt *c* was prepared by adding 1 atm CO to the protein solution held at pH 13 with or without 4 M GdnHCl, and the millisecond refolding rate constants were measured by stopped-flow dilution of the unfolded protein solution into refolding buffers containing different concentrations of GdnHCl at pH 7, 22 ^{o}C. Figure 2 reports single exponential fits of two representative refolding traces monitored by fluorescence. The observed refolding rate constant in the presence of 0.2 M GdnHCl (k_{obs} = 11.5 s^{-1}) is ten-fold slower than that for 2 M GdnHCl (k_{obs} = 120 s^{-1}). Both traces also show burst loss of fluorescence signals. Notably, at higher concentration of GdnHCl in the refolding milieu the burst loss is less.

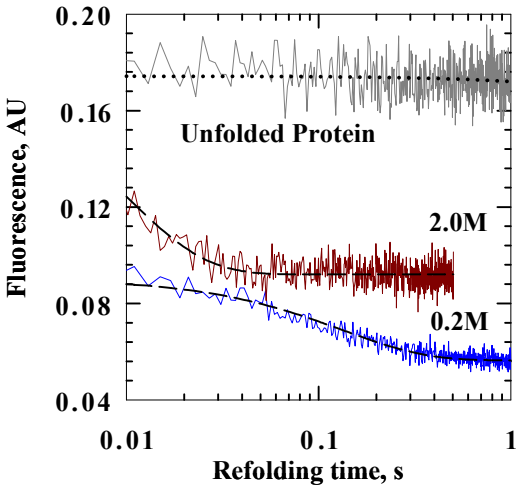


Figure 2. Representative 1-exponential refolding kinetic traces of carbonmonoxyferrocyt *c* obtained from stopped-flow experiments. Refolding shows the burst loss of fluorescence with reference to the unfolded signal. The lower the denaturant concentration in the refolding medium, the larger is the signal loss.

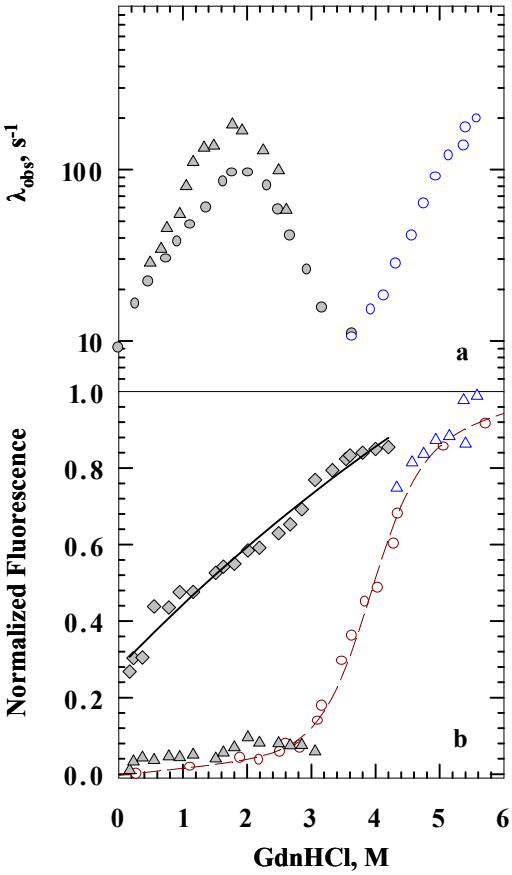


Figure 3. (a) Observed relaxation rate constants versus GdnHCl. (○) and (△) represent refolding rate constants when the alkali-unfolded CO-bound protein solution contained 0M and 4M GdnHCl, respectively. (○) represents unfolding rates. (b) Denaturant distribution of burst fluorescence. (△) and (◇) are baselines and amplitudes of refolding experiments and (△) are baselines of unfolding experiments. (○) is from equilibrium unfolding transition and the line represents two-state fit.[46]

The excellent quality of data allowed quantification of the folding–unfolding rate constants and the kinetic fluorescence amplitudes as a function of GdnHCl, pH 7 (Figure 3). Very interestingly, under strongly native-like conditions (<2 M GdnHCl) the refolding rate constant decreases linearly with decreasing concentration of the denaturant, producing a sharp inversion in the folding limb of the chevron (Figure 3a). This reverse denaturant effect does not depend on the content of GdnHCl in the initial unfolded protein solution at pH 13, although the refolding rates in the inversion zone are marginally faster when the denaturant is included. The accentuated inversion of the rate-denaturant profile provides an evidence for extreme misfolding.

2.4.2b. *Kinetic Signal Amplitudes*: The analysis of denaturant dependence of fluorescence signal amplitudes is presented in figure 3b. The plot depicts the fact that there is a considerable signal loss within the

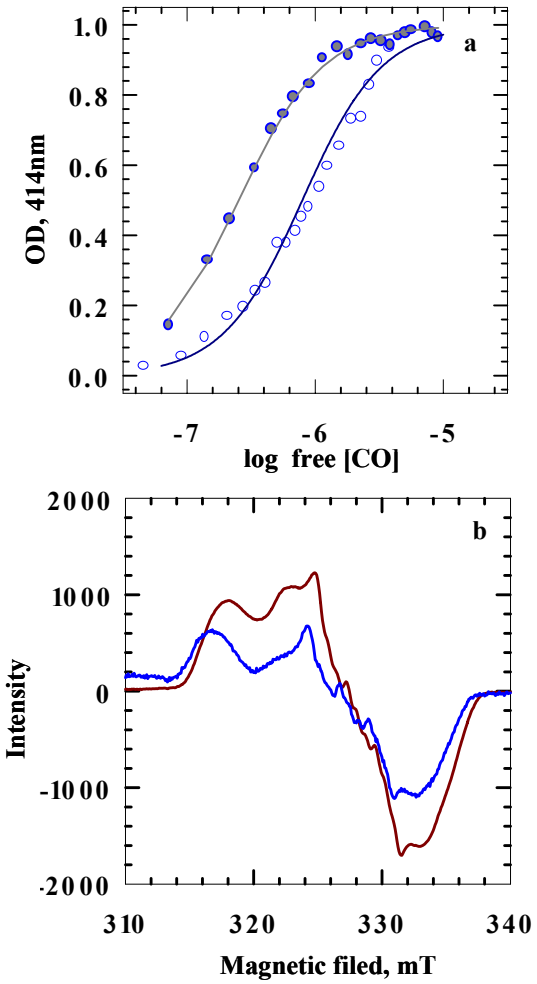


Figure 4. (a) CO titration of alkaline ferrocyt *c* with (●)and without (○) 3M GdnHCl. The solid lines are fits to equation 2. In each case the number of CO bound is unity. (b) EPR spectrum of the frozen nitrosylated intermediate (—) in 45% glyserol, pH 7. The control spectrum (—) is generated by direct nitrosylation of ferrocyt *c* at pH 7. Both spectra are characterized by the "triplet of triplet" hyperfine structure.

dead time of the instrument at the incipient stage of refolding, particularly at lower denaturant concentrations. The denaturant dependence of burst signal of unfolded protein describes the titration of abortive intermediate whose energetic stability is ~2 kcal mol^{-1}. The baselines of folding and unfolding kinetic traces fall exactly on the equilibrium denaturant transition curve.

2.4.3. *Binding of CO to Alkaline Ferrocyt c:*

To examine the protein dimerization or aggregation, titration of CO with alkaline form of ferrocyt *c* was carried out. The variation of absorbance signal between CO bound and unbound protein was used to derive the binding curve (Figure 4a). The data were fitted to the equation 2. Value of *n*, the number of CO bound, is unity in each case. Apparent dissociation constant,

K, is 1.64 and 22 μM in the presence and the absence of the denaturant, respectively.

2.4.4 EPR of Flash-frozen Intermediate:

Ligation of a single molecule of NO to the ferrous heme of cytochrome *c* can be evidently characterized by the typical "triplet of triplet" hyperfine structure.[41] The flash-frozen NO-bound alkaline ferrocyt *c* spectrum is quite similar to the directly nitrosylated ferrocyt *c* at pH 7 (NNO) (Figure 4b).

2.4.5 *Interrupted Refolding Kinetics:*

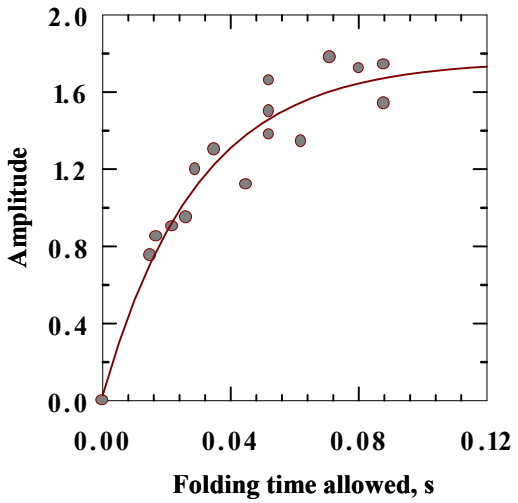


Figure 5. Amplitude of unfolding traces as a function of folding time. Folding was allowed in 0.5 M GdnHCl and the final unfolding was initiated by transferring the folding mixure to 5M GdnHCl. Single-exponential fit of the amplitude versus refolding time gives a rate constant of 31 s^{-1}. Errors in the values of amplitude is ~ 16%.

In these experiments, the unfolded protein is allowed to refold for a finite period of time, called aging time, in a native-like environment, and then transferred again into the unfolding condition.[94,95] If an unfolding reaction is initiated at some point of time while folding is in progress, then only those molecules that have already achieved the native structure will undergo unfolding, and the total fluorescence change observed for this unfolding process will be proportional to the amount of native molecules that were present at the time of interruption. The objective of this interruption during the course of refolding is to resolve the rate of formation of N in the depicted folding scheme I ⇌ U' ⇌ N. A plot of the total

fluorescence change for unfolding as a function of time of folding interruption is

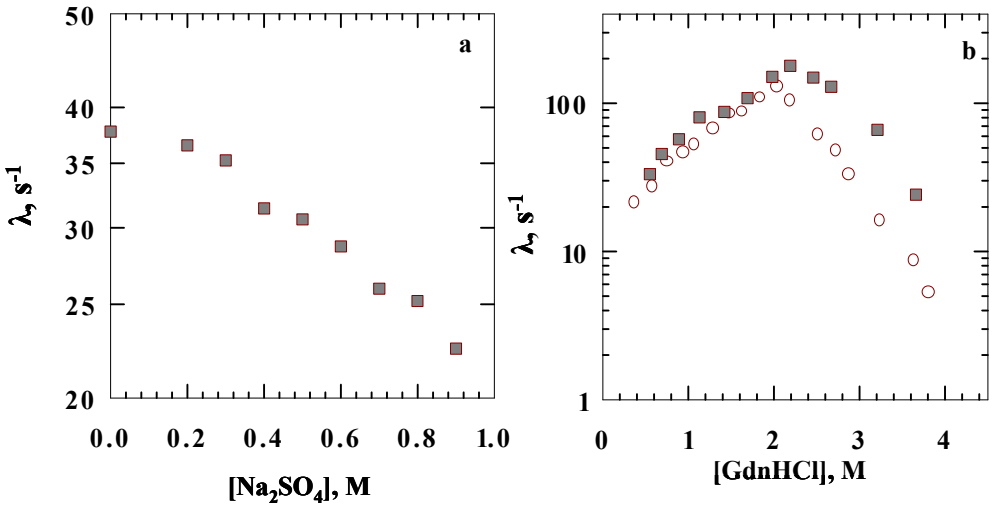


Figure 6. Sodium sulfate-induced stabilization of the abortive intermediate. (a) Variation of the rate of folding (■) in the presence of 0.5M GdnHCl as a function of Na_2SO_4. Stabilization of I retards the folding rate. (b) Denaturant dependence of folding in the absence (○) and the presence (■) of 0.4M Na_2SO_4. The same alkali-unfolded CO-bound protein solution containing 4M GdnHCl was used for both sets of refolding experiments.

essentially a plot of native molecules formed as a function of time of folding, and it provides the rate constant for the formation of N in a straightforward manner. The alkali unfolded cyt-CO was allowed to refold in the presence of 0.5 M GdnHCl for different aging times, t_i, and then driven back to the initial unfolded state. The amplitude of the unfolding reaction plotted as a function of t_i produces an exponential time course described by a rate constant, $k \approx 31$ (±2) sec^{-1} (Figure 5), which is equal to the rate constant that is obtained in direct refolding experiments (Figure 3a).

2.4.6 *Refolding in Presence of Sodium Sulfate:*

Ionic salts such as sodium sulfate are known to stabilize intermediate species. They will retard the conversion of I to N via U' if I is a dead-end intermediate. But salts are expected to stabilize the native state as well. This also happens, and is detectable in the range of GdnHCl concentrations where the folding mechanism is essentially two-state, U' ⇌ N. As moderate to strongly native-like conditions are approached and the mechanism shifts from U' ⇌ N to I ⇌ U' ⇌ N, the dead-end intermediate accumulates, and the refolding kinetics are not affected by the stabilization of native state by the salt. This is because the I → U' conversion rate-limits the formation of native state according to the I ⇌ U' ⇌ N mechanism. Figure 6a illustrates the slowing down of folding rates as the concentration of sodium sulfate is increased. The chevron data for refolding (Figure 6b) in the presence and the absence of 0.4 M Na_2SO_4 exemplifies the effect of salt on the proposed dead-end intermediate. In the presence of denaturing concentrations of GdnHCl (>2 M), addition of Na_2SO_4 accelerates folding. Below 2 M GdnHCl the presence of Na_2SO_4 does not considerably change the refolding rate.

2.4.7 *Simulation of the Abortive Pathway:*

The experimental data were simulated by numeric calculation of the denaturant dependences of the eigenvalues, λ_i, and the concentrations of each of the species corresponding to the I ⇌ U' ⇌ N mechanism. Figure 7a shows one of the solutions. The two chevrons resulting from the two eigenvalues (λ_1 and λ_2) produce a saddle point region (~2 M GdnHCl). As progressively refolding conditions are achieved (<2 M GdnHCl), I is increasingly stabilized and hence accumulates. Consequently, the refolding limb of the λ_2 chevron inverts instead of continuing linearly to the ordinate, and no kinetic phase is observed at longer times. GdnHCl dependences of the four microscopic rate constants that provide

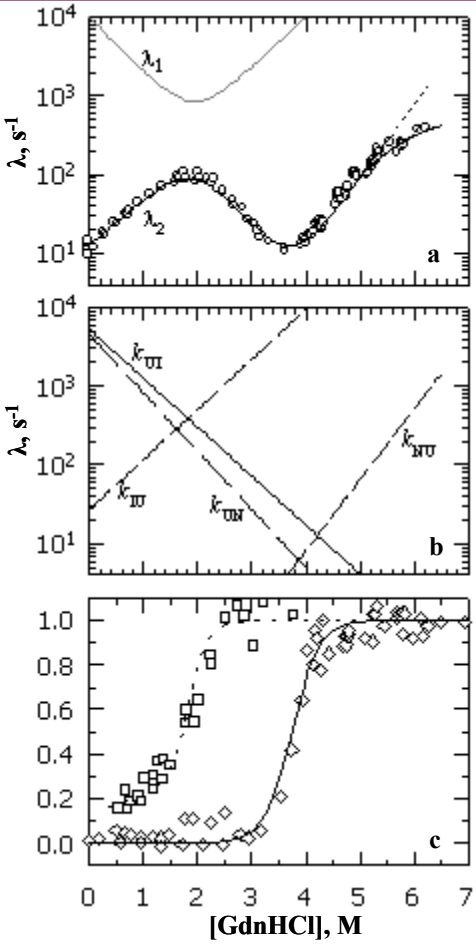


Figure 7. (a) Variation of eigenvalues (λ_1 and λ_2) as a function of the denaturant. The inverted dependence of λ_2 on GdnHCl in the folding limb suggests the unfolding of I to U. The reason for the slight rollover in the unfolding limb of the chevron is unknown. (b) The k_j-denaturant space corresponding to denaturant dependence of the two eigenvalues. (c) GdnHCl induced transitions of I (□) and N (◊) predicted according to the dead-end mechanism.

the eigenvalue solution are shown in figure 7b. Values of these rate constants in water, k_j^o (s^{-1}), and the kinetic *m*-values, $m_j^{\ddagger}$ (kcal mol^{-1} M^{-1}), are: $k^o_{U'I}$ = 5250, $m^{\ddagger}_{U'I}$ = -0.83, $k^o_{IU'}$ = 27, $m^{\ddagger}_{IU'}$ = 0.85, $k_{U'N}$ = 4550, $m^{\ddagger}_{U'N}$ = -0.99, $k^o_{NU'}$ = 0.002, and $m^{\ddagger}_{NU'}$ = 1.2. The saddle region contains the transition midpoint of I. The simulated denaturant distribution of the burst phase has already been shown in figure 7c. This simulation also is based on the scheme I ⇌ U' ⇌ N. The time constants for U' → I (190 μs) and U' → N (220 μs) predicted by the simulation appear quite small. In none of the previous studies a time as fast as 220 μs was obtained for the U' → N process. The reason for this could be the use of linear *k*-denaturant relationship. Stopped-flow chevrons of ferrocyt *c* and cyt-CO under different conditions of initial and/or final pH, except for the present ones, invariably show pronounced

curvatures in the folding limbs (Figure 8). Those results bind one to use non-linear or quadratic *k*-denaturant interpretations, and the folding time in water is found in the 1-2 ms range. Due to the inversion of the chevron in the present case (Figure 3a) the curvature is not detected, so the linear *k*-denaturant relation was chosen to simulate the data. Thus, the simulated values for the rate constants for U' → I and U' → N, especially the latter, are substantially larger.

The analytical solution of the I ⇌ U' ⇌ N reaction in water was calculated (data not shown). The obtained values are: $\tau_{(U' \rightarrow I)} \sim 15$ μs and $\tau_{(U' \rightarrow N)} \sim 33$ μs, with $K_{(IU'/U'I)} = 4.1 \times 10^{-4}$.

2.5 Discussion

2.5.1 *Chevron Inversion and Protein Misfolding*:

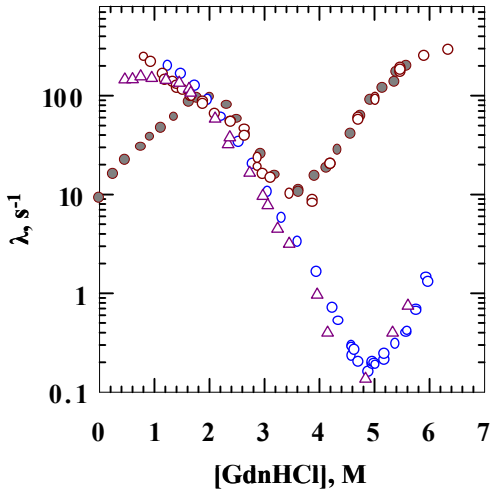


Figure 8. Kinetic chevrons for ferrocyt *c* and cyt-CO. Folding and unfolding of cyt-CO, pH 12.9 → pH 7 (●) and pH 7 → pH 7 (○). Folding and unfolding of ferrocyt *c*, pH 12.9 → pH 7 (△) and pH 7 → pH 7 (○). All the experiments were done at 22ºC. (Ref. *Chapter* 1)

Ferrocyt *c* is completely unfolded in alkaline medium when it binds to extrinsically added CO ligand (Figure 1). When the unfolded carbonmonoxyferrocyt *c* is allowed to refold from pH 12.9 to pH 7, the rate of refolding decreases linearly with decreasing concentration of GdnHCl (Figure 3a). This apparent inversion of folding chevron exemplifies possible misfolding of the protein during the refolding process. The misfolded species is increasingly stabilized and hence accumulates

as the denaturant concentration in the refolding medium falls lower. The misfolded product cannot, however, achieve forward folding because of the preponderance of non-native interactions. In this sense, it is a dead-end intermediate, and it needs to undergo a large scale unfolding to resemble the initial nascent state, both structurally and energetically, in order to have a chance to refold correctly. The gradient of the rate-denaturant plot, $m^{\ddagger}$, is generally calculated from $\ln k_{obs} = \ln k^{o} + (m^{\ddagger} / RT)\,[D]$, where k_{obs} is the observed rate constant, $\ln k^{o}$ is the rate constant in the absence of denaturant, and $[D]$ is the concentration of denaturant.[97] The positive slope of the inverted part of the refolding limb, $m^{\ddagger}_{f}$ (≈ 0.85 kcal mol^{-1}), is pretty close to the slope defined by the unfolding limb of the main chevron ($m^{\ddagger}_{u} \approx 1$ kcal mol^{-1}). This comparison reflects the enormity of the unfolding of the dead-end species. Since the intermediate unfolds in milliseconds, its rate of formation must be faster. Hence, the formation of the intermediate is expected to occur within the burst-associated submillisecond phase.

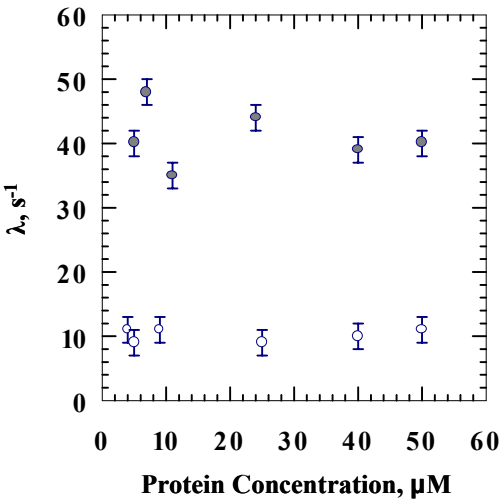


Figure 9. Refolding rate of different concentrations of carbonmonoxy ferrocyt *c* from pH 12.8 to pH 7 in the presence (●) and the absence (○) of GdnHCl.

2.5.2 *Aggregation or Misfolding*:

It is necessary to show that the observed misfolding is not related to protein dimerization or aggregation at any stage, since short-lived aggregates may cause a downward curvature in the folding limb of the chevron plot.[98] Several experiments were done to dispel the idea of formation of aggregation during refolding.

By titrating the alkaline form of ferrocyt *c* with CO in the presence or absence of GdnHCl, it is found that only one molecule of CO binds to the protein molecule (Figure 4a). It suggests that the alkaline unfolded form of the protein is monomeric. The possibility of aggregation during the course of refolding can be examined by performing the refolding reaction with different concentrations of protein. Figure 9 shows that the refolding rate constant is invariant to the protein concentration, confirming that the structural attribute of the intermediate whose accumulation produces the inversion in the folding chevron is unlikely to be an aggregate. The EPR spectrum of the flash-frozen intermediate prepared from the nitrosyl derivative of the unfolded protein shows neither dipolar broadening nor a significant reduction in intensity (Figure 4b), implying no close proximity of the unpaired electron spins due to the NO ligands bound the heme moieties of two different chains. The intermediate rather exhibits hyperfine pattern characteristics of monomeric nitrosyl cytochrome *c*, [99] and demonstrates that only one NO binds to the ferrous heme. Also, the TOF-mass spectra of the alkali unfolded protein (Figure 2c, Chapter 1) evidences against protein modification. These results do not suggest protein dimerization, aggregation or chemical modification at different stages of folding.

2.5.3. *Rate of Formation of Native Molecules*:

Interrupting refolding protein at different time and then unfolding into the same final unfolding condition will hint at the rate of formation of native structure, when intermediates are expected to involve in the refolding system. The fluorescence amplitudes for unfolding of interrupted refolding molecules at different time intervals provide a rate constant $k \approx 31$ sec^{-1} (Figure 5). This number is nearly equal to the rate constant obtained in direct refolding experiments (Figure 3a), suggesting that the formation of the native state is rate limited by the unfolding of I to U'.

2.5.4. *Reverse Sodium Sulfate Effect*:

The refolding rates of the alkaline form of carbonmonoxyferrocyt *c* to neutral pH in presence of sodium sulfate (Figure 6b) specify two important points. (1) The folding rates are accelerated by the addition of Na_2SO_4 in the presence of denaturing concentrations of GdnHCl (>2 M), where the folding mechanism is U' ⇌ N. The acceleration of folding rate indicates the stabilization of native state (N) by Na_2SO_4. Under these conditions, folding is a two-state process and is directly controlled by U' → N conversion. (2) In low denaturing concentrations of GdnHCl (<2 M), intermediate I is more populated and the folding mechanism shifts to I ⇌ U' ⇌ N. Now Na_2SO_4 stabilizes the I, hence the unfolding rate of I to U' is decelerated. Since, I → U' conversion rate-limits the overall folding, the stabilization of N does not affect the folding rate so long as the I ⇌ U' ⇌ N mechanism operates.

2.5.5 *Simulation Studies*:

The eigenvalues simulation (λ_1 and λ_2) of the rate-denaturant data based on the folding mechanism I ⇌ U' ⇌ N produces a saddle point region at ~2 M GdnHCl. Under highly refolding conditions (<2 M GdnHCl), I is increasingly stabilized and hence accumulates. Consequently, the refolding limb of the λ_2 chevron inverts. The positive slope of the inverted region of the folding limb is due to the unfolding of I to U'. At higher concentrations of the denaturant, the refolding does not involve any intermediate, and the mechanism is effectively U' ⇌ N. Also the simulation of the denaturant distribution of the burst amplitude according to the I ⇌ U' ⇌ N scheme shows the same transition midpoint for I (~ 2M GdnHCl, Figure 7c).

2.5.6. *Dead-end Intermediate*:

The phenomenon of deviation from the predicted linear dependence of logarithm of the folding rate constant on denaturant can be both qualitatively

interrupted and quantitatively modeled in terms of on-pathway as well as off-pathway placement of an intermediate.[95,101] Whether the intermediates are on the pathway or are merely side products is one of the longest-running controversies.[102] The experimental record of a fundamental kinetic criterion, namely, inversion of the denaturant dependence of the observed folding rate constant reveals the off-pathway dead-end mechanism. This direct observation has obviated the need for testing the other off-pathway kinetic criteria based on Φ value analysis,[103] pulse-chase competition experiments[104] and the evaluation by experiments of various microscopic rate constants.[105]

The protein can nearly completely misfold to abortive states in the test tube as some do in the cell.[91] Clearly, not all kinetic intermediates play a productive role. Even so, the vulnerability of the abortive product especially in the cytosolic milieu adds on to the inefficiency of folding. It is not surprising that the cell engages a variety of molecular chaperones to achieve folding efficiency.[92,93] The arrival of a protein chain at the dead-end is possibly by a consequence of unsuccessful search for the correct transition state for folding.

Dead-end intermediates associated with intrinsically slow steps on protein folding have been reported for the oxidative folding pathway of BPTI,[106] and for an intermediate with non-native prolyl isomers in RNAse T1.[107] In carbonmonoxy ferrocyt *c*, such a possibility can be considered by examining the influence of the axial non-native heme ligation at different stages of folding of cytochrome *c*.

2.5.7. *Oxidation State Dependent Folding and Misfolding*:

Refolding of oxidized cytochrome *c* (ferricyt *c*) at neutral pH has been shown to proceed through an intermediate, I_{NC}, in which the N- and C- terminal helices dock against each other in an orthogonal geometry. There have, however, been conflicting viewpoints concerning the label of I_{NC}, whether off-pathway or on-pathway. The native-state axial ligation of M80 to the oxidized heme is

replaced by H26 and H33 under denaturing conditions.[108] Even though non-native, the Fe^{3+}–H26 and Fe^{3+}–H33 contacts are so stable that the dissociation rates of these two ligands from the iron are far smaller than the protein folding rate. A quantitative kinetic model of these processes has been described earlier.[28,30] As a result, the ferricyt polypeptide collapses with persisting Fe^{3+}–H26 and Fe^{3+}–H33 contacts. This act of trapping of the non-native histidines displaces the histidine-resident chain segments from the normal distal side of the heme to the proximal side of M80. The misconfiguration arrests folding because the frustrated chain organization stays in a kinetically frozen intermediate state, I_{NC}, until thermal dissociation of non-native histidines from the heme iron gives way to the ligation of M80, and hence polypeptide reconfiguration. In this picture, an unfolding event disperses the incorrect chain organization in order to facilitate correct folding, so the intermediate could be viewed as a trapped misfolded intermediate.[23,28,30,31] On the other hand, the intermediate may be described, perhaps inappropriately, just at the level of the relevant N- and C-terminal helices without paying any attention to wrong chain organizations in other parts of the molecule.[82] I_{NC} then carries the "on-pathway" label.[49] The problem of kinetic trap or chain misfolding in ferricyt *c* can be nearly fully eliminated by allowing refolding at low pH or in the presence of an extrinsic heme ligand that blocks non-native histidine ligation. Folding kinetics are then simply modeled as a two-state reaction with no population of I_{NC}, and the burst phase is attributed to the initial chain contraction.[24]

These conflicts do not surface in the case of refolding of the virtuous ferrocyt *c* under normal conditions of refolding, since non-native heme ligands do not interfere with inherent accelerated folding of this redox state of cyt *c*. It has been shown earlier[30] that 45% of the unfolded ferrocyt *c* molecules have no equilibrium occupancy at the sixth coordination site of the heme, so ligands do

nothing to their folding. For the remaining 55%, a liganded form will enter the folding pathway only if the rate of dissociation of the ligand from the heme iron is slower than the folding rate.[109] The dissociation time constants for all four intrachain ligands of ferrocytochrome *c* are smaller than the polypeptide refolding time.[30] Thus, heme-ligation in the unfolded-state of ferrocytochrome *c* does not interfere with refolding. More detailed studies have indeed shown that stopped-flow observable millisecond kinetics for ferrocyt *c* are fast, simple, and essentially monophasic.[28,110] Reassuringly, ferrocyt *c* under physiological conditions is a two-state fast-folder.[31,38]

2.5.8. *Misfolding is not Directly Related toHeme Ligation*:

At neutral pH, the Fe^{2+}–M80 $\rightarrow$ Fe^{2+}–CO engineering lowers the stability of ferrocyt *c* from 19.3(±0.5) to 11.65(±1.13) kcal mol^{-1} without altering the rate of refolding (Chapter 1), but with a substantial increase in the unfolding rate of the protein.[28] Folding of ferrocyt *c* is thus unaffected by the extrinsic non-native ligand; the protein simply folds completely in a two-state manner with the Fe^{2+}–CO bond intact, and slow thermal dissociation of CO (~0.02 s^{-1}, 22 °C) eventually drives the Fe^{2+}–CO $\rightarrow$ Fe^{2+}–M80 reaction.[110] However, this parity in the folding of ferrocyt *c* and cyt-CO is lost when the refolding reaction is allowed starting from pH 12.9(±0.1) to a final pH of 7 in the presence of subdenaturing concentrations of GdnHCl. Figure 8 illustrates these behavior. There are three major points. (1) The denaturant dependences of refolding rates for ferrocyt *c* and cyt-CO observed for the pH 7 $\rightarrow$ 7 folding reaction are very similar, indicating that heme ligands, irrespective of native, non-native, extrinsic or intrapolypeptide ligands, do not influence folding substantially. (2) The disparity between ferrocyt *c* and cyt-CO for the pH 12.9 $\rightarrow$ 7 refolding, more so in the subdenaturing region, suggests that the extrinsic CO ligand causes misfolding. (3) The chevron features for pH 7 $\rightarrow$ 7 and pH 12.9 $\rightarrow$ 7 refolding of ferrocyt *c* are virtually identical,

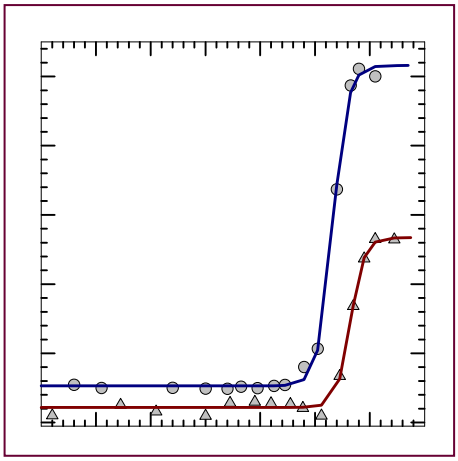
except for very minor changes under strongly refolding conditions (Figure 8), suggesting that the initial pH of the unfolded protein solution makes no significant difference. It then follows that polypeptide misfolding is conditional to the combination of the initial alkaline pH and the Fe^{2+}–CO interaction. The exact interaction(s) that facilitate severe misfolding of alkaline cyt-CO is hard to figure out with the available data. It is possible to invoke a change in the average chain configuration of cyt-CO in the initial alkaline medium, or a charge-dipole interaction between the CO electric dipole (0.112 Debye) and one of a number of possible charged groups in the proximity. But these could be mere speculations. More remains to be done to obtain a mechanistic description of the observed misfolding at the level of intramolecular interactions.

2.6. Conclusion

Refolding of carbonmonoxycyt *c* from alkaline pH to neutral pH occurs through an intermediate which leads to chevron inversion in rate-denaturant profile. At low denaturant concentrations the formation of the intermediate is facilitated, and at higher concentrations of denaturant the intermediate is destabilized. The intermediate has been characterized as an off-pathway product. Occurrence of dimers or aggregates of the initial unfolded state as well as in the course of refolding has not been found. The carbonmonoxide ligation dispels the possible misligation at the heme-iron during folding. The interrupted refolding experiments and the eigenvalues simulation studies exactly fit the refolding data into the $I \rightleftharpoons U' \rightleftharpoons N$ model for highly native-like conditions (<2 M GdnHCl), The model predicts that the protein must unfold from its misfolded intermediate to reach its native structure. However, at higher concentrations of GdnHCl (>2 M GdnHCl) the protein folds through simple two-state $U' \rightleftharpoons N$ model. Even though the structural description of the intermediate is not well exemplified, the results provide compelling evidence for the formation of a dead-end intermediate.

Chapter 3

The Alkali Molten Globule State of Ferricytochrome *c*



3.1 Abstract

Molten globule states of globular proteins and their relationship with folding intermediates has been stimulating subject area in the experimental studies of protein folding. This chapter presents the basic structural properties and a thermodynamic description of a previously unknown alkali molten globule state of horse ferricytochrome *c*. Both sodium and guanidinium cations stabilize the alkali-denatured state at pH 13, presumably by a charge screening mechanism. The Na^+-stabilized conformation at high pH (B state) clearly meets with the molecular organizational definition of the generic molten globule state and can be considered as a complementary cation-stabilized state to the A state of cyt *c*. The B state exhibits highly cooperative reversible thermal unfolding transitions monitored by both near and far-UV CD. Analyses of these transitions show substantial heat capacity change, suggesting that the hydrophobic effect contributes considerably to its energetic stability. At low salt concentration where molten globules are less stable, the B state undergoes reversible cold denaturation.

3.2 Introduction

Detection and characterization of folding intermediates, by both kinetic and equilibrium methods have been pressing issues in protein folding studies. The equilibrium molten globule (MG) state classically defined as "a compact and largely mobile molecular state containing native-like secondary structure and hydrodynamic radius, but without rigid tertiary structure" is now considered a distinct thermodynamic state of proteins.[111-113] Several reports support the molten globule state as a possible late folding intermediate, [112,114,115] and the study of the MG state may provide a better understanding of the mechanisms that stabilize the native protein. It is difficult to study them under normal folding conditions at

neutral pH because of their transient lifetime. For structural studies, they are stabilized and populated either at extremes of pH by the use of added counterions or intermediate denaturant concentrations.

The MG state is also important for protein function in the living cell.[111,116] In recent years, a class of intrinsically disordered proteins that resemble the MG state and lack tertiary structure has been studied in some detail.[117-121] These proteins, the abundance of which increases with evolution and complexity of organisms, are involved in cell signaling and regulatory function by protein–DNA and protein-protein interactions.[122-126] In general, the unstructured regions of the MG-like intrinsically disordered proteins undergo structured-to-folded transitions when bound to target sites. Another most recent example of their implication in protein function is the demonstration that a pH-dependent molten globule transition is required for the activity of the steroidogenic acute regulatory protein that stimulates steroid biosynthesis.[127] The MG state has also served as a paradigm in studies of temperature function of protein stability, since pH-destabilized states in the presence of salt are amenable to cold denaturation in the experimentally accessible temperatures.[128,129] Thus, the study of the MG state continues to be a thrust area in experimental protein research.

Evidently, molten globules have been documented for a large set of proteins.[111-113] Test-tube studies of ion-stabilized molten globules require one of the following transitions: $U_A \rightarrow A$ and $U_B \rightarrow B$, where U_A and U_B are acid and base-denatured proteins, respectively, and A and B are the corresponding MG states. In most cases, the A state has been studied in the presence of anions; the complementary cation-stabilized B state has received meager attention. In fact, the B state has been described for only a very few proteins, including β-lactamase,[132] and barstar.[133] More specifically, the B state of ferricyt *c* remained unnoticed, even though the first conclusive evidence for the existence of the MG

state was based on some detailed experiments with the A state of ferricyt *c*.[128,131,134-139]

More recently a set of experiments have been carried out to characterize the B state of cyt *c*. In those experiments,[140] carbonmonoxy derivative of the reduced form of cyt *c* at pH 13 was used to achieve the $U_B \rightarrow B$ transition in the presence of Na^+ ions. The experiments were difficult because ferrocyt *c* hardly denatures even at the extreme of basic pH, and therefore the CO adduct was used. But in the case of ferricyt *c*, where the U_B state is directly prepared at pH 13, the $U_B \rightarrow B$ transition is achieved in the presence of NaCl. The B state of ferricyt *c* exhibits all of the basic physical and structural properties known for the A state. Interestingly, in the presence of low concentration of salt the B state shows a cold denaturation transition, as does the A state.[128] Analyses of the thermal transitions suggest that the hydrophobic effect might contribute substantially toward the energetic stability of the B state.

3.3 Materials and Methods

3.3.1 *Alkali Denaturation in the Presence of NaCl*:

Solutions of 0, 0.02, 0.5 and 1.5 M NaCl were prepared in an aqueous medium containing 10 mM each of Tris, disodium hydrogen phosphate, and CAPS (3-[cyclohexylamino]-1-propane sulfonic acid). Each of the NaCl solutions was titrated to different pH values in the pH 7 – pH 13 range by the use of minimal volume of NaOH. Cyt *c* was added to these solutions to obtain samples each containing either 9 μM protein for fluorescence or ~20 μM for Far UV-CD. The dilutions did not upset the initial salt concentrations greatly. Samples were incubated for ~30 min at 22 ^oC. Fluorescence emission spectra (excited at 280 nm) were recorded in a FluoroMax-3 instrument (Jobin-Yvon, Horiba) and the

emission maximum at 343 nm was considered for analysis. Ellipticity changes were measured with JASCO, J-810 spectrometer.

3.3.2. *Titrations of NaCl, GdnHCl and Urea with Alkali-denatured Cyt c*:

Titrations of cyt *c* with NaCl in the 0-2.5 M range were carried out at pH 13, 22^{o}C, and were monitored by both fluorescence and Far UV-CD. Protein concentrations were 9 μM and 20 μM for the two probes, respectively. Near UV-CD was done with higher protein concentration (~ 40 μM). GdnHCl and urea induced equilibrium unfolding of cyt *c* held at pH 13 was followed by fluorescence alone. CD measurements were not done because the signals are obscured with GdnHCl and urea at high pH.

Activity of GdnHCl from its molar concentration was calculated by using the following polynomial function given by Aune and Tanford: [185]

$$\log a = -0.5191 + 1.4839 \log C - 0.2562(\log C)^2 + 0.5884(\log C)^3$$

where, *a* is activity and *c* the molar concentration.

3.3.3 *Thermal Unfolding Experiments*:

These experiments involved temperature-dependent CD measurements at both aromatic and peptide absorption bands (282 nm and 222 nm, respectively). Cyt *c* concentrations were 35 μM and 70 μM for use at 222 nm and 282 nm, respectively. The protein was held at pH 13 in the presence of variable NaCl concentration in the 0-2 M range. The Peltier-controlled heating and cooling rate was 0.5 ^{o}C /min.

3.4 Results

3.4.1. *Na^+ Induced Formation of Alkali Molten Globule State*:

Figure 1(a) indicates that the large-scale denaturation of cyt *c* that occurs at pH 13 results in a very substantial loss of tertiary structure as judged by the recovery of nearly all of the unfolded-state fluorescence. Interestingly, the CD_{222}

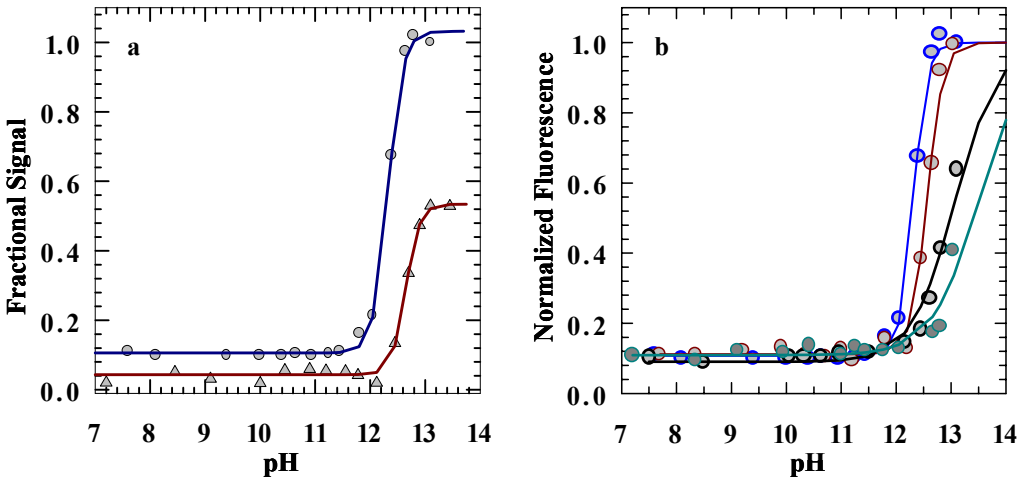


Figure 1. pH titration of ferricytochrome *c*. (a) Alkaline pH induced titration at 22ºC monitored by fluorescence (○) and peptide CD (△) showing the accumulation of an equilibrium intermediate at the extreme pH. (b) The same transition in different concentrations of sodium chloride at 22º C: without salt (○), 0.02M salt (○), 0.5M salt (○), and 1.5M salt (○). The solid lines are fits according to equation 1.

monitored transition not only shifts to a distinctly higher pH, but also indicates that the protein retains at least 40% of the native-state helical structure even at pH 13.5. The base-denatured state (pH ≥ 13) with these structural properties is referred to as the U_B state. But the observed non-coincidence of the pH transitions monitored by fluorescence and CD_{222} nm, a result consistent with an earlier report of the response of ferricyt *c* to alkali exposure, [35] provides the classic evidence for the presence of an equilibrium intermediate.[141] This base-denatured intermediate is labeled I_{B1}. The fact that the U_B state possesses nearly half of the native-state helical structure but is dimensionally expanded, sets the stage for the molten globule study. In fact, the early discovery of the molten globule state of proteins was based on the observation that the acid-pH form of α-lactalbumin[142] structurally resembles an intermediate form of the protein which populates in guanidine hydrochloride induced transition at neutral pH.[143] In general,

intermediates may not accumulate in denaturant-induced transitions, but it has been shown for a sizable set of proteins that extremes of pH can transform them into states possessing optical parameters similar to those for often elusive denaturant-induced intermediates.[111] The alkali-pH induced transitions of ferricyt *c* (Figure 1a) thus indicate that the U_B state could possibly be stabilized and transformed by cations to the alkali molten globule state, B, for structural characterization. This possibility is readily tested. Figure 1b shows that the fluorescence-monitored alkali-denaturation transition of cyt *c* shifts toward higher pH with increments of Na^+. Whereas the titration is apparently complete at pH 13 in the saltless medium, just about 40% molecules are denatured in the presence of 1.5 M NaCl. The fit to the data by the use of the equation:

$$y = \frac{c_u + c_f[10^{n(pH-C_m)}]}{1+10^{n(pH-C_m)}} \qquad (1)$$

where, C_u and C_f are normalized fluorescence signals for the unfolded and refolded state, respectively, n the number of OH^- titrated, and C_m the pH-midpoint for the transition, indicates titration of roughly three OH^-. This suggests that cations increase the stability of the protein under denaturing alkaline conditions. Ion-stabilized denatured states generally fall in the category of the MG state.[111, 123] The following experiments, show that ferricyt *c* at extreme alkaline pH and high NaCl indeed exists as a state that is not just a mixture of the native and the unfolded state, but is distinguishable as a molten globule because it displays a combination of properties of the two states.

3.4.2. *Characterization of B-state of Ferricyt c*:

The MG state is nearly as compact as the native state.[111] Na^+-driven molecular compaction of cyt *c* is easily verified by fluorescence emission due to the lone tryptophan, since compact states of cyt *c* are less fluorescent due to excitation energy transfer from W59 to the heme.[40, 136] In highly alkaline medium though, the four tyrosine residues of cyt *c* are also expected to contribute to the

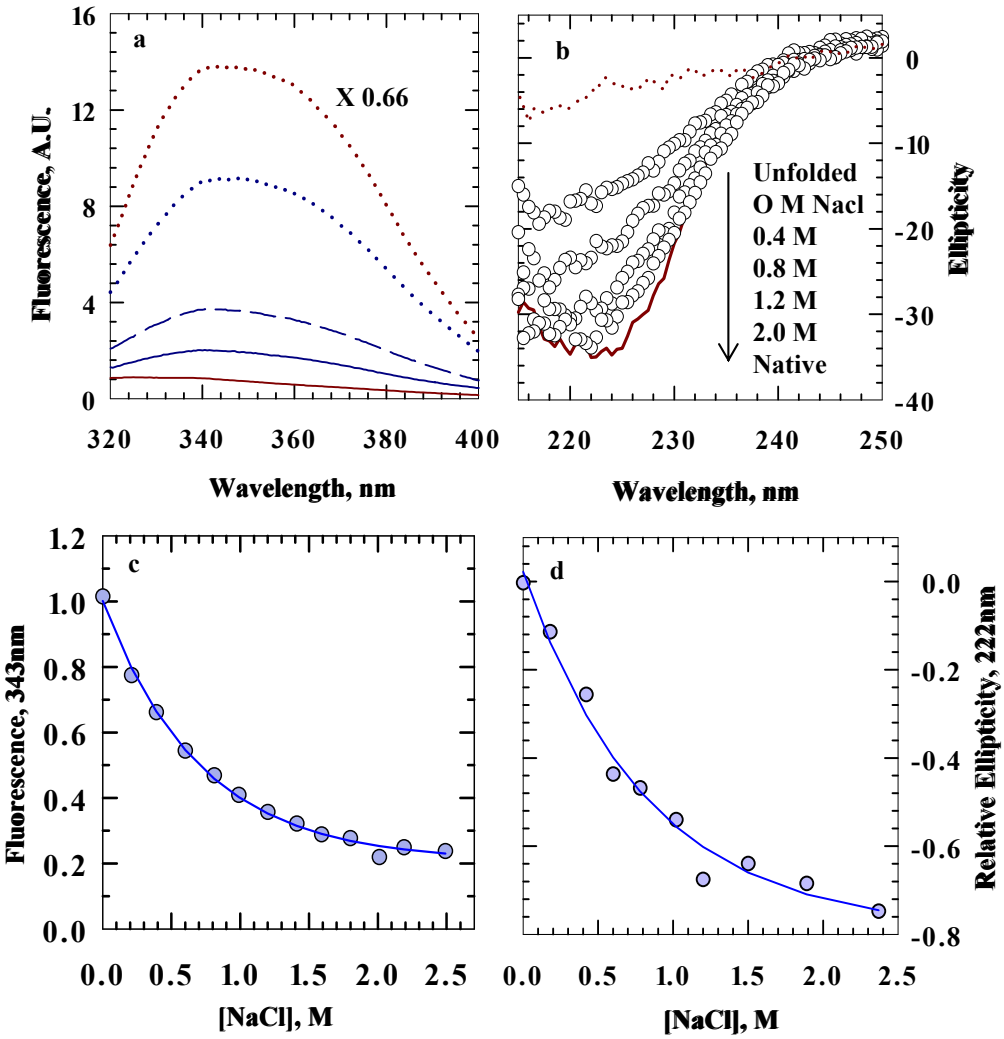


Figure 2. Formation of salt induced molten globule state, B from base-denatured U_B state. (a) Fluorescence spectra of different states of cyt *c*. cyt *c* at pH 7 (—), at pH 13 without salt (••••), at pH 13 in 1M NaCl (- - -), at pH 13 in 2M salt (—) , at pH 7 in 5M GdnHCl (••••). (b) Far UV-CD relative ellipticity of cyt *c*, native at pH 7 (—), unfolded at pH 7 in 5M GdnHCl (••••), at pH 13 with different salt concentrations (○). (c) Relative fluorescence quenching of W59 as a function of salt concentration at pH 13. (d) Relative change in peptide ellipticity as a function of salt concetration at pH 13.

fluorescence of the protein, since the excitation wavelength employed to probe

the tryptophan (280 nm) serves to excite the tyrosinate side-chain as well. Similarly, the emission maximum due to the W59 fluorescence closely coincides with that for the tyrosinates. Notably, the fluorescence intensity of the denatured protein at pH 13 is at least two fold less than at pH 7, possibly due to either quenching by hydroxyl ions or ionization of the phenolic hydroxyl group of tyrosine. In fact, alkaline quenching of tryptophan and tyrosine fluorescence has been known for a long time. In spite of these two effects at alkaline pH, the total observed fluorescence could still be used as a marker of molecular compaction of alkaline cyt *c*, because the present work relies on quenching of fluorescence, presumably because of increasing proximity of the fluorophore and the heme as the Na^+ concentration in the medium increases. Also, because all experiments are performed at pH 13, the effect of alkali on the fluorescence is expected to be uniform for all samples.

The emission observed for the alkali-denatured cyt *c* held at pH 13 (U_B) decreases when 2 M NaCl is included (Figure 2a), indicating the formation of a stable state (B). The native-state spectrum at pH 7 is of course fully quenched. Figure 2a also shows that the U_B state considered here is not as expanded as the fully unfolded state, since the fluorescence of U_B increases considerably in the presence of 5 M GdnHCl. This is not unexpected, since the pH-denatured states in general are partially compact. Similarly, the B-state undergoes molecular expansion in the presence of GdnHCl, indicating globular unfolding of the molten globule state.

The titration curve of Figure 2c shows NaCl-induced molecular compaction of U_B. The data, normalized with respect to the fluorescence of the initial U_B state and fitted to an expression similar to equation 1, indicate binding of one or two Na^+ ions. A value of 1.75 for the difference in the number of cations bound between the U_B state and the B state was also estimated from the

relation: $\Delta n = \partial \ln K_{app} / \partial \ln a_x$, where K_{app}, is the observed equilibrium constant for the salt-induced transition and a_x is the activity of the ligand.

Molten globule states have substantial content of secondary structure, often comparable to that in the native states.[111,144] The CD spectra in figure 2b show the secondary structure content in the relevant protein states. At 222nm, the U_B state has considerable value of ellipticity than that of the GdnHCl-unfolded cyt *c* at pH 7, indicating that the U_B state is not fully unfolded. Thus, the U_B state is characterized by significant amount of secondary structure. With increments of NaCl the secondary structure is strengthened. In the presence of 2 M NaCl, the ellipticity is comparable with that for the native protein at pH 7, suggesting that the MG state is fully formed. Typically, far-UV CD spectra of proteins in the MG state are similar to those of the native states.[144] The control CD spectrum for unfolded state was recorded at pH 7, because the spectra at pH 13 are obscured when GdnHCl is added to the protein.

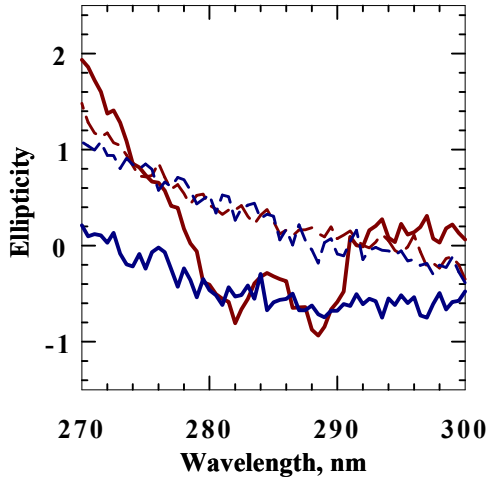


Figure 3. Near-UV CD spectra. Cyt *c* at pH 7 (—) , at pH 7 in 5M GdnHCl (- -), at pH 13 without salt (- -), at pH 13 with 2M salt (—).

Figure 2d shows the variation of the 222-nm ellipticity with NaCl in the 0-2.5 M range. The transition involves titration of one or two Na^+ ions. Clearly, alkali-denatured cyt *c* acquires extensive secondary structure in high ionic strength medium, suggesting the occurrence of the $U_B \rightarrow B$ transition.

Figure 3 presents the aromatic CD spectra of different states of cyt *c*. The spectra for the GdnHCl-unfolded protein at pH 7 and the base-denatured protein at pH 13 are nearly identical; no CD absorption is observed at 282 nm and 289 nm, indicating that the U_B state is equivalent to the unfolded state in terms of the near-UV CD signature. However, the spectrum for the B state appears to indicate a little gain in the CD absorption, a sign of formation of mobile tertiary structure.

All these data (both fluorescence and CD) suggest that the fraction of the molten globule state at pH 13 and 2 M salt is roughly 0.9. Further, the protein concentration used for various spectroscopic probes differed widely, according to the transition dipole strength of spectroscopic probes, spanning the range 9 – 70 μM. In spite of this difference, the properties and parameters derived from these data are consistent with the known properties of molten globules. The invariance of the properties to protein concentration suggests that the observed attribute of the state is unlikely due to protein aggregate or a non-monomeric form.[145] Therefore, it is clear that at pH ≥ 13, ferricyt *c* forms an extreme base-denatured state with ~40% helical structure. By adding NaCl up to 2 M, the U_B state at pH 13 can be converted to the molten globule state, B, which has native-like secondary structure, a small amount of tertiary structure and a compact size.

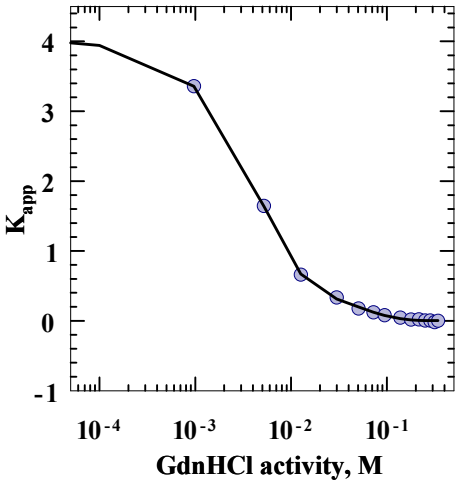


Figure 4. Stabilization or folding of the B state of cyt *c* in the presence of low concentrations of GdnHCl, pH 13, 22°C. The observable equilibrium constant, K_{app}, for the $U_B \rightleftharpoons B$ is plotted against the activity of GdnHCl. The solid line is fit to equation 2.

3.4.3. *GdnHCl-induced Stabilization and Subsequent Unfolding of the U_B state*:

An interesting behavior was observed when GdnHCl was titrated with cyt *c* at pH 13 in the absence of salt. As the denaturant concentration is raised, the fluorescence initially decreases in a sharp cooperative manner (Figure 4),

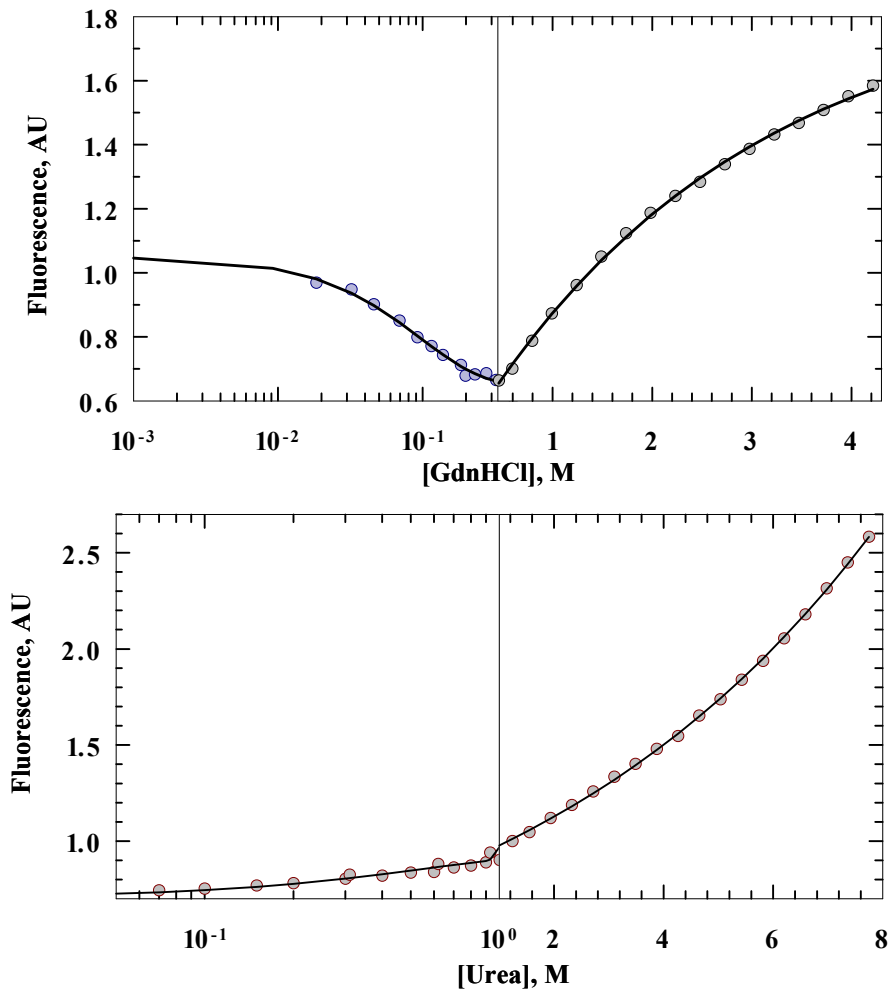


Figure 5. Interaction of ionic and neutral denaturants with cyt *c*, at alkaline pH. (a) The GdhHCl-induced folding (●) and unfolding (●) transition. Ionic denaturant stabilizes the protein at low concentrations. (b) The urea-induced equilibrium unfolding (●) does not show any stabilization effect.

and then increases relatively monotonously displaying a very weak unfolding transition (Figure 5). GdnHCl induced folding transitions for the acid-denatured form of cyt *c*, myoglobin, and a synthetic amphiphilic peptide have been reported.[150] In the present case, the initial folding transition (Figure 4) is attributed to a $GdnH^+$ mediated $U_B \rightarrow I_{B2}$ transition, where I_{B2} is a structural intermediate. Since the titration of urea, a neutral denaturant, with alkaline cyt *c* does not show such stabilization (Figure 5b), the stabilization appears to arise from screening of protein charges by GdnHCl.

Although the I_{B2} intermediate is more compact than the U_B state, it is more expanded than the B-state. Assuming that GdnHCl exerts its effect by direct protein-denaturant interaction,[151-153] the number of $GdnH^+$ ions associated with the $U_B \rightarrow I_{B2}$ transition can be estimated for a model where I_{B2} interacts preferentially with the $GdnH^+$ ions with the assumption that the binding constants for interaction of these cations with different binding sites offered by the I_{B2} state are comparable.[131] The fit of the initial folding transition data to the expression:

$$K_{app} = K(1 + K_{bind} a_x)^{\Delta n} \qquad (2)$$

where, Δn is the difference in the number of $GdnH^+$ ions bound to U_B and I_{B2} states, a_x is the activity of GdnHCl, K_{bind} is the binding constant, and K is the true equilibrium constant for the $U_B \rightleftharpoons I_{B2}$ transition in the absence of GdnHCl, yields $\Delta n \approx 2.2$, $K_b \approx 200$, and $K \approx 4.2$. At 22 ^oC the value of K corresponds to the free energy of 0.84 kcal mol^{-1}.

Above 0.5 M GdnHCl, the stabilizing effect of the $GdnH^+$ ions is overrun by their own unfolding action. The unfolding transition of the I_{B2} state is apparently shallow (Figure 5a), indicating small amount of solvent exposure of residues accompanying the unfolding process ($I_{B2} \rightarrow U$). This suggests that I_{B2}

lacks a significant hydrophobic core although patches of exposed and buried hydrophobic surfaces may be present. The GdnHCl titration data were fit to:[138]

$$\Delta G = \Delta G° - RT \ln\left(\frac{1 + K^{u}_{bind}[GdnHCl]}{1 + K^{B}_{bind}[GdnHCl]}\right)^{n} - m[GdnHCl] \qquad (3)$$

where m is the equilibrium surface exposure of residues during unfolding, n is the number of $GdnH^{+}$ binding sites, K^{u}_{bind} and K^{B}_{bind} are binding constants of $GdnH^{+}$ to the unfolded and the I_{B2} state, respectively, and ΔG^{o} is the free energy of

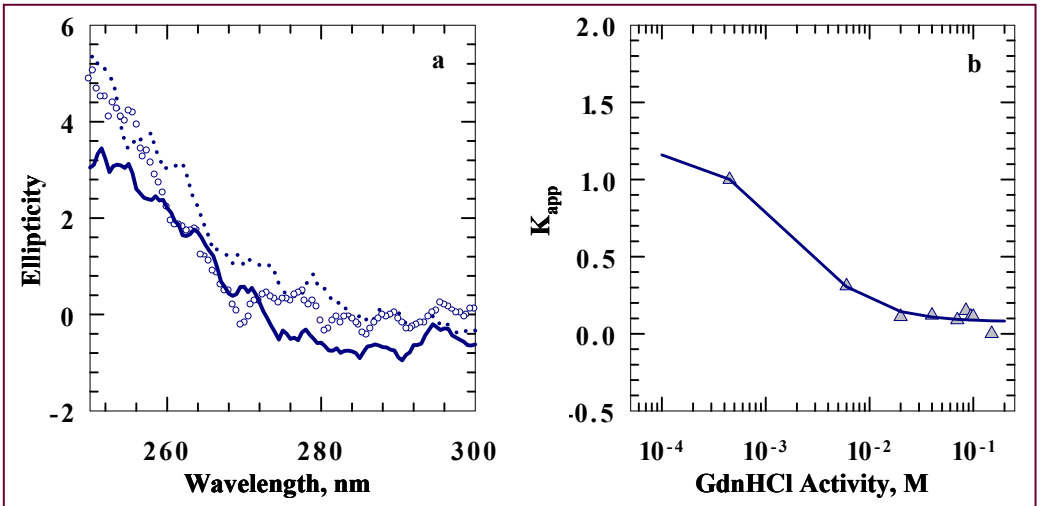


Figure 6. The GdnHCl induced stabilization of U_B state (a) Near UV-CD spectra of cyt *c* at pH13, in the presence of 0M (•••), 0.05M (o o), and 0.4M GdnHCl (—). (b) The apparent equilibrium constant for the $U_B \leftrightarrow B$ reaction is plotted against the activity of GdnHCl. The transition wasmonitored by near UV-CD and the continuous line represents the fit to the equation 2.

unfolding in the absence of the denaturant. The fit yields m=0.35 kcal mol^{-1} M^{-1} and ΔG^{o} = 1.02 kcal mol^{-1}. Compared with the values reported for the unfolding of the A state of cyt *c* (1.61 kcal mol^{-1} M^{-1} and 2.04 kcal mol^{-1}, respectively),[150] the present results show that the $GdnH^{+}$-stabilized I_{B2} state is relatively less stable. At the same time, urea does not show any stabilizing effect at low

concentrations and cannot completely unfold the U_B state within its solubility range at pH 13.

The GdnHCl-induced $U_B \rightarrow I_{B2}$ transition should be verifiable by CD as well. However, alkali solutions containing even small amount of GdnHCl show very large noisy signal due to an elevation of the HT voltage; especially, peptide signals are increasingly obscured with GdnHCl content. This difficulty has also been noted in an earlier report of alkaline ferricyt *c*.[35] Nevertheless, as shown in the Figure 6, the increase of 282 nm CD absorption with increasing concentration of GdnHCl up to 0.4 is an indication of the $U_B \rightarrow I_{B2}$ transition.

3.4.4. *Denaturant Unfolding of the B state*:

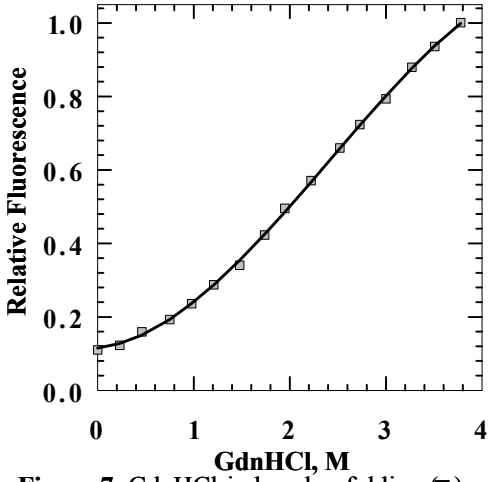


Figure 7. GdnHCl-induced unfolding (□) of the alkali molten globule state, pH13,2M salt at 22° C. The continuous line is the fit according to a two-state transition.[46]

In order to have a better understanding of the conformational stability of the B state stabilized by sodium ions, GdnHCl titration of the B state was performed. Figure 7 presents the relative fluorescence change occurring during the unfolding of the B state by GdnHCl. Assuming that the B → U unfolding transition involves no intermediate, the fit to the data using the standard two-state analysis[46] yields m = 0.53 kcal mol^{-1} M^{-1} and ΔG^o = 1.8 kcal mol^{-1}. The energetic stability of Na^+-stabilized B state is slightly higher than the $GdnH^+$ stabilized I_{B2} intermediate. A comparison with the A state analysis presented by Hagihara *et al*.[150] indicates that the conformational stability of the NaCl-stabilized A and B states of ferricyt *c* match fairly well, although the global *m*-value for the A state is about four-fold higher than that for the latter.

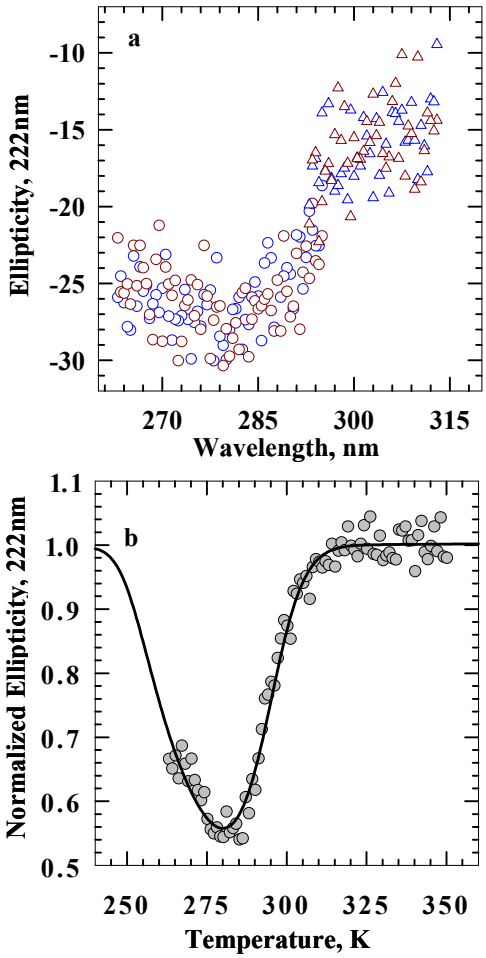


Figure 8. (a) Thermal reversibility of cyt *c* in 0.01M NaCl at pH 13 monitored at 222nm. Cooling from 293 to 263K (○) and heating from 263 to 293K (○). Heating from 293 to 313K (△) and cooling from 313 to (△) 293K. (b) Representative thermal transition curve monitored at 222nm for cyt *c* at pH 13 in 0.01M NaCl. The solid line is the fit to equation 5.

3.4.5. *Heat and Cold Denaturation of the B state*:

To access the role of hydrophobic interactions in the stability of the B state, and to determine the qualitative extent of hydration of its apolar surfaces, a series of CD-monitored temperature denaturation studies were carried out for both native and B states of ferricyt *c* in the presence of variable concentration of NaCl (pH 13). Figure 8 shows some basic details of the thermal unfolding experiments monitored by far UV-CD (222 nm) for the B state.

To ascertain that the temperature-induced changes in the CD signal are reversible so that some of the basic principles of equilibrium thermodynamics can be used to interpret data, the peptide ellipticity of a solution of cyt *c* at 0.1 M NaCl at pH 13, was continuously recorded while the sample cooled down gradually to 263 K from 295 K (Figure 8a). The same values of ellipticity were traced when the sample was heated up from

263 K to 295 K. Similarly, the heat-induced denaturation was checked for reversibility in the 293-313 K range. The unfolding was found to be irreversible when heated to temperatures higher than ~315 K, perhaps due to aggregation of the heat-unfolded chains. However, 315 K would appear sufficient for substantial unfolding of the MG state (Figure 9a), allowing one to rely on equilibrium thermodynamics for data interpretation. Figure 9a presents thermal denaturation curves of cyt *c* at pH 13 in the presence of 0.01M – 2M range of NaCl observed by 222 nm CD signal. The CD amplitude increases with salt increments, indicating increasing fraction of the B state. Because the low-temperature ellipticity values are dependent on salt concentration but the high-temperature ones are not, the data presented were normalized according to:

$$\theta = \frac{\theta_{obs} - (m_{pre}T + c_{pre})}{(m_{post}T + c_{post}) - (m_{pre}T + c_{pre})} \qquad (4)$$

where, θ_{obs} is the observed ellipticity, T is the temperature, m_{pre} and c_{pre} are slope and intercept, respectively, of the pre-transition baseline for the 2 M salt curve, and m_{post} and c_{post} are slope and intercept of the post-transition baseline in the presence of a given concentration of NaCl. The normalized thermal curves were analyzed by the Gibbs-Helmholtz equation:

$$\theta(T) = \frac{(c_{pre} + m_{pre}T) + (c_{post} + m_{post}T)\exp\left[\dfrac{\Delta H_m(T/T_m - 1) + \Delta C_p[T_m - T + \ln(T/T_m)}{RT}\right]}{1 + \exp\left[\dfrac{\Delta H_m(T/T_m - 1) + \Delta C_p[T_m - T + \ln(T/T_m)}{RT}\right]} \qquad (5)$$

where, ΔH_m is the enthalpy at the transition midpoint, T_m, and ΔC_p is the change in heat capacity. Table 1 lists the values obtained from the 222 nm data. The thermal curves exhibit uniform baselines at high temperatures, indicating that the temperature dependence of the ellipticity values of the unfolded protein measured

at 222 nm is independent of NaCl concentration. There are two temperature dependent unfolding transitions, one at low temperature and the other at high temperature. The low temperature unfolding transition is increasingly noticeable with decreasing salt concentration. Figure 8b magnifies this feature for the 10 mM NaCl data, indicating that the B state of cyt *c* undergoes cold denaturation. It becomes more obvious at lower concentration, because proteins undergo cold denaturation under conditions where the conformation is relatively destabilized

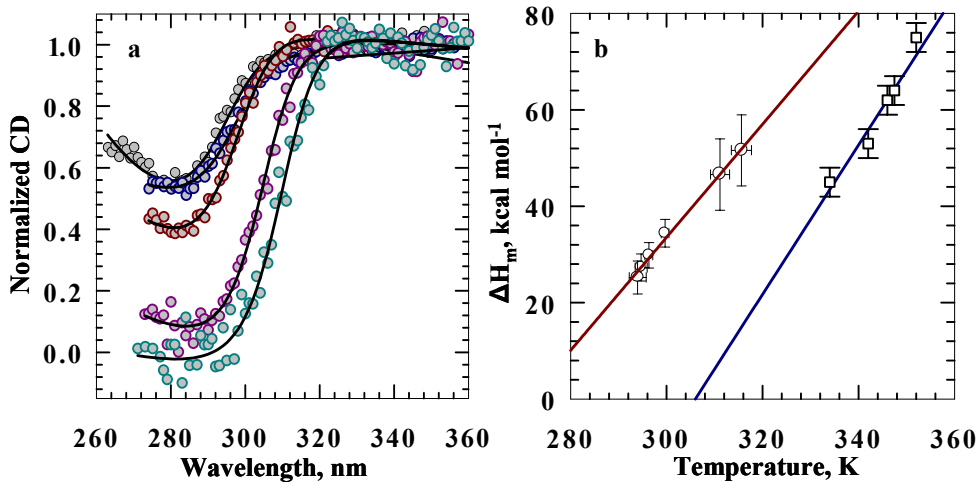


Figure 9. (a) Thermal transitions of cyt *c* at pH 13 monitored with far-UV CD. The data were normalized according to equation 4. The symbols (○), (○), (○), (○), and (○) represent the transitions with 0.01M, 0.1M, 0.25M, 1.0M and 2.0M salt respectively. The continuous lines are fit according to equation 5. (b) Determination of ΔC_p for thermal unfolding of the B-state (○) by linear analysis of ΔH_m obtained from panel (a) as a function of T_m. For comparison, variation of ΔH_m with T_m for native ferricyt *c* at pH 7 is also plotted. Slope of the lines are listed in Table 2.

by extremes of pH or ionic strength, or both.[154]

3.4.6. ΔC_p *for the Unfolding of the B State*:

The thermal denaturation data facilitate the exrraction of ΔC_p by linear analysis of ΔH_m as a function of T_m. Values of ΔH_m were obtained in three ways: (1) from van't Hoff analysis of thermal denaturation data (ln K versus $1/T$) assuming that the temperature unfolding of B is two-state in nature;[155-157] (2) from the entropy change at T by using $\Delta S = (\partial\ \Delta G/\partial T)$, and $\Delta H_m = T_m \Delta S_m$; and (3) from the fits of the thermal data by the use of equation 5. For brevity, only the third approach to find the ΔC_p value is presented. Figure 9b shows the ΔH_m versus T_m graph for those vales of ΔH_m that were derived from the transition curves presented in the Figure 9a. For comparison, data for native cyt *c* recorded at pH 7 without salt are also presented. Values of ΔC_p calculated from the slopes of all plots are listed in Table2.

Table 1. Thermodynamic parameters for unfolding of the B state of ferricyt *c* at pH 13, calculated from 222nm CD data

From Gibbs-Helmholtz eqn					vant' Hoff eqn		Entropy-enthalpy compensation		
NaCl (M)	T_m (K)	T_c (K)	ΔH_m†	ΔC_p‡	T_m (K)	ΔH_m†	T_m (K)	ΔH_m†	ΔS_m‡
0.01	289.0	257.0	20.1	1.20	289.6	20.1	293.1	19.5	-0.07
0.05	294.0	250.0	26.7	1.15	294.0	27.8	298.1	23.0	-0.07
0.25	295.3	246.0	28.9	1.02	295.8	32.9	300.0	32.2	-0.10
1.00	303.1	245.0	34.4	1.10	304.0	35.8	305.0	32.4	-0.10
2.00	308.8	241.5	36.1	0.98	309.8	39.3	309.0	35.2	-0.11

† in kcal mol^{-1} ‡in kcal mol^{-1} K^{-1}

The values of ΔC_p for the unfolding of the B state reported here is significantly higher than that in the literature for the thermal unfolding of the A state of cyt *c* (Table 2). For the denaturation of the native state though, the ΔC_p value estimated from this work (1.56 kcal mol^{-1} K^{-1}) is in fair agreement with earlier reported values of 1.26 kcal mol^{-1} K^{-1} for horse ferricyt *c* and 1.37 kcal

mol^{-1} K^{-1} for yeast iso-1-ferricyt *c*.[158] These comparisons establish the following order for the extent of hydrophobic interactions in different states of cyt *c*: N>B>A>U.

Table 2. ΔC_p values from thermal unfolding of A, B and N states of ferricyt *c* determined from dependence of ΔH_m on T_m

Transition	Method	Probe	ΔC_p	Reference
B → U_B	van't Hoff	CD, 222nm	0.85	This work
	ΔS_m at T_m	CD, 222nm	0.63	
	Gibbs-Helmholtz	CD, 222nm	0.91	
A → U_A	van't Hoff	CD, 222nm	0.38	138
	van't Hoff	CD	0.40	128
N → U	van't Hoff	CD, 222nm	1.26	138
	van't Hoff	Soret Abs	1.55	128
	Gibbs-Helmholtz	Soret Abs	1.57	This work

3.4. Discussion

3.4.1. *The Alkali Denatured State of Ferricyt c*:

The base-denatured state, U_B is not the end product (sometimes referred to as state IV) of the commonly known alkaline transition of ferricyt *c*. Most recent studies indicate that the alkaline transition involves the deprotonation of some unidentified buried groups controlled by unfolding of the ω-loop foldon.[159-160] This local unfolding event results in a swap of the heme ligand (M80 → K, pK≈9.3) with the structural readjustments at the site of the heme.[161] Major structural perturbations occur at above pH 12.[35,162] The tertiary structure is fully lost at pH 13 (Figure 1a), but a fraction of destabilized helical structure is retained. This is the gross structural identity of the U_B state, which has also been likened to an equilibrium intermediate.[35] It is well known that pH-denatured states, although largely structureless, do not simulate ideal random coils,[63,111] and

in many cases they are transformed to the molten globule state by the charge screening effect of added ions.

3.4.2. *A and B States of Proteins: Charge-stabilized Molten Globules*:

The A state is well known as a general anion-stabilized thermodynamic state of proteins.[111-114,130,131] Although the complementary cation-stabilized state under extreme basic conditions has been reported for only a few proteins,[132,133,140] the present results for ferricyt *c* provide growing evidence for the generality of the B state as the cation-stabilized alkali equilibrium state. Since A and B states accumulate under extreme pH conditions, they should be stabilized by the added counter ions through the same mechanism, the solvent ions either reduce the intrapolypeptide electrostatic repulsion or bind directly to the protein charges to form ion pairs, or possibly both.[131,138,163] They are also expected to share at least qualitatively similar structural and dynamic properties. Indeed, the B state exhibits all of the basic properties known for the A state of ferricyt *c*.[131,134,136] Thus, A and B states are complementary charge-stabilized molten globules.

3.4.3. *Folding/Stabilization of Alkali Denatured Cyt c by GdnHCl*:

It is observed that the U_B state is transformed to an intermediate, I_{B2}, by $GdnH^+$ ions (Figure 4). The complementary $U_A \rightarrow A$ transition for ferricyt *c* is also driven by GdnHCl.[150] In both cases, the ions dissociated from GdnHCl influence the stability. The mechanism of interaction of $GdnH^+$ with negative protein charges, and of Cl^- with the positive protein charges need not be the same, however. In acid medium (pH < 2), where the net charge of cyt *c* is +24, Cl^- can shield the positive protein charges by Debye-Huckel or ion-pair or both types of interactions. In the basic medium (pH > 12.5), where the net protein charge is -17, $GdnH^+$ ions are more likely to form ion-pairs with anionic sites of the protein. Formation of Debye-Huckel type spheres will perhaps be unfavorable due to large size of the $GdnH^+$ ions. Crystal structures of proteins with low concentrations of GdnHCl do show direct interaction of $GdnH^+$ ions with proteins.[164-166] This

argument is advanced only in the context of GdnHCl stabilization of base-denatured cyt *c*. Na^+ ions, far smaller than $GdnH^+$ by radius, may form charge spheres as well.

Stabilization of native or native-like states by sub-denaturing concentrations of GdnHCl at neutral pH has also been shown for a few proteins.[110,166,167] Although ionic interactions between the ions of GdnHCl and the protein are implicated in all cases, the structural, dynamical, and thermodynamic consequences of protein-denaturant interactions are more involved than the simple charge screening argument.[151,152,166] The interactions of the denaturant molecules with different groups of the protein can establish non-specific networks of intramolecular interactions, as observed in crystal structures of proteins with low concentrations of denaturants.[165] These cross-linking interactions may serve to decrease the motional freedom in different parts of the protein, but to increase the barriers to motions in the more compact conformers of the protein. Existing X-ray data for proteins do show a significant decrease in the *B*-factor of side-chain and backbone atoms in the presence of low concentrations of GdnHCl.[157] In this scenario, $GdnH^+$-induced stabilization of the alkali-denatured protein ($U_B \rightarrow I_{B2}$ transition; Figure 4) would appear to originate largely from a reduction in fluctuations in the positions of individual or clusters of atoms around their average. The relative contribution of these two effects– electrostatic shielding and conformational entropy– to the free energy of stabilization of proteins remains to be assessed.

3.4.4 *Thermal Unfolding of the B State*:

Although molten globules in certain cases have not been found to exhibit a cooperative thermal unfolding transition,[144,169] data for other proteins, including cytochrome *c*, apomyoglobin, staphylococcal nuclease, equine lysozyme, and canine milk lysozyme, show cooperative thermal transitions.[113] Consistent with

the latter observation, the B → U_B transitions reported here are rather cooperative; the steepness increases with increase in the NaCl concentration (Figure 9a), reflected clearly by the sizable values of ΔH_m listed in Table 1. Thus, the cation-promoted stability proceeds with an increase in enthalpy. The value of ΔC_p for the B → U_B transition reported here is at least 0.4 kcal mol^{-1} K^{-1} higher than that for the A → U_A thermal transition of ferricyt *c* (Table 2), suggesting that the hydrophobic residues in the B state are relatively more shielded from water. Taking the water accessibility of hydrophobic surfaces as a measure of unfoldedness, the data suggest that the A state of cyt *c* is more unfolded than the complementary B state.

Another interesting observation is a decrease in ΔS_m associated with the B → U_B transition regardless of the amount of NaCl added to the medium (Table 1). Since the conformational entropy of the protein in the B state cannot be higher than that in the U_B state, the observed entropy suppression must arise from the entropy of water. The change in the water entropy has been suggested to originate from the hydrophobic effect,[170] which appears to contribute considerably to the forces stabilizing the B state. This is consistent with the large value of ΔC_p for the B → U_B transition. The molten globule state is thought to be stabilized by the same forces that are cooperative in the native state, i.e. hydrophobic and hydrogen bonding interactions,[138] and the influence of the hydrophobic effect has been observed.

3.4.5 *Cold Denaturation of Molten Globules*:

Although the phenomenon of cold denaturation is firmly based on the thermodynamic principles of protein stability,[154,155,171,172] perhaps no native protein under normal conditions of pH and pressure has yet been found to undergo cold denaturation without the assistance of destabilizing solvent additives. This is because the native-state stability does not depreciate sufficiently in the experimentally accessible low temperatures. Destabilizing

conditions produced by denaturant or extreme pH facilitate observation of cold denaturation.[154] Thus, even though cold denaturation of native ferricyt *c* was predicted many years ago[171] experiments have not yet witnessed low-temperature unfolding of this protein at neutral pH. However, the A state of cyt *c* has been seen to denature at low temperature.[128] Similarly, cold denaturation of the molten globule state of apomyoglobin,[129] the low-pH form of holomyoglobin,[173] a partially unfolded state of yeast iso-1-cytochrome *c*,[174] α-lactalbumin,[175] and β-lactoglobulin[176] has been reported.

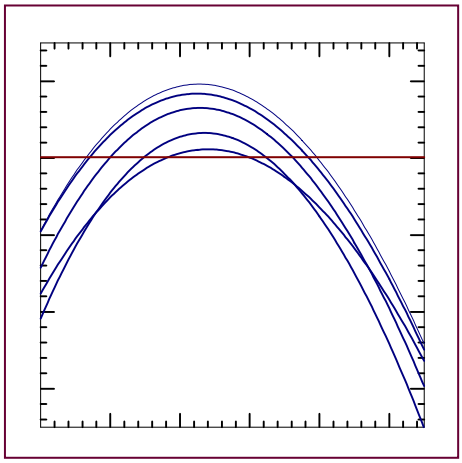
The molecular basis of cold denaturation of proteins is not fully understood. According to one proposal, cold denaturation arises from decreasing hydrophobic contribution to protein stability as the temperature is lowered.[177] This is because cold water destabilizes hydrophobic interactions, and the use of solvent additives like alkali and salt, as for the experiments reported here, is likely to alter the general mechanism of the process.[155] A recent experimental study of the structure of a cold denatured protein lends support to this view. The cold-denatured state appears to unfold in a non-cooperative fashion, as may be thought to arise from a weakening of tertiary, hydrophobic-driven interactions.[178] The role of hydrophobic effect in cold-denaturation derived from a two-state model of water structure has also been proposed.[179,180] On the other hand, reports that one can account for ΔC_p of unfolding for many proteins by summing up hydration contributions from individual groups[181-184] implies that both hydrophobic and hydrophilic interactions are responsible for cold denaturation. If so, the relative contributions of the two effects is expected to depend on the content of polar and apolar residues. But the issue of molecular basis of cold denaturation is deep and more remains to be revealed.

3.5. Conclusion

To complement the well known A state, there exists a B state of ferricyt *c*. In extreme basic medium, cations transform the denatured state of ferricyt *c* (U_B) into a charge-stabilized molten globule state (B). The B state meets with the basic molecular and structural organizational definition of the molten globule state. Hydrophobic effect appears to contribute substantially toward the energetic stabilization of the B state, since the thermally driven B → U_B transition is characterized by a large heat capacity change. The energetic stability of the B state depreciates with decreasing temperature, and it undergoes cold denaturation.

Chapter 4

Cold Denaturation of Ferricytochrome *c*: Protein Destabilization at Low Temperature



4.1 Abstract

Determining the forces stabilizing a protein is one of the primary issues in protein chemistry. Thermal denaturation of proteins has successfully been applied to analyze the stabilizing interactions in proteins. However, protein denaturation at low temperature has not been studied as widely as heat denaturation, particularly in the case of cyt *c*. Cold denaturation under amenable temperature range can be achieved by destabilizing the protein with GdnHCl. At low temperatures cyt *c* unfolds with moderate amount of GdnHCl and shows less free energy change. It shows maximum stability near 288 K. The results explain the nature of hydrophobic interactions and water structure in cyt *c* during thermal denaturation. It also compares the enthalpy and entropy changes occurring during heat and cold denaturations. The denaturant dependence of thermodynamic properties insinuates the probable cold denaturation transition temperature (~235 K) for cyt *c*.

4.2 Introduction

A delicate balance between hydrogen bonding, hydrophobic effect, and salt bridges, which operate in both intramolecular protein interactions and protein-protein or protein-solvent interactions, stabilize proteins in their native conformation.[186] Thermodynamic analysis of protein unfolding can provide clues to the effects of these interactions in terms of entropy, enthalpy, and heat capacity. Denaturation of proteins by chemical denaturants, pH changes, and thermal variation has been adapted as valuable techniques to estimate the thermodynamic properties of proteins. Among these, thermal denaturation obtains a direct relation between two of the most fundamental variables- temperature and enthalpy (heat capacity).[171] At high temperatures proteins are more prone to

aggregation. Thus, the process of denaturation of proteins can be studied only under conditions that prevent aggregation.

Calorimetric experiments on proteins have led to the conclusion that the denaturation free energy change shows a parabolic-like profile as a function of temperature.[187] This suggested the existence, beyond the well-known heat denaturation of globular proteins, of a second conformational transition caused by a temperature decrease, which would usually occur at temperatures lower than room temperature. This intriguing phenomenon is called 'cold denaturation'.[188] Although cold denaturation of globular proteins was suggested long back,[187] the experimental procedures have intrinsic problem of inaccessibility of very low temperatures at which globular proteins can be unfolded. However, experiments under slightly destabilizing conditions can be performed to extract the cold denaturation profile. The low temperature instability studies have been carried out for metmyoglobins in acid,[173] for ribonuclease[189] and metmyoglobin[190] under high pressure, for α-chymotrypsinogen by the emulsion droplet technique,[191] for T4 lysozyme[172] and β-lactoglobulin[176] in urea solution, and for phosphoglycerate kinase[192] and Che Y[193] in guanidinium hydrochloride solution. More recently, ubiquitin encapsulated in reverse micelle was used to examine the cold denaturation.[194]

Low temperature unfolding of cytochrome *c*, one of the well documented protein in folding and stabilization studies, has not been carried out. Even though a few cold destabilization experiments with the molten globule states of cyt *c*,[128,195] as also described in Chapter 3 of this thesis, have been described, experiments with the native state at neutral pH remains unreported, perhaps due to large free energy of stabilization of cyt *c* at pH 7. This can be overcome by adding a destabilizing agent such as GdnHCl that will lower the stability of the protein. The low temperature instability can then be brought into the range of

experimentally accessible temperatures. The denaturant also serves to lower the freezing temperature of water and accelerates the unfolding reaction.[172]

As expected, the cold denaturation of cyt *c* is associated with decrease in enthalpy and entropy with a positive heat capacity change. These thermodynamic changes are explained by hydrophobic interactions and water structure of cyt *c* at various temperatures. Although, molecular nature of hydrophobicity of proteins has not been well understood, these experiments emphasize the importance of hydrophobic interactions in stabilizing the native conformation of protein.

4.3 Materials and Methods

4.3.1. *GdnHCl Induced Equilibrium Unfolding of Ferricyt c*:

Stock solutions of 100 mM phosphate buffer with (unfolding buffer) and without (native buffer) 5 M GdnHCl were prepared, and pH of the solutions was adjusted to 7(±0.5). Equal volumes of cyt *c* in phosphate buffer without denaturant was added to both solutions to obtain a final protein concentration of ~15 μM. The solutions were kept at the required experimental temperatures for ~1 hour in a water bath. In case of low temperature studies, the native buffer contained low concentrations of GdnHCl (upto 1 M) so as to avoid freezing of the samples. Native and unfolded buffers were mixed in different volumes to obtain different final concentrations of GdnHCl. The solutions were kept in the experimental temperature for ~15 min before recording the spectra. The spectra were taken between 250 and 210 nm, with 50 nm/min acquisition rate using a JASCO J-810 CD instrument. The final concentration of GdnHCl in each solution was determined by refractive index measurement.

4.3.2. *Cold and Heat Denaturation Studies*:

Temperature induced changes in aromatic CD absorption (at 222 nm) were measured for cyt *c* at pH 7 in different concentrations of GdnHCl. The

protein concentration was ~20 μM. Temperature changes between 263 and 353 K were controlled by Peltier with cooling and heating rate of 0.5 ^{o}C/min. For control experiments, protein without GdnHCl was monitored from 273 to 313 K, and with 5 M GdnHCl monitored from 263 to 343 K.

4.4 Results

4.4.1. *Unfolding Transition of Ferricytochrome c at Different Temperatures*: Denaturation effect of GdnHCl on ferricyt *c* at neutral pH was investigated at various temperatures using the peptide ellipticity change. Values obtained at different temperatures are normalized with respect to the native and unfolded ellipticites of cyt *c* at room temperature (295 K). The equilibrium unfolding transitions do not hint at accumulation structural intermediates. Therefore, thermodynamic analysis based on two-state process is applied to analyze the data.[46] Figure 1 presents the two-state equilibrium fit of the relative ellipticities. The equation and the procedure used for the analysis of such data are described in Chapter 1 (equation 1).

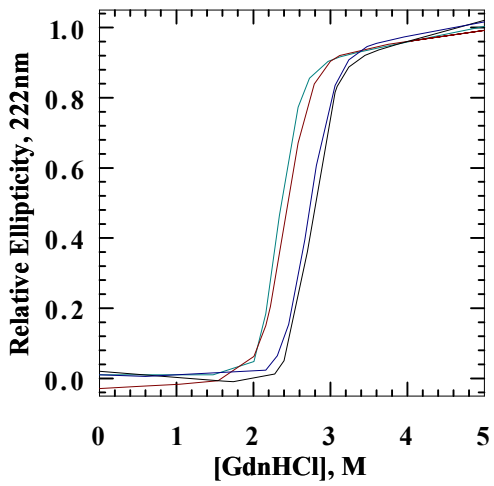


Figure 1. GdnHCl induced unfolding of ferricyt *c*, at pH 7, followed by changes in ellipticity at 222nm. (—), (—), (—), and (—) represent the unfolding transition at 263K, 273 K, 288 K and 295 K respectively. Cyt *c* shows maximum stability at 288K (15 ^{o}C).

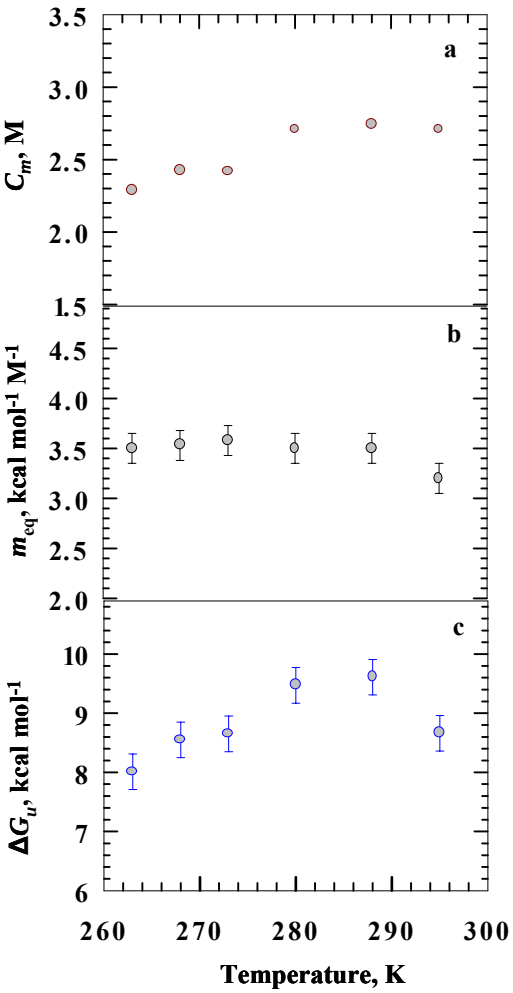


Figure 2. Temperature dependence of various parameters obtained from two-state equilibrium unfolding of cyt *c* at pH 7. (a) The transition midpoint value, C_m for GdnHCl unfolding transition. (b) Equilibrium *m*-value corresponds to the exposure of buried surface area during unfolding. (c) Free energy change of unfolding reaction.

Parameters obtained from the unfolding transition curves are shown in Figure 2. Transition midpoint (C_m value) and free energy of unfolding (ΔG_u) attains maximum value near 288 K. Above and below this temperature, both parameters decrease. The loss of free energy from the maximum stability point at higher as well as lower temperature regions of the plot predicts the possibility of heat and cold denaturation of the protein, respectively. However, the m_{eq} value which corresponds to the slope of the unfolding transition and expresses equilibrium surface exposure of buried amino acid residues, is not notably altered with temperature changes.

4.4.2. *Thermal Reversibility of Cyt c*:

While heating protein solutions, solubility of the protein drastically decreases over a narrow temperature range. It may lead to aggregation of proteins at higher temperatures[171] and produces

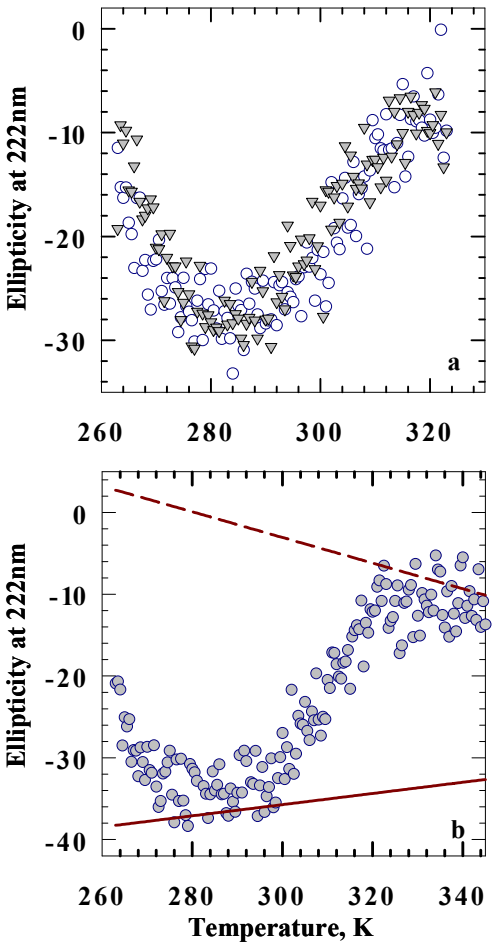


Figure 3. (a) Reversible thermal transition of cyt *c* in 2.6M GdnHCl at neutral pH. Ellipticity is monitored for the protein heated from 263K to 330K (○). The sample is then cooled down to 263K (▽). (b) Representative thermal denaturation of ferricyt *c* in 2.5M GdnHCl followed by 222nm CD ellipticity (○). Native (—) and unfolded (- - -) baselines at pH 7 are also shown.

irreversibility of thermal denaturation. Incidentally, when ferricyt *c* is heated from low temperature (263 K) to high temperature (330 K), where it is transformed into a completely thermally denatured state, it does not produce any aggregated form (or produces insignificantly low concentration of non-monomeric forms). Figure 3a shows this reversible behavior where cyt *c* was heated from 263 K to 330 K and then continuously cooled down to 263 K from 330 K The result allows the use of basic equilibrium thermodynamics for analyses of thermal transition data for ferricyt *c* held at neutral pH.

4.4.3. *Heat and Cold Denaturation*:

Since ferricyt *c* does not unfold at low temperatures due to its high free energy of unfolding, it is slightly destabilized with GdnHCl to extract its cold denaturation properties. Figure 3b represents one such thermal transitions for cyt *c*. One notices that both cold and heat denaturation processes are cooperative in nature.

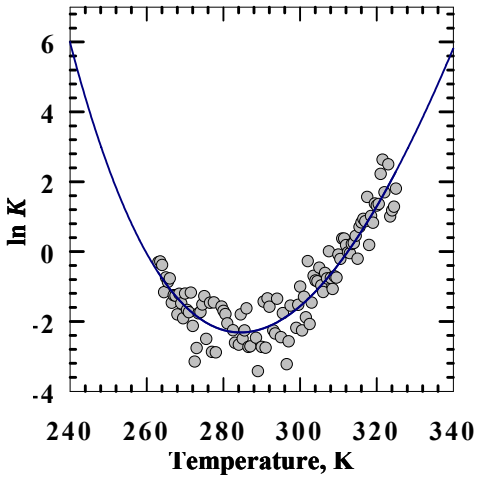


Figure 4. Thermal stability of cyt *c* in 2.5M GdnHCl. Equilibrium constant, *K,* for two-state thermal transition N⇌U is plotted against temperature. The solid line represents fit to equation 1.

For reversible melting, low-temperature and high-temperature melting cannot be regarded as qualitatively different processes. The low-temperature melting can be inferred from the extrapolation of data from high-temperature melting.[171,172,187] In these experiments, the unfolding reaction is assumed to be two-state and the change in partial molar heat capacity, ΔC_p, is assumed to be constant over the range of experimental temperature.[171] Applying the usual procedures for two-state transitions, the ellipticity values are transformed to extract the unfolded fraction with respect to temperature. To evaluate the native baseline, the thermal melt of the native protein without denaturant has been measured from 273 to 313 K and linearly extrapolated to both ends of the temperature ranges studied for the experiment. The changes observed for cyt *c* in 5 M GdnHCl solution from 263 to 343 K has been taken as fully unfolded state baseline (Figure 3b).

The equilibrium constant, *K*, determined for the unfolding reaction as a function of temperature is presented in figure 4. The derived data are fitted to the constant heat capacity model with a rational assumption that low-temperature melting is qualitatively similar with high-temperature melting for the reversible melting process and the low-temperature melting can be inferred from the extrapolation of the high-temperature melting data.[172] It can be expressed in terms of constant coefficients as:

$$\ln K = A + B(T_o/T) + C\ln(T_o/T) \qquad (1)$$

with,

$$A = [-\Delta C_p + \Delta S_m(T_o)]/R$$

$$B = [\Delta C_p - \Delta S_m(T_o)]/R - [\Delta G(T_o)/RT_o]$$

$$C = -\Delta C_p/R$$

where, T_o is an arbitrary reference temperature, ΔG and ΔC_p are Gibbs free energy change and partial molar heat capacity of unfolding, respectively. ΔS_m is unfolding entropy at transition midpoint temperature T_m.

Table 1. Coefficients and thermodynamic parameters obtained from thermal transitions of cyt *c*

	GdnHCl concentration (M)				
	2.31	2.37	2.52	2.79	2.82
T_m^{cold} (K)	253	254	260	269	276
-*A* or *B*	563.5	540.9	593.9	638.4	465.5
-*C* or $(\Delta C_p/R)^{\S}$	502.7	485.3	542.7	599.6	446.4
$-\Delta S_m$ (cal mol^{-1} K^{-1})	120.4	110.2	101.4	76.8	37.8
$-\Delta H_m$ (kcal mol^{-1} K^{-1})	30.4	27.9	26.4	20.7	10.4
T_m^{heat} (K)	320	317	312.5	305	300
-*A* or *B*	444.6	433.4	494.1	563.1	428.3
-*C* $(\Delta C_p/R)^{\S}$	502.7	485.3	542.7	599.6	446.4
$-\Delta S_m$ (cal mol^{-1} K^{-1})	114.8	102.7	96.2	72.3	35.9
$-\Delta H_m$ (kcal mol^{-1} K^{-1})	36.8	32.6	30.1	22.1	10.8

§ - ΔC_p is partial heat capacity and *R* is gas constant equals to 1.98 cal mol^{-1} K^{-1}

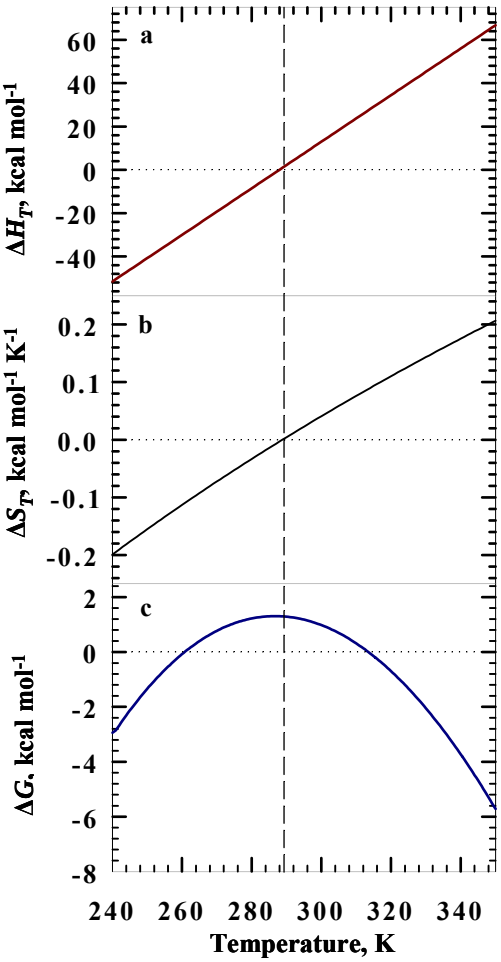


Figure 5. Temperature dependence of enthalpy, entropy and free energy change between the native and denatured state of ferricyt *c* at pH 7 in 2.5M GdnHCl. The vertical line corresponds to the temperature of maximum stability, 288.5K.

The equation indicates that if T_o corresponds to either of the transition midpoint temperatures (T_m^{cold} or T_m^{heat}), then ΔG becomes zero by definition, and hence $A = -B$. Applying these relations to the obtained ln K values, the coefficients A, B and c and subsequently ΔS_m and ΔC_p are computed for cyt *c* in different GdnHCl concentrations (Table 1). Enthalpy of unfolding at T_m can be calculated from the relation, $\Delta H_m = T_m * \Delta S_m$, since at transition midpoint the free energy change is zero.

4.4.4 *Temperature Dependence of Thermodynamic Parameters*:

It is well known from the basic thermodynamic relations that free energy, enthalpy, and entropy are temperature dependent functions. Gibbs free energy change of standard two-state transition can be derived from the equilibrium constant (K) using the relation $\Delta G = -RT \ln K$. Since the temperature dependent change on K is experimentally measurable (Figure 4), estimation of free energy change is viable from this equation (Figure

5c). Similarly, enthalpy and entropy changes can also be evaluated (Figure 5a and b) from the following equations, [173]

$$\Delta H_T = \Delta H_m - (T_m - T)\Delta C_p \qquad (2)$$

$$\Delta S_T = \frac{\Delta H_m}{T_m} - \Delta C_p \ln(T_m / T) \qquad (3)$$

It is clear from figure 5 that the enthalpy of unfolding changes linearly with temperature; the entropy though deviates a little. In addition, both enthalpy and entropy change their signs from positive to negative at low temperature, called inversion temperature, T_{inv}.[173] The inversion of sign for entropy change happens at the temperature of maximal stability T_{ms}. Concerning enthalpy, the sign inversion is seen to occur for many proteins at a slightly lower temperature. The same is expected for cytochrome *c* as well. However, with the uncertainty of measurement it is not emphasized here.

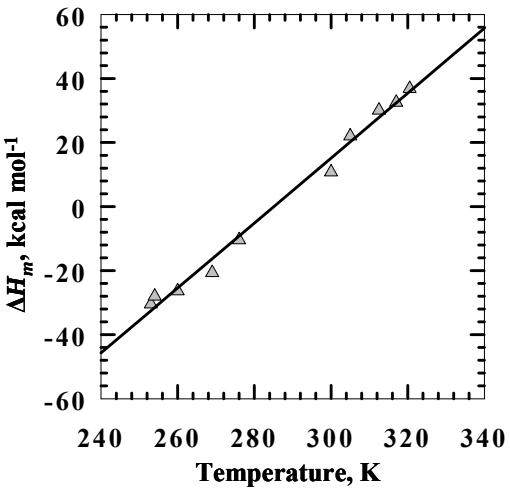


Figure 6. Plot of enthalpy of unfolding, ΔH_m versus transition temperature T_m. The data are obtained from Table 1. The solid line represents the linear fit to the data. Slope of the curve corresponds to ΔC_p of unfolding (1.014 kcal mol^{-1} K^{-1})

4.4.5. *Heat Capacity of Unfolding*:

Linear analysis of enthalpy change (ΔH_m) with respect to temperature provides the heat capacity change. Enthalpy of unfolding acquired from the analysis of thermal denaturation of cyt *c* destabilized with different GdnHCl concentrations (Table 1) has been used to calculate the ΔC_p value (Figure 6).

The obtained ΔC_p value (1.014

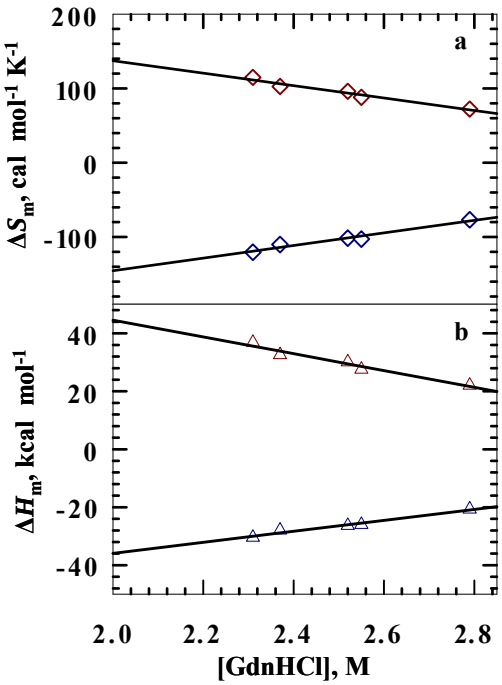


Figure 7. Variation of enthalpy and entropy with denaturation concentration. (a) Entropy of unfolding at T_m decreases for heat denaturation (◇) and increases for cold denaturation (◇). (b) Enthalpy of unfolding at T_m also decreases for heat denaturation (△) and increases for cold denaturation (△). The straight lines are linear fits of the corresponding data.

kcal mol^{-1} K^{-1}) is slightly lower than the values already reported.[128,138] The techniques used to calculate the heat capacity change are different. Destabilization of cyt *c* at lower pH values was extracted to evaluate the ΔC_p of unfolding (1.38 kcal mol^{-1} K^{-1}) by Kuroda *et al*.[128] Hagihara *et al*.[138] employed the acetylated cyt *c* to calculate the same (1.26 kcal mol^{-1} K^{-1}). Even though different procedures have been adapted to compute the unfolding heat capacity, the values are quite similar.

4.4.6. *Denaturant Effect on Enthalpy and Entropy*:

The added GdnHCl alters the thermodynamics of unfolding of cyt *c* (Table 1). Enthalpy and entropy of denaturation measured at both cold and heat transition midpoints change linearly with denaturant concentration (Figure 7). The linear extrapolations of these lines yield the corresponding thermodynamic properties at zero denaturant concentration. It must be noted that during cold denaturation both enthalpy and entropy obtain negative values whereas for heat denaturation both are positive. At transition midpoints the Gibbs free energy of unfolding is zero, hence ΔH_m equals $T\Delta S_m$. It leads to a valuable

prediction of transition midpoint temperature for cold (T_m^{cold}) and heat (T_m^{heat}) transition of the protein in the absence of denaturant (Table 2).

Table 2. Thermodynamic properties for thermal unfolding of ferricyt *c* in the absence of denaturant.

	Cold Denaturation	Heat Denaturation
ΔH_m (kcal mol^{-1})	-73.9	102.4
ΔS_m (cal mol^{-1} K^{-1})	-314.7	304.3
T_m predicted (K)	234.8	336.5
T_m experimental (K)	--	331.0§

§ - Bágeľova, J. *et al.*[201]

4.5 Discussion

4.5.1. *Thermal Stability of Cytochrome c*:

It is very convenient to estimate the stability of proteins from their unfolding free energy. Ferricyt *c* shows remarkable changes with respect to temperature which can be assessed from its denaturant induced unfolding transition curves (Figure 1 and 2). Maximum thermal stability of the protein (T_{ms}) is obtained near 288 K (15 ^{o}C). This is similar to the value already attained from urea denaturation of cyt *c* (T_{ms} = 293 K).[196] Moreover, it is consistent with the fact that many globular proteins reach their maximum marginal stabilities around room temperature,[197] since many nonpolar compounds show minimum solubility in water around room temperature.[170]

The denaturant m_{eq} value, the dependence of free energy of unfolding on denaturant concentration ($\partial\Delta G_u/\partial$[denaturant]), is independent of temperature for cyt *c* (Figure 2b). Thus, it is expected that the temperature at which C_m is

maximal will also be the temperature at which ΔG_u is maximal as per the definition of two-state unfolding transition, $C_m \approx \Delta G_u / m_{eq}$. It is apparently demonstrated from Figure 2 that both C_m and ΔG_u attain maximum at the same temperature. The non-dependency of m_{eq} with temperature has also been characterized for some other two-state proteins like Hpr[198] and cold shock protein B.[199] The m_{eq} value is directly proportional to the change in solvent accessible surface area ΔASA, the difference in solvent accessibility of the native form and the unfolded form. Therefore, it enunciates that solvent accessibility of the protein while unfolded in GdnHCl is not substantially changed with temperature, although it perceives considerable change in stability.

4.5.2. *Thermal Denaturation of Ferricyt c*:

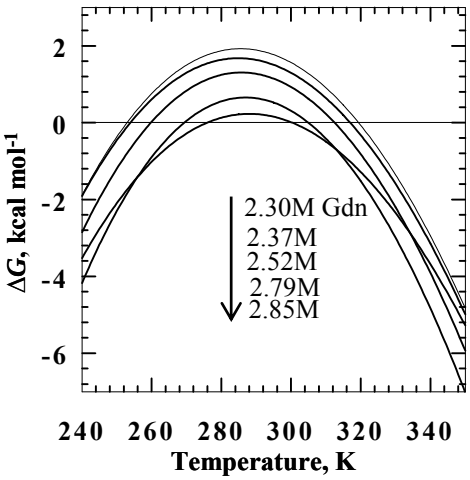


Figure 8. Free energy change as a function of temperature. As the concentration of GdnHCl increases the transition midpoint where ΔG=0 moves toward the centre and the free energy at maximum stable temperature decreases.

Decrease in free energy change, when a protein is cooled from the temperature of its maximal stability, promises for cold denaturation of the protein.[200] Ferricytochrome *c*, a two-state globular protein, also shows such a behavior (Figure 3b). The denaturation temperature of the native protein, however, falls well below the experimentally accessible temperature range. It obstructs the study of cold denaturation of many proteins. Here, the native protein is slightly destabilized with GdnHCl to draw the protein into the accessible low-temperature range. The added denaturant decreases the heat

capacity of the protein, and shifts the cold denaturation into higher temperature. Thermal melting of cyt *c* does not show aggregation at higher temperature (Figure 3a). Therefore, estimation of all thermodynamic properties becomes viable using equilibrium thermodynamic relations.

4.5.3. *Free Energy of Unfolding*:

Increase of GdnHCl concentration decreases the free energy of unfolding, and cold and heat denaturation transition midpoints are shifted to higher and lower temperatures, respectively. Maximal stability temperature of the protein does not vary with the addition of GdnHCl (Figure 7).

4.5.4 *Enthalpy and Entropy Compensation*:

Enthalpy of unfolding is a linear function of temperature (Figure 5a), but entropy is a non-linear function (Figure 5b). Both the parameters change its sign from positive to negative when temperature is reduced. In other words, during cold denaturation of the protein both enthalpy and entropy produces negative values, in contrast to the heat denaturation process for which both show positive values. It means that high temperature denaturation proceeds with heat absorption and the low temperature process proceeds with heat release. Hence, the cold denaturation of protein can be implied as enthalpy-driven and the heat denaturation as entropy-driven.[202] The decrease of entropy upon breakdown of the ordered native structure of protein molecules sounds like a paradox.[203] It emphasizes that a different molecular mechanism is required to explain the enthalpy-entropy compensation during low-temperature denaturation.

4.5.4a *Enthalpy Changes*: Polar and nonpolar residues react differently in terms of hydration enthalpy with temperature change. Although at room temperature both groups have negative entropy, the magnitude increases with increasing temperature for polar groups, but decreases for nonpolar groups. The total enthalpy is dominated by van der Walls contacts between nonpolar groups and

their hydration; thus the resultant unfolding enthalpy decrease in magnitude with increasing temperature[206] (Figure 5a).

4.5.4b *Entropy Changes*: Total entropy change of protein unfolding includes two components: [204,205] a positive conformational entropy change of the solvent-free polypeptide chain and a negative entropy change due to hydration of buried amino acid residues. Since the hydration entropies of both polar and nonpolar groups are negative at low temperatures, entropy change for cold denaturation process is negative.[207] In heating, the magnitude of hydration entropy for nonpolar residues decreases and for polar residues increases. But the slope of increase of polar residues entropy is smaller that that of nonpolar residues. Therefore, the total hydration entropy which is large and negative at low temperature decreases in magnitude with increasing temperature. Ionization of protein during unfolding imparts entropy effects. But it is negligible compared with the total unfolding entropy. Denaturation configurational arrangement also raises entropy change. This entropy decreases with increasing temperature, the total entropy, however, is dominated by the hydration energy change (Figure 5b).

4.5.4c *Gibbs Free Energy Change*: Cyt *c* shows maximal stability at 288 K, where the free energy of unfolding is higher. Above and below this temperature, the stability of the protein decreases. At the temperature of maximal stability the entropy of protein unfolding is zero and the native stabilization is achieved only by the enthalpy factor (Figure 5). At lower temperatures, decrease in the stabilization effect of hydrogen bonding and hydrophobic interactions reduce the stability. At higher temperatures, increase of the dissipative forces destabilizes the native conformation[207]. These effects are depicted by the parabolic shape of Gibbs free energy function plotted against temperature (Figure 5c).

4.5.5 *Heat Capacity and Hydrophobic Interactions*:

Protein denaturation is associated with positive heat capacity. Linear analysis of enthalpy at different temperatures gives heat capacity change (ΔC_p) for unfolding. For cyt *c*, ΔC_p corresponds to 1.014 kcal mol^{-1} K^{-1} (Figure 6). The value obtained from these studies is slightly lower than the values extracted from other techniques (Table 2, Chapter 3). The linear change of enthalpy presumed for a long range of temperature might have caused such a difference in estimating ΔC_p value.

Transfer of aliphatic compounds into water is accompanied by an increase in heat capacity.[170] Moreover, heat capacity change of denaturation is linearly correlated to the solvent accessible nonpolar surface area (ΔASA_{np}). [182,184] ΔC_p linearly correlates also to the number of amino acid residues in the protein, probably due to the tendency of large globular proteins to bury a greater fraction of the total surface than small proteins.[196] These observations suggest that the heat capacity change upon protein unfolding are defined by the nonpolar groups of the protein buried in the native conformation and become exposed to solvent in the unfolded state. Therefore, heat capacity change directly corresponds to the change in hydrophobic interactions of protein during denaturation. This predicts that proteins with more hydrophobic interaction have higher cold denaturation temperatures. Protein interior also consists of polar groups, the exposure of such groups result in negative heat capacity.[208] This suggests that the use of both polar and nonpolar ΔASA can describe adequately the heat capacity change upon unfolding.[184] As exact molecular description of hydrophobic force has not been defined, more studies will be needed to explain the mechanism of thermal denaturation of proteins. In addition, hydrophilic interactions and hydrogen bonding energies are also vital contributors to the stabilization of the native state, which makes the understanding of protein stabilization more ambiguous.

4.5.6 *Water Structure and Thermal Denaturation*:

Interaction of water molecules has vital impact on the stabilization and dynamics of proteins. Cyt *c* consists of nearly 106 hydrogen bonds and ~337 water molecules in native state. Ordered water structure of the protein is modified during denaturation. At low temperatures, exposure of the nonpolar surface area optimizes the H-bond network. But significant enthalpy gain over unfavorable entropy loss directs protein denaturation. At high temperatures increasingly disordered fluctuating water molecules cause denaturation. The hydrophobic effect and water structure may be explained by two different states of water at low (Ice-Ih type) and high (Ice-II type) temperature.[179]

4.5.7 *Prediction of Cold Denaturation Temperature*:

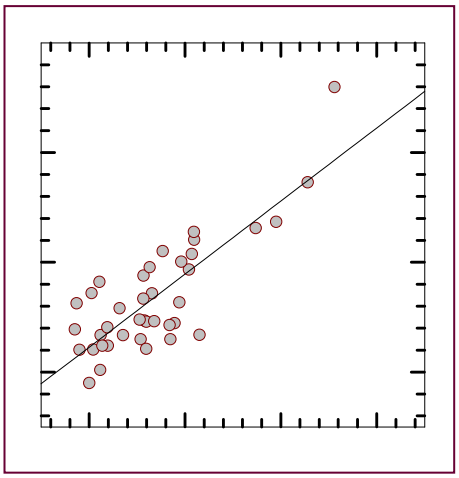
Addition of denaturant to cyt *c* changes thermodynamic characters of its thermal denaturation (Figures 7 and 8). Enthalpy and entropy change of unfolding for both heat and cold denaturation displays linear relation with denaturant concentration. For cold denaturation, enthalpy and entropy is negative and the magnitude of their values decrease with increasing GdnHCl. For heat denaturation, these values are positive and their magnitudes decrease with increasing GdnHCl. The transition midpoint temperatures extracted from linear extrapolations of these parameters to zero denaturant concentration are 234.8 K for cold and 331 K for heat denaturation. T_m^{heat} value is much closer to the value obtained from the DSC experimetnt.[201] But the temperature for cold denaturation (~ -38^{o}C) for native cyt *c* is not practically accessible without addition of solutes. Still these experiments promise a reliable prediction of T_m^{cold} for cyt *c* using the basic thermodynamic relations.

4.6 Conclusion:

Free energy loss above and below the maximal stable temperature predicts the cold denaturation of cyt *c*. However, the cold denaturation is not directly measurable; it has been obtained from slightly denatured protein with GdnHCl. The results propose that different molecular mechanisms must be followed to understand cold and heat denaturation, and emphasize the collective nature of hydrophobic and hydrophilic groups in stabilizing the native conformation of proteins. It also predicts the cold denaturation temperature for native cyt *c*.

Chapter 5

Empirical Relations to Predict Folding Rates of Small Two-state Proteins



5.1 Abstract

Delineating the determinants of folding of small proteins is essential for decoding the folding code. By considering the literature data for the folding of 45 two-state proteins that differ widely in sequence, structure, and function, this work suggests two empirical relations for the prediction of the folding rate. One relation is based on the content of secondary structure elements, and the other uses the number of residues of each of the following types: hydrophobic, positively charged, and negatively charged. Both relations incorporate the chain length as an indirect descriptor. The correlation between experimental values for folding rates and free energy is poor, providing little support to the "stability gap" hypothesis based on the folding of simple lattice and off-lattice polymers. The average rate-determining barrier height for these two-state proteins is much larger than the small barriers for the transition-state ensemble envisaged by theoretical models.

5.2 Introduction

Excitement in the protein folding field over the past 20 years has been fueled by very significant development along several lines of research. The most notable examples include atomic-level characterization of temporally resolved folding intermediates for proteins that apparently fold via multistate kinetics,[87,100] the finding of a substantially large set of two-state or class II proteins[209] for which no structural intermediate is populated to a detectable level,[210-212] and the emergence of theoretical models inspired by the folding of lattice polymers.[213-216] The idea that obligates intermediates and the observation of straight two-state folding appear mutually restrictive. While folding models that engage kinetic intermediates sequentially[111,217] seem to provide clues to Levinthal's conformational search enigma, the volume of experimental data for two-state

proteins raises doubts that the accumulation of kinetic intermediates resolves the problem. The theoretical landscape model leaves it to the sequence-based funneled organization of the energy landscape that dominates the folding kinetics.[218] Folding on a sufficiently smooth funnel is rapid with no traps, the analogue of the phenomenological two-state model. Folding on a rugged funnel would appear to be multistate, since discrete intermediates, labeled "frustrated" or "misfolded", appear trapped at much later stages of folding.[65,68]

A unified mechanism is needed for the conceptual understanding of the folding problem. Theory and simulations suggest that the funnel landscape does provide a unified picture, because the great variety of detailed mechanisms at work in the funnel influences the folding trajectories.[218,219] However, conflicts between experimental observations and theory-based tenets have been found.[212,220,221] Another basic problem with the funnel perspective is that it does not attach much significance to the decoding of the primary structure in the three-dimensional native structure.[27] In a recent work, Kuwajima and co-workers have attempted to unify the folding mechanisms, in a classical sense, of non-two-state and two-state proteins.[222]

One approach to understanding the basic mechanism of folding, first taken up in Baker's laboratory,[223] is to define the determinants of folding kinetics of small two-state proteins by considering the available literature on sequence, structure, and folding. For a data set of 12 proteins, percent contact order was found to correlate strongly with folding rate constants.[223] A later analysis using 24 proteins showed a stronger correlation between relative contact order and folding rate.[75] Several related studies, including the prediction of folding rates based on the first principles of protein folding,[294] calculation of folding rates from three dimensional structure,[225] and correlation of folding rates with stability and contact order[226], long-range order,[227,228] and total contact distance,[229] have been reported

since then. In the most recent study of this kind, transition-state contact orders for 10 proteins determined by molecular dynamics simulation have been found to correlate with folding rates.[230] In a somewhat different approach, prediction of folding rates by analysis of local secondary structure content for a set of 24 proteins has been described.[231]

The scaling of folding times, τ_f, with protein size has also been considered. While theoretical studies have projected the predictive value of the number of residues in the protein, N, through the simple relation $\log(\tau_f) \approx N^{\beta}$, in which the value of β lies in the range of 0.5-0.67,[232-234] a recent database analysis of various size proteins and peptides with sizes ranging from 16 to 396 residues has shown that $\beta = 0.5$.[235,236] This scaling simply involves the folding rate and the number of residues and does not consider any structural or chemical details of the proteins. One is interested to know how the folding rates of simple benchmark proteins are correlated with the details of secondary structure and random coil content, the content of amino acid residues classified according to their chemical nature, and the total number of residues in the proteins. Equally interesting is the dependence of the equilibrium surface exposure of residues in global unfolding (m_{eq}) and the approximate position of the folding transition state along the reaction coordinate ($\alpha^{\#}_f$) on protein length.

This chapter considers the experimental folding parameters in a data set of 45 two-state proteins (Figure 1 and Table 1) and establish empirical relations between chain length and equilibrium m value, between chain length with percent secondary structure elements and folding rate, and between chain length with classified residue types and folding rate. The correlations are sufficiently strong to enable prediction of folding rates from the primary sequence data. The data have also been viewed from the energy landscape perspective. Consistent with earlier finding of Plaxco et al.[75] and others,[211,237] the native-state stability is poorly

correlated with the folding rate. Additionally, there is no considerable correspondence between theory and experiments regarding the size of the barrier energy, although the location of the classical transition-state barrier along the folding coordinate derived from experiments agree fairly well with the theoretical prediction of the region for the transition-state ensemble.

Table 1. Proteins used in this study. The labels correspond to those shown in Figure 1. Additional information is provided in the Appendix.

Label	Protein	PDB code	Structure type
1	λ-repressor	1lmb	α
2	ACBP bovine	1hb6	α
3	ACBP rat	Models	α
4	ACBP yeast	Models	α
5	ACBP *Plasmodium falciparum*	1hbk	α
6	Ferricytochrome *c* horse (oxidized)	1hrc	α
7	Ferrocytochrome *c* horse (reduced)	1giw	α
8	Ferricytochrome *c* yeast (oxidized)	1yic	α
9	Ferrocytochrome *c* yeast (reduced)	1ycc	α
10	Psb D	2pdd	α
11	Im 9	1imq	α
12	Cytochrome b_{562} reduced	1m6t	α
13	Engrail homeo domain	1enh	α
14	Vilin head piece	1vii	α
15	Tendamistat	2ait	β
16	Csp *B. subtilis*	1csp	β
17	Csp *B. caldolylicus*	1c9o	β
18	Csp *B. Thermotoga Maritima*	1g6p	β
19	Csp A	3mef/1mjc	β
20	α-spexctrin	1aey	β
21	Src SH3 domain	1srl	β
22	SH3-Fyn	1a0n	β
23	9-Fibronectin-III	1fnf -3	β
24	10 Fibronectin-III	1fnf –4	β
25	Twitchin	1wit	β

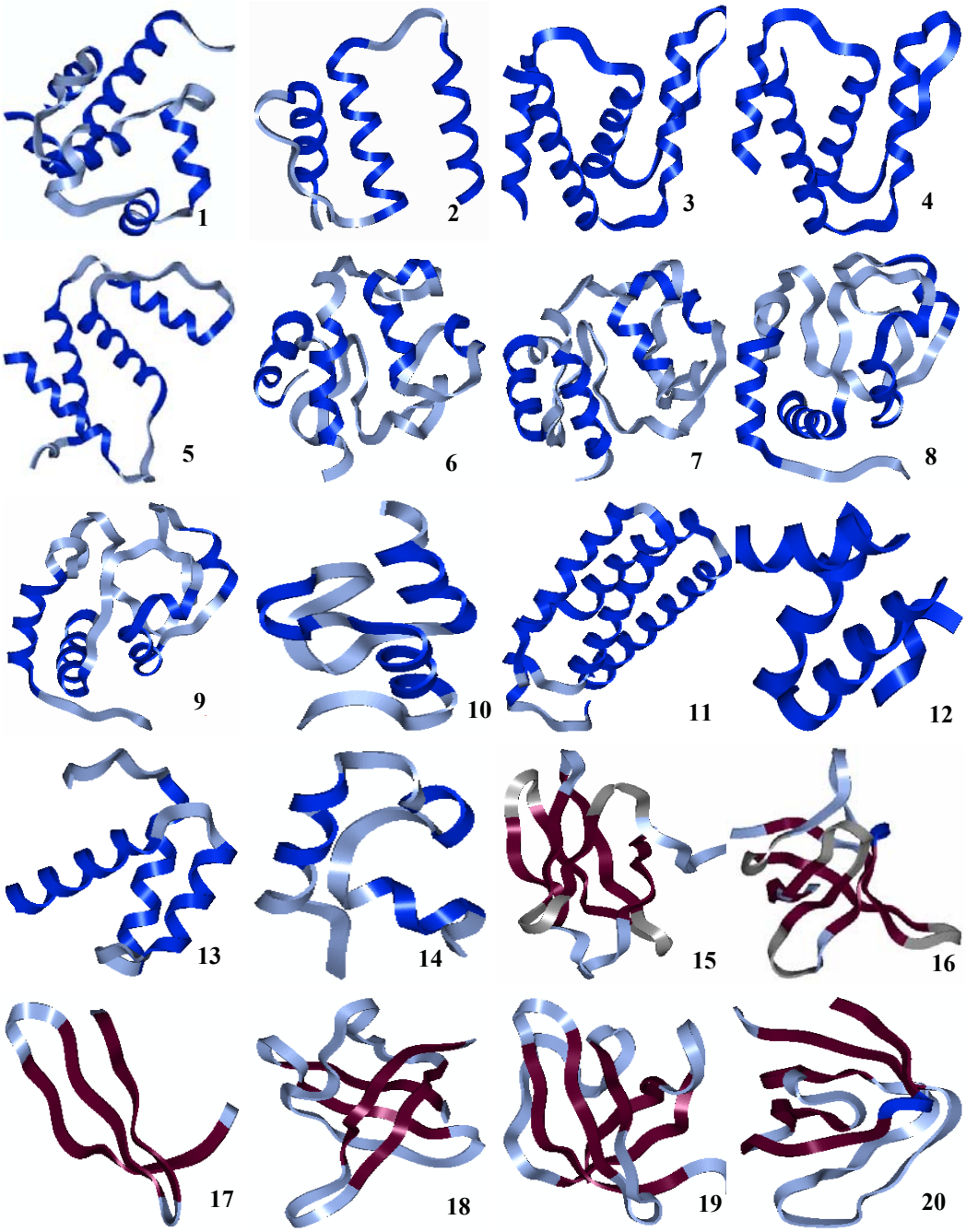
continued…

continued…

26	Tenascin (short)	1ten	β
27	Tenascin (long)	-	β
28	Psa E	1pse	β
29	Ti-I27	1tit	β
30	SH-3 pl 3-kinase	1pks	α and β
31	CI2	2ci2	α and β
32	Procarboxy Peptidase A2	1pba	α and β
33	Arc repressor	1arr	α and β
34	Ubiquitin	1ubq	α and β
35	Protein L	2ptl	α and β
36	Protein U1A	1urn	α and β
37	Hpr	1hdn	α and β
38	FKBP 12	1fkb	α and β
38A	FKBP 12	1fkb	α and β
39	Muscle AcP	1aps	α and β
40	Villin 14T	1vik	α and β
41	Che Y	3chy	α and β
42	B1-domain ProteinG	1em7	α and β
43	Protein L9 (1-56)	1div	α and β
44	Sso 7D	1bnz	α and β
45	Mer P	1osd	α and β

5.3 Materials and Methods

Forty-five two-state proteins from different families (Table 1) with different secondary structures were taken for the analysis. Structural properties of the proteins were obtained from Protein Data Bank. All thermodynamic and kinetic properties were collected from the literature. Secondary structure propensities were calculated from the PROSS program developed in the laboratory of G. Rose. Contact orders were obtained from Plaxco's Contact Order Calculator. The amino acid composition and numbers of hydrophobic, positively charged, and negatively charged residues were calculated with SAPS.


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

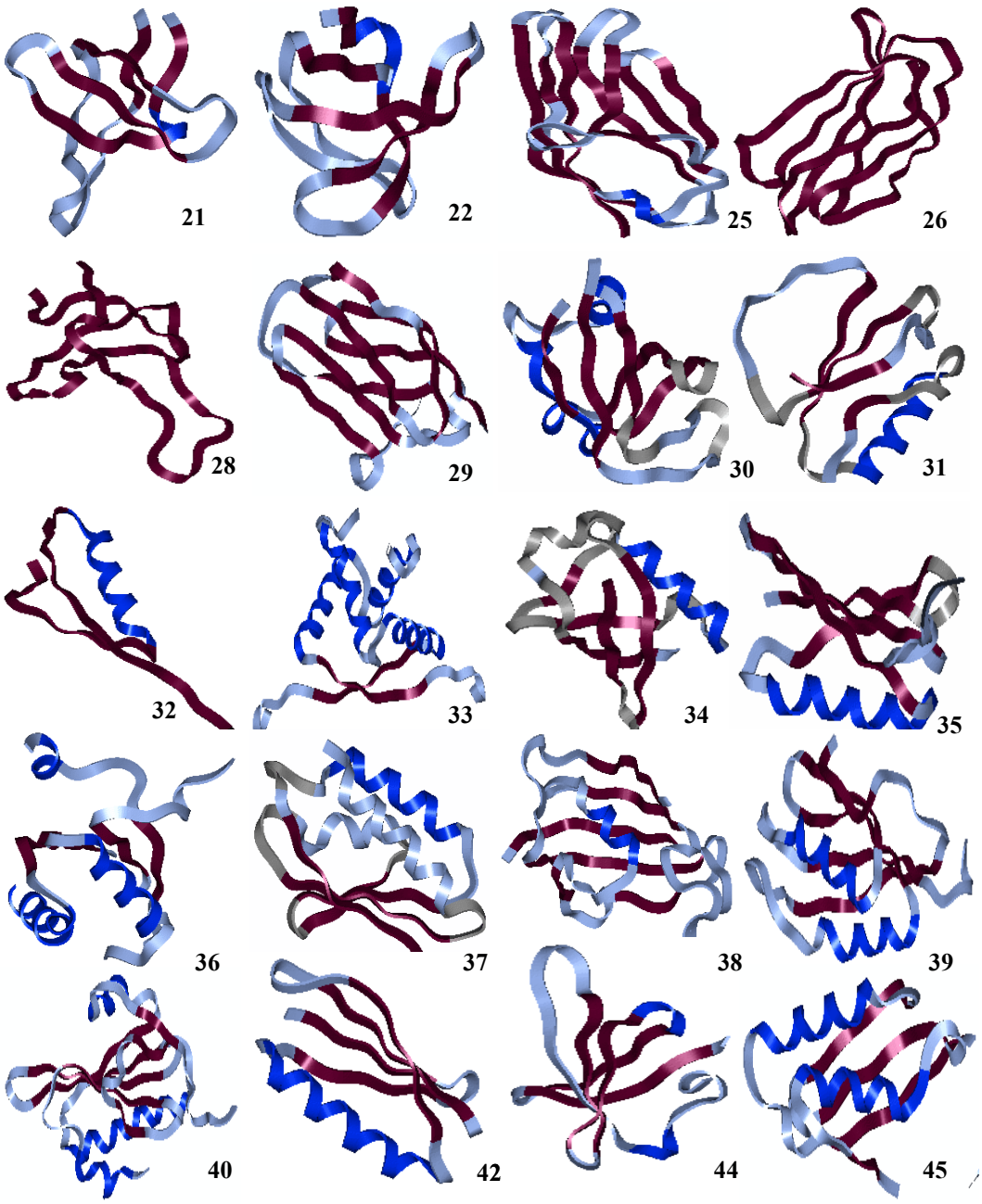


Figure 1. Ribbon diagrams of the proteins listed in Table 1.

5.4 Results

5.4.1. *Length of Two-State Proteins and Folding Parameters*:

The number of amino acid residues (N) is not directly correlated with any of the folding parameters, namely, equilibrium surface exposure of residues during unfolding (m_{eq}), free energy of unfolding (ΔG^{o}), rates of folding and unfolding (k_f and k_u, respectively),and position of the folding transition state($\alpha^{\#}_f$). N may have an indirect contribution to all or some of these parameters. For example, Figure 2 shows a linear analysis of N and m_{eq} for 44 proteins by the empirical relation,

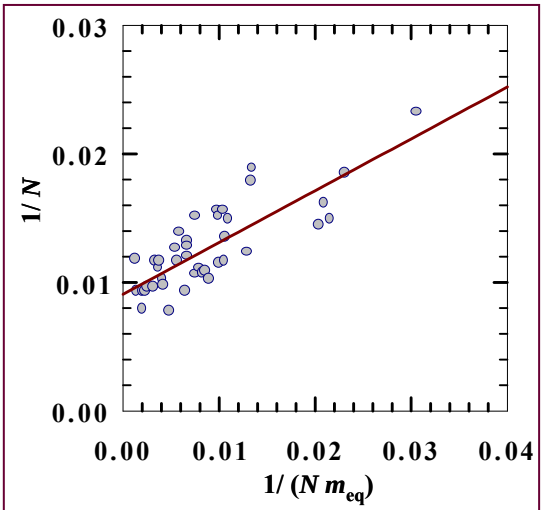


Figure 2. Correlation between $1/N$ and $1/(Nm_{eq})$ for 44 proteins given by equation 1. The solid line is the best fit for these data.

$$\frac{1}{N} = a + \frac{b}{Nm_{eq}} \qquad (1)$$

where a = 0.009 and b = 0.403. Both the correlation coefficient (r = 0.81) and the statistical significance (p<0.0001) are fairly good, suggesting that N could be an indirect determinant for folding parameter values.

5.4.2. *Residue Type and Folding Rate*:

Percentage of amino acid residues forming each secondary structure elements, α-helix, β-strand, β-turn, polyproline-II, and random coil was calculated using the PROSS program, which is based solely on backbone dihedral angles.[231] The correlation between secondary structure content with random coil and experimentally observed folding rates for 39 two-state proteins in the data set was analyzed. The correlation coefficient was poor (r=0.58). When the total number

of residues was introduced as an indirect determinant the correlation coefficient

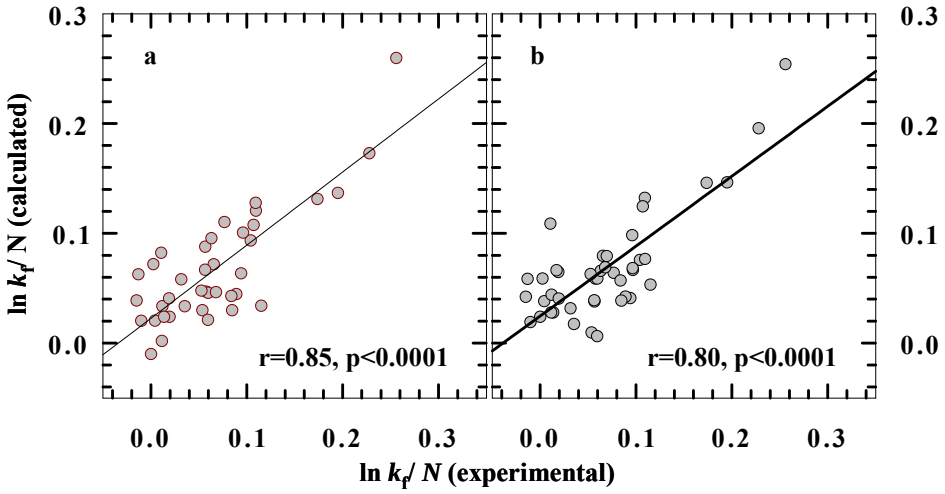


Figure 3. Predicted and experimental folding rates. (a) Prediction based on secondary structure and random coil content for 39 proteins, with correlation coefficient from equation 2. (b) Prediction of folding rates on the basis of hydrophobic, positively charged and negatively charged residues for 43 proteins. The coefficients are from equation 3.

increased. Figure 3a shows plot of the result. The best-fit linear relationship between (ln k_f)/N and the percentages of residues forming different secondary structural element is

$$\frac{\ln k_f}{N} = -0.136(\pm 0.0802) + 0.0013(\pm 0.008)a + 0.007b - 0.001c + 0.0016(\pm 0.0014)d + 0.001e + \frac{10.7374(\pm 1.5193)}{N} \qquad (2)$$

where a-e are percentages of residues forming α-helix, β-turn, β-strand, polyproline-II, and random coil, respectively. The numbers in parentheses represent standard errors associated with the coefficients. The correlation is strong (r=0.85) with a good level of statistical significance (p<0.001).

In search for a residue-type trend, amino acids were classified as hydrophobic, positively charged, and negatively charged. The linear correlation between the sum of numbers of classified residue types and the folding rate was poor (r=0.43). When the chain length is incorporated as a descriptor, the correlation is fairly strong (r=0.80, p<0.0001). The best-fit relationship between (ln k_f/N) and the number of different residue types of 43 two-state proteins was found to be:

$$\frac{\ln k_f}{N} = -0.2094(\pm 0.0712) + 0.0001l + 0.0043(\pm 0.0015)m + 0.0023(\pm 0.0018)n + \frac{15.3449(\pm 2.62)}{N} \qquad (3)$$

where l, m, and n are the number of hydrophobic, positively charged and negatively charged residues, respectively (Figure 3b).

5.4.3. *Secondary Structure Based Classification and Analyses*:

Previous studies have shown that contact order,[223,230] relative contact order[75] or secondary structure content[231] can predict the folding rates of two-state proteins. These studies were however performed with data for less number of proteins (Table 2). These relations were verified here with a relatively large set of data. To analyze the contact order-rate relationship, 29 proteins were taken, 16 from the data set of 24 proteins reported previously[75] and 13 from this data set. The analysis showed poor linear relationships between ln k_f and contact order and between ln k_f and relative contact order (Table 2).

All the relations predicted were analyzed separately for proteins classified according to their secondary structure. It does not improve the correlation between any of the parameter and folding rate. But, the analysis of the linear relation between rate and secondary structure content or rate and hydrophobic and charged residue types shows reasonably good correlation coefficients.

Table 2. Predicted and Experimentally Observed Folding Rates

		r value from this work (Table 1)				literature		
Independent variable(s)	Dependent variable	all proteins	all α	all β	α+β	no. of proteins	*r*-value	ref
Secondary structure and random coil content and N	$(\ln k_f)/N$	0.85	0.97	0.83	0.77			
No. of hydrophobic, positive and negatively charged residues and N	$(\ln k_f)/N$	0.80	0.95	0.68	0.80			
Secondary structure content and N	$\ln k_f$	0.58	0.60	0.79	0.34	24	0.91	31
Contact Order	$\ln k_f$	0.62	0.68	0.64	0.44	12	0.81	22
Relative contact order	$\ln k_f$	0.63	0.68	0.64	0.44	24	0.92	23
ΔG^o	$\ln k_f$	0.31	0.18	0.30	0.52	24	0.40	23

5.4.4. *Analysis of "Standard" Data Set*:

The lower records of the strength of the linear relationships could be due to a gross error of measurement or an error in recording data for some of the proteins in the current data set, producing many outliers in the final correlation plots. The data for various proteins comprising the data set used were acquired under various pH, temperature, denaturant and solvent conditions. To exclude the effect of different experimental conditions for different proteins, the correlation analyses were carried out using data from the recently reported kinetic data set of 30 two-state proteins that were studied under a "standard" set of experimental conditions.[240]

Of these 30 proteins, only those that also figured in the data set of this study were chosen. The contact order versus ln k_f and the relative contact order versus ln k_f relationships were further analyzed using 19 proteins, all from the standard data set of Maxwell et al.[240] Table 3 lists the correlation coefficients that were found. The coefficients for these two determinants are poor again.

Table 3. Analyses Using the "Standard" Two-state Data Set Reported by Maxwell et al.[240]

Independent variable(s)	Dependent variable	*r* value	no. of proteins
Secondary structure and random coil content and *N*	$(\ln k_f)/N$	0.81	14[a]
No. of hydrophobic, positive and negatively charged residues and *N*	$(\ln k_f)/N$	0.50	14[a]
Secondary structure content and N	$\ln k_f$	0.78	14[a]
Contact order	$\ln k_f$	0.79	11[a]
Relative contact order	$\ln k_f$	0.69	11[a]
ΔG^o	$\ln k_f$	0.26	13[a]
Contact order	$\ln k_f$	0.33	19[b]
Relative contact order	$\ln k_f$	0.49	19[b]

[a] Number of proteins common to this data survey and the experimental data set of Maxwell et al.
[b] Number of proteins taken from the study of Maxwell et al.

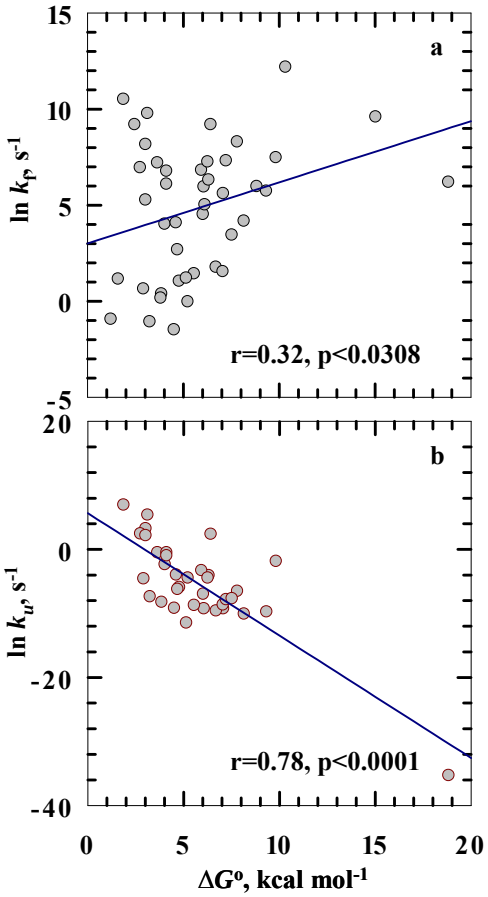


Figure 4. Test of stability hypothesis with data for 45 proteins. (a) Natural log of refolding rate is very poorly correlated with the free energy of folding. (b) The correlation between natural log of unfolding rate and ΔG^o is fairly good. The straight lines represent the best linear fits for data.

5.4.5. *Free Energy and Rates*:

Funnel model of protein folding essentially envisages that folding speed of a protein is directly related to the free energy gradient between folded and unfolded form, or the funnel depth. To examine this hypothesis refolding rates (ln k_f) of 45 two-state proteins were linearly correlated with the free energy of unfolding (ΔG^o). The correlation is very poor (r=0.32, Figure 4a) indicating that folding speed is hardly dependent on the protein stability. Surprisingly, the observed rate constant for the unfolding in water is considerably correlated with ΔG^o (Figure 4b), apparently suggesting that higher equilibrium stability retards the unfolding rate.

5.4.6. *Activation Energy of Folding and Unfolding*:

By assuming that the apparent activation energies for folding and unfolding in water are given by an expression of the type E_a=-RT ln(k/A_o), where k is the rate coefficient (k_f and

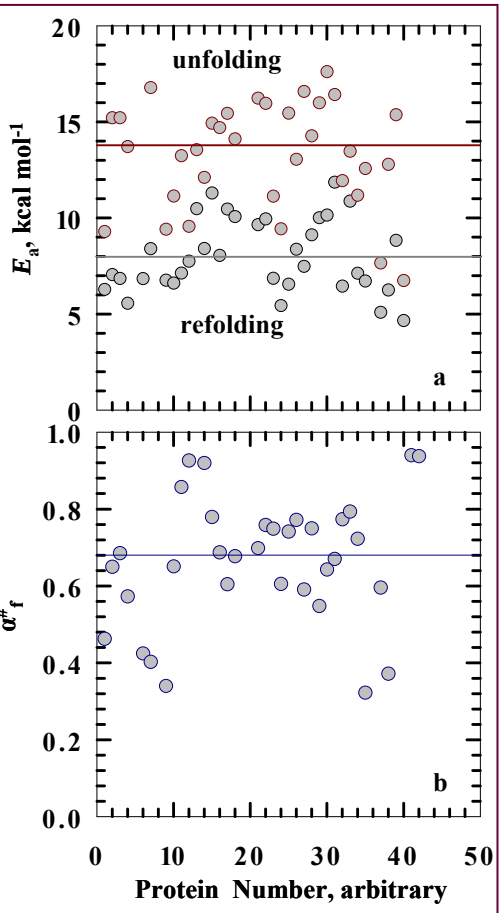


Figure 5. (a) Activation energies of folding (○) and unfolding (○) for 42 and 35 two-state proteins, respectively. (b) Barrier locations shown for 41 proteins. The solid lines in both panels show the mean values.

k_u) and A_o is the front factor, the heights of activation barriers for folding and unfolding can be estimated, although the appropriateness of the use of this thermally activated rate law in its general form is not quite clear.[56,57] It is assumed that A_o as the time needed for two residues of the unfolded chain to come together to form a contact or geminate pair. The recent ultrafast folding experiment has shown that for ferrocyt *c* the diffusion time constant for formation of a contact between two regions separated by 46-60 residues is ~400ns.[38] The diffusion times for the first steps in formation of short turns, loops and helices fall in the 5-40 ns range.[61,245] From these considerations, the value of A_o can be set in the range of 10^7-10^8 s^{-1}. Using an A_o of 1×10^8 s^{-1}, E_a values associated with refolding and unfolding reactions for 42 and 35 proteins, respectively, were estimated (Figure 5a). The barrier heights average to 7.92 ± 2.1 and 13.3 ± 2.77 kcal/mol, respectively. These averages decrease by 1± 0.5 kcal/mol when an A_o of 1×10^7 s^{-1} is used to estimate the barrier sizes.

5.4.7. *Position of Transition State*:

Classically, an approximation of the extent of structure formation in the transition-state ensemble (TSE) along the folding coordinate is made from the relation $\alpha = m_{f(u)}^{\#}/m_g$, where $m_f^{\#}$ and $m_u^{\#}$ are kinetic m values ($=2.3RT\,\partial\log k_{f(u)}/\partial[\mathrm{GdnHCl}]$) associated with refolding and unfolding respectively.[63] The average α values, estimated from folding data for 41 two-state proteins, is 0.63 (Figure 5b). This average must be taken at a coarse level, since curvature in the chevron limbs for some two-state proteins introduces large errors in the estimation of $m_f^{\#}$. Nevertheless, it does provide the general idea that the folding transition states for small proteins structurally resemble the native states to an extent of ~ 60%, suggesting that roughly two-thirds of the surface area that is buried in the native state becomes buried in the transition state.

5.5. Discussion

5.5.1. *Protein Length and Solvent Accessible Surface Area*:

A very strong correlation between N and the calculated change in total surface area in unfolding ($A_u - A_n$, where A_u and A_n are total solvent accessible surface areas in unfolded and native states, respectively) has been shown for a mixed set of two- and higher-state proteins.[184] Then, a linear correlation of N with m_{eq} is also expected, since m_{eq} is directly proportional to ($A_u - A_n$).[151] The reason this direct correlation is not observed here to any significant extent (r=0.37, Figure 2) must be due to different levels of unfoldedness and chain configuration for different proteins. Incomplete chain expansion and the persistence of loops or local structural elements in the unfolded state affect the total surface area accessible to water. Equation 1 provides a means of estimating the effective m_{eq} value based on the length of any two-state protein.

5.5.2. *Prediction of Folding Rate*:

Although a correlation between the folding rate and the chain length has been reported for Gộ model proteins,[238] no direct relationship is found between the length and the folding rate of naturally occurring two-state proteins, consistent with earlier reports.[75,239] Here, a reliable linear correlation has been obtained for folding rate with percentages of residues forming secondary structural propensities (Equation 2, Figure 3a). A similar analysis using a set of 24 two-state proteins has been reported, where % helix, % turn, and % hairpin are used in conjunction with chain length.[231] The analysis of this study also considers % polyproline II and % random coil. This has improved the correlation when relatively large number of proteins is taken to predict the rate. In addition, folding rate is also predicted from the number of hydrophobic, positively charged and negatively charged residues (Figure 3b, Equation 3).

These empirical relations provide bases for predicting folding rates from sequence alone, without a detailed knowledge of the three-dimensional molecular structure. While the number of hydrophobic and charged residues can be obtained by inspection of primary structure, the fraction of the sequence in each secondary structural element can be calculated by any available secondary structure identification programs, PROSS, for example. It is notable that equations 2 and 3 provided to calculate folding rate of two-state proteins use the protein length as an indirect descriptor and are clearly distinct from the general relation $\log(\tau_f)=N^{\beta}$ that scales the folding time simply with the number of amino acids.[233,234] By using a β of 0.5 as prescribed,[235] one finds r=0.6 for 36 of the two-state proteins considered here. Clearly, the folding rate is predicted well when N is used in conjunction with secondary structural content or residue type details.

Proteins were classified according to their secondary structural content and analyzed. This analysis does not increase the correlation coefficient

effectively for the relations such as contact order-rate, relative contact order-rate, and secondary structure content-rate. However, the predicted rates from equations 2 and 3 produce reliable linear correlations with the experimental values for this analysis (Table 2). This fact argues that formation of different secondary structure contents influences the folding rate. Inclusion of a parameter for secondary structure content with topological parameter like contact order may increase the correlation coefficient, thus can effectively predict the folding rate. The lower correlation coefficients may then arise from those proteins for which chain topology is not a critical determinant of folding kinetics. These proteins may be the outliers in the final correlation plots. Indeed, mutations that do not significantly alter the topology, and hence the contact order, may affect the folding rate.[75]

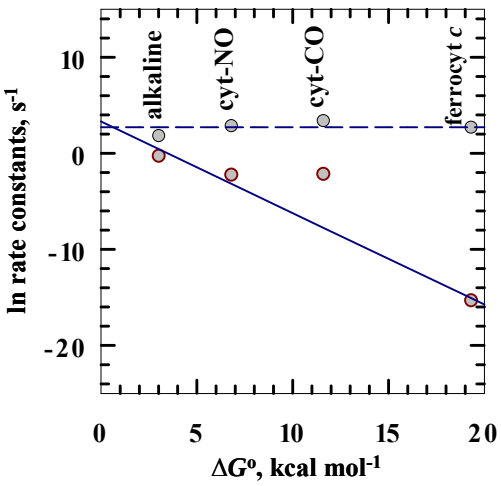


Figure 6. Correlation of natural logarithm of folding and unfolding rates with ΔG^o for four structurally very similar test tube variants of horse ferrocyt *c*: (○) folding rate and (○) unfolding rate.

5.5.3. *Experimental Results and Theoretical Predictions*:

Numerous theoretical and computer simulation studies have endeavored to provide a general frame work, popularly called funnel perspective, for the understanding of the protein folding problem. Just as the fast folding lattice- and simplified off-lattice polymer models[213-216] results from a smooth energy landscape, one can imagine a smooth funnel for rapid folding of a perfect two-state folder. As folding progresses, the energy of the folding ensemble decreases in a

smooth and monotonic manner all the way to the bottom of the funnel. The large data set compiled in this work for two-state proteins provides an opportunity to check the extent to which the predictions of funnel-guided folding properties are met.

5.5.3a. *Funnel Depth and Folding Speed*: The correlation between folding free energy and folding rate is very poor (r=0.32) whereas unfolding rate in water significantly correlates with ΔG^o (r=0.78, Figure 4). These relations suggest that folding speed is hardly dependent on the stability and higher equilibrium stability retards the unfolding rate of a protein. An extreme case of this relationship is provided by the folding-unfolding properties of the four test tube variants of ferrocyt *c* that have already been described in Chapter 1. Figure 6 shows the dependences of ln k_f and ln k_u on ΔG^o for ferrocyt *c*, alkaline ferrocyt *c*, carbonmonoxyferrocyt *c* (cyt-CO)and nitrosylferrocyt *c* (cyt-NO). They share identical primary and secondary structures, and very similar native-state tertiary structures, but are far apart in terms of equilibrium stability.[212,221] It would then appear that the natural sequence rather than the stability controls the folding speed. On the other hand, it is the native-state stability that dictates the unfolding speed.

The energy landscape approach associates the fast folding speed with the "energy gap" between the native or native-like states and the set of states with little structural similarity to the native state.[5,67,70] Translation of this contextual energy gap, called the "stability gap" by Bryngelson et al.,[5] to ΔG^o ($=G_U$-G_N) would associate higher equilibrium stability with larger free energy gradient or energetic drive leading to the native state, especially when no glassy states exist, and hence rapid folding. The energy gap hypothesis, in a subtly different form, has also been advanced by others,[241,242] but the effective predictions are the same. The data survey indicated that the folding speed is not correlated with the free

energy gradient or the funnel depth. Studies show that proteins from the same family can produce a good correlation between the folding rate and free energy.[75,243] When all the proteins are considered together, the relation is not significant. Thus, even though the rapid folding of simple lattice polymers may be related to the size of the stability gap, the folding speed of natural two-state proteins is effectually independent of the equilibrium stability. It should be mentioned that a better measure of foldability, for lattice models of proteins at least, is the *Z*-score parameter that estimates the energy of the native conformation relative to the average energy of compact nonnative conformations.[244] Folding rate of small natural proteins may show good correlation with *Z*-scores. The test of the hypothesis would require experimental determination of the "relative energy of the native conformation".

The energy gap rather appears to be a determinant of the unfolding rate of these small proteins (Figures 4 and 6). The considerable correlation between unfolding rates and protein stability may reflect the bearing of the native-state stability on the unfolding activation barrier. It seems the differences in the rate-limiting unfolding barrier heights are similar to the differences in the equilibrium stabilities for these small proteins. This proportionality requires that the unfolding transition state be native-like structurally, and presumably thermodynamically, so that the energetic stabilization of the transition state is controlled by the native-like stability. More experimental data will be needed to test this conjecture.

5.5.3b. *Barrier Heights*: Activation energies for folding and unfolding reactions are estimated from the Arrhenius relation (Figure 5a). The barrier height is ~7.9 kcal/mol for folding and ~13.3 kcal/mol for unfolding. The landscape perspective is not strict about such barriers. Folding in an extremely smooth funnel may not encounter any significant barrier;[66] the retardation of folding here arises from the integral over a number of changing dynamic processes, including motional

diffusion, as folding progresses.[4,5] According to simulation studies, general downhill folding may encounter, at later stages of folding, tiny barriers (~$3k_BT \approx$ 1.68 kcal/mol at 10 ^oC)[5,65] that are significantly smaller in magnitude than the average value of 7.92 kcal/mol shown by the experimental data.

5.5.3c *Barrier Location*: The mean value for the relative position of refolding transition state ($\alpha_f^{\#}$) is 0.63. The α –value as such does not indicate the amount of native-like structure present in the transition state. However, several studies including Φ value analysis[246] and restrained molecular dynamic simulations[230] show the presence of substantial native-like structures in the folding transition states of many proteins, suggesting that the transition barrier is placed closer to the native state along the folding coordinate.

Mapping of the behavior of lattice simulations to real proteins shows that the transition state ensemble (TSE) of small proteins has 60% of the native contacts.[65,214,247] The same result has been obtained from an all-atom simulation study; the TSE has ~80-85% of its surface area buried.[68,247,248] Thus, a good correspondence between experimental data and theoretical predictions exist for the barrier location.

5.5.4. *Limitations of the Data Set Analysis*:

Understandably, the strength of the equations provided for the prediction of folding rates is subject to the quality of data in the data set that has been used to obtain the equations. Proteins in the database (Appendix) have been studied at various temperatures ranging from 5 to 40 ^oC. Inherent errors in the experiments vary with different techniques. Also, the minimum denaturant concentration used for different proteins in the set is not the same, and the effects of urea and guanidinium hydrochloride have not been considered separately.

5.6 Conclusion:

Although the number of residues in the sequence does not directly determine the folding rate, it can be used as an indirect descriptor for prediction. For this purpose, two empirical relations have been provided: one requires the percentage of residues separately forming α-helix, β-turn, β-strand, polyprolineII, and random coil and the other the number of residues classified as hydrophobic, positively charged, and negatively charged. The size of the classical rate-determining folding barrier is substantially larger than the tiny TSE barriers on the order of few $k_{\mathrm{B}}T$ encountered by the funnel model, although there is a fair correspondence between experiments and theory regarding their location along the folding coordinates.

Appendix 1

Structural, thermodynamic, and kinetic properties of two-state proteins

No	Name	PDB	Secondary structure	*N*	Hyd	+ ve	- ve	*T*	ΔG^0
1	λ repressor	1lmb	α only	87	25	10	12	37	3
2	ACBP bovine	1hb6	α only	86	18	16	17	5	7.06
3	ACBP rat	Models	α only	86	19	17	15	5	6.05
4	ACBP yeast	Models	α only	86	19	13	17	5	7.78
5	ACBP *P.falciparum*	1hbk	α only	89	22	13	13	-	-
6	Ferricyt *c* horse	1hrc	α only	104	21	21	12	10	8.8
7	Ferrocyt *c* horse	1giw	α only	104	21	21	12	10	18.8
8	Ferricyt *c* yeast	1yic	α only	108	21	19	11	-	5.7
9	Ferrocyt *c* yeast	1ycc	α only	108	21	19	11	20	15
10	Psb D	2pdd	α only	43	13	8	5	25	3.1
11	Im 9	1imq	α only	86	20	7	16	10	6.24
12	Cyt b_{562}	1m6t	α only	106	23	16	20	20	10.3
13	Engr. homeo domain	1enh	α only	54	12	13	6	25	1.85
14	Villin head piece	1vii	α only	36	12	6	4	25	2.44
15	Tendamistat	2ait	β only	74	14	4	9	25	8.13
16	Csp *B.subtilis*	1csp	β only	67	20	6	12	25	2.72
17	Csp *B. caldolylicus*	1c9o	β only	66	20	7	9	25	3.63
18	Csp *T. maritima*	1g6p	β only	66	19	11	11	25	6.3
19	Csp *A*	1mjc	β only	69	18	7	8	25	3
20	α - spectrin	1aey	β only	62	17	10	11	25	2.9
21	Src SH3	1srl	β only	64	14	4	7	22	4

Appendix

m_{eq}	ln k_f	ln k_u	$\alpha^{\#}_f$	CO	RCO	α - helix	β – sheet	β – turn	PP-II	RC	Ref
1.15	8.2	3.3	0.46	--	--	68.9	5.75	0	6.9	18.3	249,250
3	5.6	-9.2	0.65	10.8	0.13	68.6	6.98	0	13.9	10.4	251,252
3.4	6.0	-9.2	0.69	--	--	-	-	-	-	-	252
2.03	8.3	-6.5	0.57	--	--	-	-	-	-	-	252
-	-	-	--	10.3	0.12	-	-	-	-	-	211
3	6.0	6.0	0.42	--	--	38.1	19.0	9.52	11.4	20.9	25,253
3.7	6.2	-35.2	0.4	10.5	0.1	21.9	26.7	0	7.62	43.8	30
4.52	-	-	--	--	--	-	-	-	-	-	--
3.96	9.6	9.6	0.34	--	--	36.1	20.3	2.78	14.8	25.7	253
0.76	9.8	5.4	--	4.79	0.11	46.5	11.6	0	13.9	27.9	254
1.1	7.3	-4.4	0.94	10.4	0.12	43.0	10.4	0	2.33	44.1	255
-	12.2	-	--	9.45	0.09	78.3	8.49	0	5.66	7.55	256
0.8	10.5	7.0	--	7.18	0.13	70.3	0	7.41	7.41	14.8	257
-	9.2	-	--	4.02	0.11	30.5	25	0	8.33	36.1	258
1.27	4.2	-10.0	0.65	--	--	0	8.11	32.4	10.8	48.6	259
0.693	7.0	2.5	0.86	--	--	0	17.9	28.3	13.4	40.3	260
1.53	7.2	-0.4	0.93	10.6	0.16	0	18.1	33.3	18.1	30.3	73,261
2	6.3	-4.0	0.86	11.4	0.17	0	0	21.2	4.55	74.2	262,211
0.71	5.3	2.2	0.92	11.0	0.16	0	10.1	31.8	5.97	52.1	263
0.77	0.7	-4.5	0.78	11.5	0.2	0	17.2	43.1	6.9	32.7	264
1.6	4.0	-2.3	0.69	10.9	0.19	0	12.5	28.5	7.14	51.7	265

Appendix

No	Name	PDB	Secondary structure	*N*	Hyd	+ ve	- ve	*T*	ΔG^0
22	SH3 –Fyn	1a0n	β only	67	15	4	13	20	6
23	9 Fibronectin -III	1fnf	β only	90	22	8	9	25	1.2
24	10Fibronectin -III	1fnf	β only	94	22	8	8	25	6.1
25	Twitchin	1wit	β only	93	26	12	12	20	3.83
26	Tenascin - short	1ten	β only	90	22	9	18	20	4.76
27	Tenascin - long	-	β only	92	23	9	18	20	6.68
28	PsaE	1pse	β only	69	20	8	7	22	1.57
29	Ti – I27	1tit	β only	98	28	8	14	25	7.5
30	SH3-Pl 3	1pks	α and β	79	17	8	16	20	3.23
31	CI2	2ci2	α and β	83	26	13	13	25	7.03
32	Procarboxy peptidase A2	1pba	α and β	81	24	7	19	25	4.1
33	Arc repressor	1arr	α and β	53	16	11	7	25	6.4
34	Ubiquitin	1ubq	α and β	76	23	11	11	25	7.2
35	Protein L	2ptl	α and β	78	13	8	17	22	4.6
36	Protein U1A	1urn	α and β	102	33	17	9	25	9.3
37	Hpr	1hdn	α and β	85	24	8	10	20	4.66
38	FKBP12	1fkb	α and β	107	29	14	13	25	5.53
38	FKBP 12	1fkb	α and β	107	29	14	13	25	5.13
39	Muscle AcP	1aps	α and β	98	24	16	10	28	4.49
40	Villin 14T	1vik	α and β	126	30	14	13	25	9.8
41	Che Y	3chy	α and β	128	43	15	19	25	5.2
42	Protein G	1em7	α and β	57	11	8	11	-	-
43	Protein L	1div	α and β	56	15	12	6	19	4.1
44	Sso 7D	1bnz	α and β	64	16	16	10	20	5.92
45	Mer P	1osd	α and β	72	18	9	6	25	3.8

Appendix

m_{eq}	ln k_f	ln k_u	$\alpha^{\#}_f$	CO	RCO	α - helix	β – sheet	β – turn	PP-II	RC	Ref
1.36	-6.9	1.5	0.68	--	--	0	15.5	34.4	10.3	39.6	266
3	--	-0.9	--	--	--	0	8.89	54.4	15.5	21.1	267
1.42	--	1.5	--	--	--	0	10.6	43.6	15.96	29.7	267
1.3	-8.2	1.1	0.7	18.7	0.2	0	13.9	38.7	4.3	43.0	268
1.4	-5.9	1.3	0.76	--	--	0	13.3	52.2	14.4	20	269,270
1.26	-9.5	1.7	0.75	--	--	-	-	-	-	-	269,270
-	-	-	--	11.5	0.17	0	13.0	24.3	8.7	53.6	271
2.5	-7.6	1.1	0.94	15.7	0.18	0	44.9	37.0	4.44	53.9	272
2.33	-7.3	0.3	0.6	20	0.2	0	26.3	27.6	5.26	40.7	273
1.79	-8.6	1.4	0.61	9.94	0.15	13.2	16.8	22.8	10.8	14.4	210
0.96	-0.4	1.4	0.74	--	--	19.7	6.17	12.3	4.94	56.7	274
1.4	2.4	1.5	0.77	2.71	0.05	52.8	7.55	13.2	3.77	22.6	275
1.95	-7.8	1.3	0.59	11.3	0.15	14.4	23.6	42.1	10.	9.21	276
1.9	-3.9	0.9	0.75	9.96	0.13	16.6	3.85	32.0	25.6	44.8	277
2.3	-9.7	1.4	0.55	--	--	32.2	10.4	18.7	15.6	22.9	21
9.2	-6.2	0.8	0.64	15.3	0.18	36.4	8.29	27.0	7.06	21.7	278
1.43	-8.7	1.4	0.67	18.6	0.17	7.48	21.5	32.7	12.1	26.1	278
6.6	-11.4	-0.2	0.77	--	--	--	--	--	--	--	279
1.14	-9.1	1.3	0.79	20.6	0.21	18.3	4.08	22.4	6.12	48.9	280
3.93	-1.8	0.9	0.72	13.5	0.14	0	10.1	50.5	7.07	32.3	281
1.6	-4.4	1.2	0.32	11.1	0.09	46.8	5.47	19.5	7.18	20.3	282
-	-	-	--	10	0.18	25	10.7	41.0	0	23.2	--
1.34	-0.9	1.1	0.6	10.3	0.07	23.2	19.6	7.14	10.7	39.2	283
1.5	-3.2	1.4	0.37	7.64	0.12	14.0	15.6	40.6	9.38	20.3	284
2.34	-	-	-	11.6	0.16	33.3	13.8	25	6.94	20.8	285

References

1.Anfinsen, C. B., Haber, E., Sein, M. & White, F. H. Jr. (1961). *Proc. Natl. Acad. Sci. USA*, **47**, 1309-1314.

2.Anfinsen, C. B. (1973). *Science*, **181**, 222-230.

3.Levinthal, C. (1968). *J. Chem. Phys.* **65**, 41-45.

4.Dill, K. A. & Chan, H. S. (1997). *Nature Struct. Biol.* **4**, 10-19.

5.Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995). *Proteins*: *Struct. Funct. Genet.* **21**, 167-195.

6.Plotkin, S. S. & Onuchic, J. N. (2002). *Quart. Rev. Biophys.* **35**, 111-167.

7.Plotkin, S. S. & Onuchic, J. N. (2002). *Quart. Rev. Biophys.* **35**, 205-286.

8.Serrano, L, Matouschek, A. & Fersht, A. R. (1992). *J. Mol. Biol.* **224**, 805-818.

9.Matouschek, A. & Fersht, A. R. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 7814-7818.

10. Matthews, J. M. & Fersht, A. R. (1995). *Biochemistry*, **34**, 6805-6814.

11. Otzen, D. E., Itzhaki, L. S., elMasry, N. F., Jackson, S. E. & Fersht, A. R. (1994). *Proc. Natl. Acad. Sci. USA*, **91**, 10422-10425.

12. Milla, M. E., Brown, B. M., Waldburger, C. D. & Sauer, R. T. (1995) *Biochemistry*, **34**, 13914-13919.

13. Grantcharova, V. P., Riddle, D. S., Santiago, J. V. & Baker, D. (1998). *Nature Struct. Biol.* **5**, 714-720.

14. Burton, R. E., Huang, G. S., Daughetry, M. A., Fullbright, P. W. & Oas, T. G. (1996). *J. Mol. Biol.* **263**, 311-322.

15. Jonsson, T., Waldburger, C. D. & Sauer, R. T. (1996). *Biochemistry*, **35**, 4795-4802.

16. Martinez, J. C., Pisabarro, M. T. & Serrano, L. (1998). *Nature Struct. Biol.* **5**, 721-729.

References

17. Dalby, P. A., Oliveberg, M. & Fersht, A. R. (1998). *J. Mol. Biol.* **276**, 625-646.
18. Otzen, D. E., Kristensen, O., Proctor, M. & Oliveberg, M. (1999). *Biochemistry* **38**, 6499-6511.
19. Bilsel, O. & Matthews, C. R. (2000). *Adv. Protein Chem.* **53**, 153-207.
20. Daggett, V. (2002). *Acc. Chem. Res.* **35**, 422-429.
21. Silow, M. & Oliveberg, M. (1997). *Biochemistry*, **36**, 7633-7637.
22. Sánchez, I. E. & Kiefhaber T. (2003). *J. Mol. Biol.* **327**, 876-884.
23. Sosnick, T. R., Mayne, L., Hiller, R. & Englander, S. W. (1994). *Nature Struct. Biol.* **1**, 149-156.
24. Sosnick, T. R., Mayne, L., Hiller, R. & Englander, S. W. (1996) *Proteins: Struct. Funct. Genet.* **24**, 413-426.
25. Chan, C. K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A. & Hofrichter, J. (1997). *Proc. Natl. Acad. Sci. USA*, **94**, 1779-1784.
26. Shastry, M. C. R., & Roder, H. (1998). *Nature Struct. Biol.* **5**, 385-392.
27. Rumbley, J., Hoang, L. Mayne, L. & Englander, S. W. (2001). *Proc. Natl. Acad. Sci. USA,* **98**, 105-112.
28. Bhuyan, A. K., & Kumar, R. (2002). *Biochemistry,* **41**, 12821-12834.
29. Krantz, B. A., Mayne, L., Rumbley, J., Englander, S. W. & Sosnick, T. R. (2002). *J. Mol. Biol.* **324**, 359-371.
30. Bhuyan, A. K. & Udgaonkar, J. B. (2001). *J. Mol. Biol.* **312**, 1135-1160.
31. Kumar, R. & Bhuyan, A. K. (2005). *Biochemistry*, **44**, 3024-3033.
32. Onuchic, J. N. (1997). *Proc. Natl. Acad. Sci. USA*, **94**, 7129-7131.
33. Goldbeck, R. A., Thomas, Y. G., Chen, E., Esquerra, R. M., & Kliger, D. S. (1999). *Proc. Natl. Acad. Sci. USA*, **96**, 2782-2787.
34. Dinner, A. R., Šali, A., Smith, L. J., Dobson, C. M. & Karplus, M. (2000). *Trends Biochem. Sci.* **25**, 331-339.

35. Hoang, L., Bédard, S., Krishna, M. M. G., Lin, Y. & Englander, S. W. (2002). *Proc. Natl. Acad. Sci. USA*, **99**, 12173-12178.
36. Krishna, M. M. G., Lin, Y., Rumbley, J. N. & Englander, S. W. (2003). *J. Mol. Biol.* **331**, 29-36.
37. Ascenzi, P., Coletta, M., Santucci, R., Polizio, F. & Desideri, A. (1994). *J. Inorg. Biochem.* **53**, 273-280.
38. Kumar, R., Prabhu, N. P. & Bhuyan, A. K. (2005). *Biochemistry*, **44**, 9359-9367.
39. Vanderkooi, J. M. & Erecinska, M. (1975). *Eur. J. Biochem.* **60**, 199-207.
40. Tsong, T. Y. (1974). *J. Biol. Chem.* **249**, 1988-1990.
41. Kon, H. (1969). *Biochem. Biophys. Res. Commun.* **35**, 423-427.
42. Bushnell, G. W., Louie, G. V. & Brayer, G. D. (1990). *J. Mol. Biol.* **214**, 585-595.
43. Sauder, J. M., MacKenzie, N. E. & Roder, H. (1996). *Biochemistry*, **35**, 16852-16862.
44. Kawahara, K. & Tanford, C. (1966). *J. Biol. Chem.* **241**, 3228-3232.
45. Jacob, M. & Schmid, F. X. (1999). *Biochemistry*, **38**, 13773-12779.
46. Santoro, M. M. & Bolen, D. W. (1988). *Biochemistry*, **27**, 8063-8068.
47. Ikeguchi, M., Kuwajima, K., Mitani, M. & Sugai, S. (1986). *Biochemistry*, **25**, 6965-6972.
48. Kumar, R. & Bhuyan, A. K. (2005). *Biochemistry*, **44**, 3024-3033.
49. Colón, W., Elöve, G. A., Wakem, L. P., Sherman, F. & Roder, H. (1996). *Biochemistry*, **35**, 5538-5549.
50. Shastry, M. C. R., Sauder, J. M. & Roder, H. (1998). *Acc. Chem. Res.* **31**, 717-725.
51. Sosnick, T. R., Shtilerman, M. D., Mayne, L. & Englander, S. W. (1997). *Proc. Natl. Acad. Sci. USA,* **94**, 8545-8550.

52. Qi, P. Q., Sosnick, T. R. & Englander, S. W. (1998). *Nature Struct. Biol.* **5**, 882-884.
53. Bhuyan, A. K., Udgaonkar, J. B. (1999). *Perspectives in Structural Biology* (Vijayan, M., Yathindra, N. & Kolaskar, A. S., eds), pp. 293-303, University Press (India) Limited, Hyderabad.
54. Akiyama, S., Takahashi, S., Ishimori, K. & Morishima, I. (2000). *Nature Struct. Biol.* **7**, 514-520.
55. Chen, E., Wittung-Stafshede, P. & Kliger, D. S. (1999). *J. Am. Chem. Soc.* **121**, 3811-3817.
56. Portman, J. J., Takada, S. & Wolynes, P. G. (2001). *J. Chem. Phys.* **114**, 5082-5096.
57. Kaya, H. & Chan, H. S. (2002). *J. Mol. Biol*. **315**, 899-909.
58. Thirumalai, D. (1999). *J. Phys. Chem. B*, **103**, 608-610.
59. Portman, J. J., Takada, S. & Wolynes, P. G. (2001). *J. Chem. Phys.* **114**, 5069-5081.
60. Yang, W. Y. & Gruebele, M. (2003). *Nature*, **423**, 193-197.
61. Krieger, F., Fierz, B., Bieri, O., Drewello, M. & Kiefhaber, T. (2003). *J. Mol. Biol.* **332**, 265-274.
62. Xu, Y., Mayne, L. C. & Englander, S. W. (1998). *Nature Struct. Biol.* **5**, 774-778.
63. Tanford, C. (1968). *Adv. Protein. Chem.* **23**, 121-282.
64. Leopold, P. E., Montal, M. & Onuchic, J. N. (1992). *Proc. Natl. Acad. Sci. USA*, **89**, 8721-8725.
65. Wolynes, P. G., Onuchic, J. N. & Thirumalai, D. (1995). *Science*, **267**, 1619-1620.
66. Schonbrun, J. & Dill, K. A. (2003). *Proc. Natl. Acad. Sci. USA*, **100**, 12678-12682.
67. Socci, N. D. & Onuchic, J. N. (1994). *J. Chem. Phys.* **101**, 1519-1528.

68. Socci, N. D., Onuchic, J. N. & Wolynes, P.G. (1998). *Proteins: Struct. Funct. Genet.* **32**, 136-158.
69. Dill, K. A. (1999). *Protein Sci.* **8**, 1166-1180.
70. Shakhnovich, E. I. & Gutin, A. M. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 7195-7199.
71. Shakhnovich, E. I. (1997) *Curr. Opin. Struct. Biol.* **7**, 29-40.
72. Jacob, J., Krantz, B., Dothager, R. S., Thiagarajan, P. & Sosnick, T. R. (2004). *J. Mol. Biol.* **338**, 369-382.
73. Magg, C. & Schmid, F. X. (2004). *J. Mol. Biol.* **335**, 1309-1323.
74. Baker, D. (2000). *Nature*, **405**, 39-42.
75. Plaxco, K. W., Simons, K. T., Ruczinski, I. & Baker, D. (2000). *Biochemistry*, **39**, 11177-11183.
76. Martínez, J.C. & Serrano, L. (1999). *Nature Struct. Biol*. **6**, 1010-1016.
77. Chiti, F., Taddei, N., White, P. M., Bucciantini, M., Magherini, F., Stefani, M. & Dobson, C. M. (1999). *Nature Struct. Biol.* **6**, 1005-1009.
78. Hagen, S. J., Hofrichter, J., Szabo, A. & Eaton, W. A. (1996). *Proc. Natl. Acad. Sci. USA*, **93**, 11615-11617.
79. Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2000). *Proc. Natl. Acad.Sci. USA*, **97**, 7220-7225.
80. Zhou, H.-X. (2003). *J. Chem. Phys.* **118**, 2010-2015.
81. Bhuyan, A. K. & Udgaonkar, J. B. (1998). *Proteins: Struct. Funct. Genet.* **32**, 241-247.
82. Elöve, G. A., Bhuyan, A. K. & Roder, H. (1994). *Biochemistry*, **33**, 6925-6935.
83. Fersht, A. R. (1995). *Proc. Natl. Acad. Sci. USA,* **92**, 10869-10873.
84. Wendoloski, J. J., Matthew, J. B., Weber, P. C. & Salemme, F. R. (1987). *Science*, **238**, 719-866.
85. Makarov, D. E., & Plaxco, K. W. (2003). *Protein Sci.* **12**, 17-26.

86. Tanford, C. (1970). *Adv. Protein Chem.* **24**, 1-95.
87. Udgaonkar, J. B. & Baldwin, R. L. (1988). *Nature*, **335**, 694-699.
88. Hamada, D., Segawa, S. I. & Goto, Y. (1996). *Nature Struct. Biol.* **3**, 868-873.
89. Troullier, A., Reinstädler, D., Dupont, Y., Naumann, D. & Forge, V. (2000). *Nature Struct. Biol.* **7**, 78-86.
90. Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2002). *Nature Struct. Biol.* **9**, 209-216.
91. Dobson, C. M. (2004). *Seminars Cell & Developmental Biol.* **15**, 3-16.
92. Hartl, F. U. & Hayer-Hartl, M. (2002). *Science*, **295**, 1852-1858.
93. Gething, M-J. & Sambrook, J. (1992). *Nature*, **355**, 33-45.
94. Schmid, F. X. (1986). *Methods Enzymology*, **131**, 70-82.
95. Wildegger, G. & Kiefhaber, T. (1997). *J. Mol. Biol.* **270**, 294-304.
96. von Hippel, P. H. & Wong, K-Y. (1965). *J. Biol. Chem.* **240**, 3909-3923.
97. Schindler, T., Herrler, M., Marahiel, M. A. & Schmid, F. X. (1995). *Nature Struct. Biol.* **2**, 663-673.
98. Silow, M., Tan, Y-J., Fersht, A. R. & Oliveberg, M. (1999). *Biochemistry*, **38**, 13006-13012.
99. Yonetani, T. Yamamoto, H., Erman, J. E., Leigh, J. S. & Reed, G. H. (1972). *J. Biol. Chem.* **247**, 2447-2455.
100. Roder, H., Elove, G. A. & Englander, S. W. (1988). *Nature*, **335**, 700-704.
101. Creighton, T. E. (1997). *Trends Biochem. Sci*. **22**, 6-10.
102. Mayor, U., Guydosh, N. R., Johnson, C. M., Grossman, J.G., Sato, S., Jas, G. S., Freund, S. M. V., Alonso, D. A. V., Daggett, V. & Fersht, A. R. (2003). *Nature*, **421**, 863-867.
103. Fersht, A. R. (1993). *FEBS Lett.* **325**, 5-16.

104. Laurents, D. V., Bruix, M., Jamin, M. & Baldwin, R. L. (1998). *J. Mol. Biol.* **283**, 669-678.
105. Bai, Y. (1999). *Proc. Natl. Acad. Sci. USA*, **96**, 477-480.
106. Weisman, J. S. & Kim, P. S. (1991). *Science*, **253**, 1386-1393.
107. Kiefhaber, T. & Schmid, F. X. (1992). *J. Mol. Biol.* **224**, 231-240.
108. Muthukrishnan, K. & Nall, B. T. (1991). *Biochemistry*, **30**, 4706-4710.
109. Zwanzig, R. W. (1997). *Proc. Natl. Acad. Sci. USA*, **94**, 148-150.
110. Kumar, R., Prabhu, N. P., Yadaiah, M. & Bhuyan, A. K. (2004). *Biophys. J.* **87**, 2656-2662.
111. Ptitsyn, O. B. (1995). *Adv. Protein Chem.* **47**, 83-229.
112. Kuwajima, K. & Arai, M. (2000). *Mechanism of Protein Folding* (Pain, R. H. ed.), pp.138-174, 2nd edition. Oxford University Press, New York.
113. Arai, M. & Kuwajima, K. (2000). *Adv. Protein Chem.* **53**, 209-271.
114. Colón, W. & Roder, H. (1996). *Nature Struct. Biol.* **3**, 1019-1025.
115. Mok, K. H., Nagashima, T., Day, I. J., Hore, P.J. & Dobson, C. M. (2005) *Proc. Natl. Acad. Sci. USA,* **102**, 8899-8904.
116. Bychkova, V. E., Pain, R.H. & Ptitsyn, O.B. (1988). *FEBS Letters*, **238**, 231-234.
117. Fink, A. L. (2005). *Curr. Opin. Struct. Biol.* **15**, 35-41.
118. Tompa, P. (2003). *Bioassays*, **15**, 847-855.
119. Dyson, H. J. & Wright, P. E. (2002). *Curr. Opin. Struct. Biol.* **12**, 54-60.
120. Tompa, P. (2002). *Trends Biochem. Sci.* **27**, 527-533.
121. Uversky, V. N. (2002). *Eur. J. Biochem.* **269**, 2-12.
122. Nakayama, K. I., Hatakeyama, S. & Nakayama, K. (2001). *Biochem. Biophys. Res. Commun.* **282**, 853-860.

123. Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M. & Uversky, V. N. (2005). *FEBS J.* **272**, 5129-5148.
124. Kokai, E., Tantos, A., Vissi, E., Szoor, B., Tompa, P., Gausz, J. *et al*. (2006). *Arch. Biochem. Biophys.* **451**, 59-67.
125. Radivojac, P., Vucetic, S., O' connor, T. R., Uversky, V. N., Obradovic, Z. & Dunker, A. K. (2006). *Proteins: Struct. Funct. Genet.* **63**, 398-410.
126. Tompa, P., Banki, P., *et al*., (2006). *Biophys. J.* **91**, 2243-2249.
127. Baker, B. Y., Yaworsky, D. C. & Miller, W. L. (2005). *J. Biol. Chem.* **280**, 41753-41760.
128. Kuroda, Y., Kidokoro, S., & Wada, A. (1992). *J. Mol. Biol.* **223**, 1139-1153.
129. Nishii, I., Kataoks, M., Tokunaga. F., & Goto, Y. (1994). *Biochemistry*, **33**, 4903-4909.
130. Goto, Y., Calciano, L. J. & Fink, A. L. (1990). *Proc. Natl. Acad. Sci. USA*, **87**, 573-577.
131. Goto, Y., Takahashi, N. & Fink, A. L. (1990). *Biochemistry,* **29**, 3480-3488.
132. Goto, Y. & Fink, A. L. (1989). *Biochemistry*, **28**, 945-952.
133. Rami, B.R. & Udgaonkar, J. B. (2002). *Biochemistry*, **41**, 1710-1716.
134. Ohgushi, M. & Wada, A. (1983). *FEBS Letters*, **164**, 21-24.
135. Jeng, M-F., Englander, S.W., Elöve, G. A., Wand, A. J. & Roder, H. (1990). *Biochemistry*, **29**, 10433-10437.
136. Jeng, M-F. & Englander, S. W. (1991). *J. Mol. Biol.* **221**, 1045-1061.
137. Kataoka, M., Hagihara, Y., Mihara, K. & Goto, Y. (1993). *J. Mol. Biol.* **229**, 591-596.
138. Hagihara, Y., Tan, Y. & Goto, Y. (1994). *J. Mol. Biol*. **237**, 336-348.

139. Chalikian, T. V., Gindikin, V. S. & Breslauer, K. J. (1995). *J. Mol. Biol*. **250**, 291-306.
140. Rao, D. K., Kumar, R., Yadaiah. M. & Bhuyan, A. K. (2006). *Biochemistry*, **45**, 3412-3420.
141. Kim, P. S. & Baldwin, R. L. (1982). *Annu. Rev. Biochem*. **51**, 459-489.
142. Dolgikh, D. A., Gilmanshin, R. I., Brazhnikov, E. V. & Ptitsyn, O. B. (1981). *FEBS Letters*, **136**, 311-315.
143. Kuwajima, K. (1977). *J. Mol. Biol*. **114**, 241-258.
144. Kuwajima, K. (1989). *Proteins: Struct. Funct. Genet*. **6**, 87-103.
145. Rao, D. K., Prabhu, N. P. & Bhuyan, A. K. (2006). *Biochemistry*, **45**, 8393-8401.
146. Ferrer, J. C., Guillemette, J. G., Bogumil, R., Inglis., S.C., Smith, M. & mauk, A.G. (1993). *J. Am. Chem. Soc*. **115**, 7507-7508.
147. Pollock, W. B. R., Rosell, F. I., Twitchett, M. B., Dumont, M. E. & Mauk, A. G. (1998). *Biochemistry*, **37**, 6124-6131.
148. Russell, B. S., Melenkivitz, R. & Bren, K. L. (2000). *Proc. Natl. Acad. Sci. USA*, **97**, 8312-8317.
149. Assfalg, M., Bertini, I., Dolfi, A., Turano, P., Mauk, A.G., Rosell, F. I. & Gray, H. B. (2003). *J. Am. Chem. Soc*. **125**, 2913-2922.
150. Hagihara, Y., Aimoto, S., Fink, A.L. & Goto, Y. (1993). *J. Mol. Biol*. **231**, 180-184.
151. Schellman, J. A. (1978). *Biopolymers*, **17**, 1305-1322.
152. Schellman, J. A. (1987). *Annu. Rev. Biophys. Biophys. Chem*. **16**, 115-137.
153. Makhatadze, G. I. & Privalov, P. L. (1992). *J. Mol. Biol.* **226**, 491-505.
154. Privalov, P. L. (1990). *Crit. Rev. Biochem. Mol. Biol*. **25**, 281-305.
155. Becktel, W. J. & Schellman, J. A. (1987). *Biopolymers*, **26**, 1859-1877.
156. Potekhin, S. & Pfeil, N. (1989). *Biophys. Chem*. **34**, 55-62.

157. van Osdol, W. W., Mayorga, O. L. & Freire, E. (1991). *Biophys. J.* **59**, 48-54.
158. Cohen, D. S. & Pielak, G. J. (1994). *Protein Sci*. **3**, 1253-1260.
159. Hoang, L., Maity, H., Krishna, M. M. G., Lin, Y. & Englander, S. W. (2003). *J. Mol. Biol*. **331**, 37-43.
160. Maity, H., Rumbley, J. N. & Englander, S. W. (2006), *Proteins: Struct. Funct. Genet*. **63**, 349-355.
161. Wilson, M. T. & Greenwood, C. (1996). In *Cytochrome c: A Multidisciplinary Approach* (Scott, R. A. & Mauk, A. G., eds), pp. 611-634, University Science Books, Sausalito, CA.
162. Chalikian, T. V., Gindikin, V. S. & Breslauer, K. J. (1996). *FASEB J.* **10**, 164-170.
163. Goto, Y. & Nishikori, S. (1991). *J. Mol. Biol.* **222**, 679-686.
164. Hibbard, L. S. & Tulinsky, A. (1978). *Biochemistry*, **17**, 5460-5468.
165. Dunbar, J., Yennawar, H. P., Banerjee, S., Luo, J. & Farber, G. K. (1997). *Protein Sci*. **6**, 1272-1733.
166. Bhuyan, A. K. (2002). *Biochemistry*, **41**, 13386-13394.
167. Pace, C. N., Laurents, D. V. & Thomson, J. A. (1990). *Biochemistry*, **29**, 2564-2572.
168. Mayr, L. M. & Schmid, F. X. (1993). *Biochemistry*, **32**, 7994-7998.
169. Ptitsyn, O. B.(1992). In *Protein Folding* (Creighton, T. E. ed), p. 234-300, W. H. Freeman and Company, New York.
170. Kauzmann, W. (1959). *Adv. Protein Chem.* **14**, 1-63.
171. Privalov, P. L. (1979). *Adv. Protein Chem.* **33**, 167-241.
172. Chen, B. & Schellman, J. A. (1989). *Biochemistry*, **28**, 685-691.
173. Privalov, P. L., Griko, Y. V., Venyaminov, S. Y. & Kutyshenko, V. P. (1986). *J. Mol. Biol.* **190**, 487-498.

174. Nelson, C. J., LaConte, M. J. & Bowler, B. E. (2001). *J. Am. Chem. Soc.* **123**, 7453-7454.
175. Mizuguchi, M., Hashimoto, D., Sakurai, M. & Nitta, K. (2000). *Proteins: Struct. Funct. Genet.* **38**, 407-413.
176. Yamada, Y., Yajima, T. *et al.* (2005). *J. Mol. Biol.* **350**, 338-348.
177. Murphy, K. P. & Freire, E. (1992). *Adv. Protein Chem.* **43**, 313-361.
178. Babu, C. R., Hilser, V. J. & Wand, A. J. (2004). *Nature Struct. Mol. Biol.* **11**, 352-357.
179. Tsai, C. J., Maizel, J. V. & Nussinov, R. (2002). *Crit. Rev. Biochem. Mol. Biol.* **37**, 55-69.
180. Prabhu, N. V. & Sharp, K. A. (2005). *Annu. Rev. Phys. Chem.* **56**, 521-548.
181. Makhatadze, G. I. & Privalov, P. L. (1995). *Adv. Protein Chem.* **47**, 307-425.
182. Spolar, R. S., Livingstone, J. R. & Record, M. T. (1992). *Biochemistry*, **31**, 3947-3955.
183. Spolar, R. S. & Record, M. T. (1994). *Science*, **263**, 777-784.
184. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995). *Protein Sci.* **4**, 2138-2148.
185. Aune, K. C. & Tanford, C. (1969). *Biochemistry*, **8**, 4586-4590.
186. Sorenson, J. M., Hura, G., *et al*., (1999). *J. Phys. Chem. B*, **103**, 5413-5426.
187. Brandts, J. F. (1964). *J. Am. Chem. Soc*. **86**, 4291-4301.
188. Franks, F., Hatley, R. H. M. & Friedman, H. L. (1988). *Biophys. Chem*. **31**, 307-315.
189. Hawely, S. A. (1971). *Biochemistry*, **10**, 2436-2442.
190. Zipp, A., & Kauzmann, W. (1973). *Biochemistry*, **12**, 4217-4228.
191. Franks, F., & Hately, R. H.M. (1985). *Cryo. Lett.* **6**, 171-180

192. Nojima, H., Ikai, A., Oshima, T., & Noda, H. (1977). *J. Mol. Biol.* **116**, 429-442.
193. DeKoster, G. T. & Robertson, A. D. (1995). *J. Mol. Biol.* **249**, 529-534.
194. Pometun, M. S., Peterson, R. W., Babu, C. R. & Wand, A. J. (2006). *J. Am. Chem. Soc*. **128**, 10652-10653.
195. Kumar, R., Prabhu, N. P., Rao, D. K. & Bhuyan, A. K. (2006). *J. Mol. Biol.* **364**, 483-495.
196. Ganesh, C., Eswar, N., Sarika, S., Ramakrishnan, C. & Varadarajan, R. (1999). *FEBS Lett*. **454**, 31-36.
197. Tsai, C. J., Maizel, J. V. & Nussinov, R. (2002). *Crit. Rev. Biochem. Mol. Biol*. **37**, 55-69.
198. Nicholoson E. M. & Scholtz, J. M. (1996). *Biochemistry*, **35**, 11369-11378.
199. Schindler, T. & Schmid, F. X. (1996). *Biochemistry*, **35**, 16833-16842.
200. Grio, Y. V., Privalov, P. L., Sturtevant, J. M. & Venyaminov, Y. S. (1988). *Proc. Natl. Acad. Sci. USA*, **85**, 3343-3347.
201. Bágeľova, J. Antalik, M. & Tomori, Z. (1997). *Biochem. Mol. Biol. Int.* **43**, 891-900.
202. Graziano, G., Catanzano, F., Riccio, A. & Barone, G. (1997). *J. Biochem.* **122**, 395-401.
203. Griko, Y. V., Privalov, P. L., Sturtevant, J. M. & Venyaminov, S. Y. (1988). *Proc. Natl. Acad. Sci. USA.* **85**, 3343-3347.
204. Makhatadze, G. & Privalov, P. L. (1996). *Protein Sci.* **5**, 507-510.
205. Hallerbach, B. & Hinz, H.-J. (1999). *Biophys. Chem.* **76**, 219-227.
206. Makhatadze, G. I. & Privalov, P. L. (1993). *J. Mol. Biol.* **232**, 639-659.
207. Privalov, P. L. & Makhatadze, G. I. (1993). *J. Mol. Biol.* **232**, 660-679.
208. Makhatadze, G. I. & Privalov, P. L. (1990). *J. Mol. Biol.* **213**, 375-384.

209. Baldwin, R. L. & Rose, G. D. (1999). *Trends. Biochem. Sci.* **24**, 77-83.
210. Jackson, S. E. & Fersht, A. R. (1991). *Biochemistry* **30**, 10428-10435.
211. Jackson, S. E. (1998). *Folding Des.* **3**, 81-91.
212. Prabhu, N. P., Kumar, R. & Bhuyan, A. K. (2004). *J. Mol. Biol.* **337**, 195-208.
213. Bryngelson, J.D. & Wolynes, P. G. (1987). *Proc. Natl. Acad. Sci. USA,* **84**, 7524-7528.
214. Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z. & Socci, N. D. (1995). *Proc. Natl. Acad. Sci. USA,*. **92**, 3626-3630.
215. Thirumalai, D., Ashwin, V., Bhattacharjee, J. K. (1996). *Phys. Rev. Lett.* **77**, 5385-5388.
216. Nymeyer, H., Garcia, A. E. & Onuchic, J. N. (1998). *Proc. Natl. Acad. Sci. USA,* **95**, 5921-5928.
217. Matthews, C. R. (1993). *Annu. Rev. Biochem.* **62**, 653-683.
218. Onuchic, J. N. & Wolynes, P. G. (2004). *Curr. Opin. Struct. Biol.* **14**, 70-75.
219. Onuchic, J. N., Nymeyer, H., Garcia, A. E., Chahine, J. & Socci, N. D. (2000). *Adv. Protein. Chem.* **53**, 87-152.
220. Gillespie, B. & Plaxco, K. W. (2004). *Annu. Rev. Biochem.* **73**, 837-859.
221. Bhuyan, A. K., Rao, D. K. & Prabhu, N. P. (2005). *Biochemistry*, **44**, 3034-3040.
222. Kamagata, K., Arai, M. & Kuwajima, K. (2004). *J. Mol. Biol.* **339**, 951-965.
223. Plaxco, K. W., Simons, K. T. & Baker, D. (1998). *J. Mol. Biol.* **277**, 985-994.
224. Debe, D. A. & Goddard, W. A. (1999). *J. Mol. Biol.* **294**, 619-625.

225. Munoz, V. & Eaton, W. A. (1999). *Proc. Natl. Acad. Sci. USA,* **96**, 11311-11316.
226. Dinner, A. & Karplus, M. (2001). *Nat. Struct. Biol.* **8**, 21-22.
227. Ghaemmaghami, S., Word, J. M., Burton. R. E., Richardson, J. S. & Oas, T. G. (1998). *Biochemistry*, **37**, 9179-9185.
228. Gromiha, M. M. & Selvaraj, S. (2001). *J. Mol. Biol.* **310**, 27-32.
229. Zhou, H. & Zhou, Y. (2002). *Biophys. J.* **82**, 458-463.
230. Paci, E., Lindorff-Larsen, K., Dobson, C. M., Karplus, M. & Vendruscolo, M. (2005). *J. Mol. Biol.* **352**, 495-500.
231. Gong, H., Isom, D. G., Srinivasan, R. & Rose, G. D. (2003). *J. Mol. Biol.* **327**, 1149-1154.
232. Wolynes, P. G. (1997). *Proc. Natl. Acad. Sci. USA,* **94**, 6170-6175.
233. Gutin, A. M., Abkevich, V. I., & Shakhnovich, E. I. (1996). *Phys. Rev. Lett.* **77**, 5433-5436.
234. Ivankov, D. N., Garbuzynskiy, S. O., Alm, E., Plaxco, K. W., Baker, D. & Finkelstein, A. V. (2003). *Protein Sci.* **12**, 2057-2062.
235. Nanganathan, A. N. & Munoz, V. (2005). *J. Am. Chem. Soc.* **127**, 480-481.
236. Li, M. S., Klimov, D. K. & Thirumalai, D. (2004). *Polymer* **45**, 573-579.
237. Scott, K. A., Batey, S., Hooton, K. A. & Clark, J. (2004). *J. Mol. Biol.* **344**, 195-205.
238. Koga, N. & Takada, S. (2001). *J. Mol. Biol.* **313**, 171-180.
239. Galzitskaya, O. V., Garbuzynskiy, S. O., Ivankov, D. N., & Finkelstein, A. V. (2003). *Protein. Sci.* **51**, 162-166.
240. Maxwell, K. A., Wildes, D., Zarrine-Afsar, A., de Los Rios, M. A., *et al*. (2005). *Protein. Sci.* **14**, 602-616.
241. Sali, A., Shakhnovich, E. & Karplus, M. (1994). *Nature*, **369**, 248-251.

242. Sali, A., Shakhnovich, E. & Karplus, M. (1994). *J. Mol. Biol.* **235**, 1614-1636.
243. Clarke, J., Cota, E., Fowler, S. B. & Hamill, S. J. (1999). *Struct. Folding Des.* **7**, 1145-1153.
244. Gutin, A. M., Abkevich, V. I. & Shakhnovich, E. I. (1995). *Proc. Natl. Acad. Sci. USA*, **92**, 1282-1286.
245. Ivankov, D. N. & Finkelstein, A. V. (2001). *Biochemistry*, **40**, 9957-9961.
246. Fersht, A. R. (1997). *Curr. Opin. Struct. Biol.* **7**, 3-9.
247. Wolynes, P. G. (2004). *Proc. Natl. Acad. Sci. USA,* **101**, 6837-6838.
248. Alonso, D. O., Daggett, V. (2000). *Proc. Natl. Acad. Sci. USA,* **97**, 133-138.
249. Huang, S. G. & Oas, T. G. (1995). *Proc. Natl. Acad. Sci. USA,* **92**, 6878-6882.
250. Huang, S. G. & Oas, T. G. (1995). *Biochemistry*, **34**, 3884-3892.
251. Kragelund, B. B., *et al*. (1995). *Biochemistry*, **34**, 7217-7224.
252. Kragelund, B. B., *et al*. (1996). *J. Mol. Biol.* **256**, 187-200.
253. Mines, G. A., Pascher, T., Lee, S. C., Winkler, J. R. & Gray, H. B. (1996). *Chemistry & Biology*, **3**, 491-497.
254. Spector, S., *et al*. (1998). *J. Mol. Biol.* **276**, 479-489.
255. Ferguson, N., *et al*. (1999). *J. Mol. Biol.* **286**, 1597-1608.
256. Wittung-Stafshede, P., Lee, J. C., Winkler, J. R. & Gray, H. B. (1999). *Proc. Natl. Acad. Sci. USA,* **96**, 6587-6590.
257. Mayor, U., Jhonson, C. M., Daggett, V. & Fersht, A. R. (2000). *Proc. Natl. Acad. Sci. USA,* **97**, 13518-13522.
258. Wang, M., *et al*. (2003). *J. Am. Chem. Soc.* **125**, 6032-6033.
259. Schönbrunner, N., Koller, K. P. & Kiefhaber, T. (1997). *J. Mol. Biol.* **268**, 526-538.

260. Schindler, T., *et al*. (1999). *J. Biol. Chem.* **274**, 3407-3413.
261. Perl, D., Holtermann, G. & Schmid, F. X. (2001). *Biochemistry,* **40**, 15501-15511.
262. Wassenberg, D., Welker, C. & Jaenicke, R. (1999). *J. Mol. Biol.* **289**, 187-193.
263. Reid, K. L. Rodriguez, H. M., Hillier, B. J. & Gregoret, L. M. (1998). *Protein Sci.* **7**, 470-479.
264. Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L. & Serrano, L. (1994). *Biochemistry*, **33**, 2142-2150.
265. Grantcharova, V. P. & Baker, D. (1997). *Biochemistry*, **36**, 15685-15692.
266. Plaxco, K. W. *et al*. (1998). *Biochemistry*, **37**, 2529-2537.
267. Plaxco, K. W., Spitzfaden, C., Campbell, I. D. & Dobson, C. M. (1997). *J. Mol. Biol.* **270**, 763-770.
268. Fong, S., *et al*. (1996). *J. Mol. Biol.* **264**, 624-639.
269. Hamill, S. J. Meekhof, A. E. & Clarke, J. (1998). *Biochemistry*, **37**, 8071-8079.
270. Clarke, J., Hamill, S. J. & Jhonson, C. M. (1996). *J. Mol. Biol.* **270**, 771-778.
271. Grantcharova, V. P., Alm, E. J., Baker, D. & Horwich, A. L. (2001). *Curr. Opin. Struct. Biol.* **11**, 70-82.
272. Carrion-Vazquez, M., *et al*. (1999). *Proc. Natl. Acad. Sci. USA,* **96**, 3694-3699.
273. Guijarro, J. I., Morton, C. J., Plaxco, K. W., Campbell, I. D. & Dobson, C. M. (1998). *J. Mol. Biol.* **276**, 657-667.
274. Villegas, *et al*. (1995). *Biochemistry*, **34**, 15105-15110.
275. Robinson, C. R. & Sauer, R. T. (1996). *Biochemistry,* **35**, 13878-13884.

276. Khorasanizadeh, S., Peters, I. D., Butt, T. R. & Roder, H. (1993). *Biochemistry*, **32**, 7054-7063.

277. Scalley, M. L., Yi, Q., Gu, H., McCormack, A., Yates, J. R. III & Baker, D. (1997). *Biochemistry*, **36**, 3373-3382.

278. van Nuland, N. A. J., *et al*. (1998). *Biochemistry*, **37**, 622-637.

279. Main, E, R. G., Fulton, K. F. & Jackson, S. E. (1999). *J. Mol. Biol.* **291**, 429-444.

280. van Nuland, N. A. J., *et al*. (1998). *J. Mol. Biol.* **283**, 883-891.

281. Choe, S. E., Matsudaira, P. T., Osterhout, J., Wagner, G. & Shakhnovich, E. I. (1998). *Biochemistry*, **37**, 14508-14518.

282. Muñoz., Lopez, E. M., Jager, M. & Serrano, L. (1994). *Biochemistry*, **33**, 5858-5866.

283. Kuhlman, B., Luisi, D., Evans, P. A., Raleigh, D. P. (1998). *J. Mol. Biol.* **284**, 1661-1670.

284. Guerois, R. & Serrano, L. (2000). *J. Mol. Biol.* **304**, 967-982.

285. Aronsson, G., Brorsson, A., Sahlman, L. & Jonsson, B. (1997). *FEBS Lett.* **411**, 359-364.

List of Publications:

1. **Prabhu, N. P.**, Kumar, R. & Bhuyan, A. K. (2004). *J. Mol. Biol.* **337**, 195-208.
2. Kumar, R., **Prabhu, N. P.**, Yadaiah, M. & Bhuyan, A. K. (2004). *Biophys. J.* **87**, 2656-2662.
3. Bhuyan, A. K., Rao, D. K. & **Prabhu, N. P.** (2005). *Biochemistry* **44**, 3034-3040.
4. Kumar, R., **Prabhu, N. P.** & Bhuyan, A. K. (2005). *Biochemistry* **44**, 9359-9367.
5. **Prabhu, N. P.** & Bhuyan, A. K. (2006). *Biochemistry* **45**, 3805-3812.
6. Rao, D. K., **Prabhu, N. P.** & Bhuyan, A. K. (2006). *Biochemistry* **45**, 8393-8401.
7. Kumar, R., **Prabhu, N. P.**, Rao, D. K. & Bhuyan, A. K. (2006). *J. Mol. Biol.* **364**, 483-495.