EXPLORING METAL-LIGAND BONDS IN METALLOPROTEINS BY MECHANICAL FORCE

by

Peng Zheng

B.Sc., Nanjing University, China, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2013

Abstract

The nature and properties of chemical bonding are at the heart of chemical reactivity, where bond rupture (activation) processes are the first step in chemical reactions. Thus, knowledge concerning the bond strength and characteristics is fundamentally important to understanding general chemical processes.

Metalloproteins are ubiquitous in biology. Through the formation of multiple metal-ligand bonds, metal centers are incorporated into metalloproteins, greatly expanding their stability and functionality. Consequently, determining the characteristics of a metal center and individual an metal-ligand bond are critical towards understanding their unique stability and function.

This thesis presents a series of studies on the mechanical rupture processes of metal-ligand bonds in the metal center of metalloproteins using single molecule Atomic Force Microscopy (AFM). Specifically, the investigation of the ferric-thiolate bonds in a simple iron-sulfur protein rubredoxin is presented.

First, we develop a protein-level chemical coupling method to build poly-metalloproteins suitable for AFM studies. The ability to build a pure iron-form poly-rubredoxin molecule paves the way to investigating metalloproteins using AFM.

This thesis then shows that the FeS_4 center in rubredoxin can be ruptured by mechanical force, the first example demonstrating that a naturally occurring metal

center in a biological system can be ruptured using AFM. Moreover, the measured rupture force of the highly covalent ferric-thiolate bond is surprisingly low.

The rupture mechanism of multiple ferric-thiolate bonds is then explored. Using a loop insertion rubredoxin variant, we found that the FeS_4 center ruptures in a stochastic manner, where bonds can break both sequentially and cooperatively. In addition, by stretching the metal center from different directions, the mechanical anisotropy of the FeS_4 metal center is studied.

The interplay between metal-ligand bond strength and the protein structure is also presented here, where we found that the amide hydrogen bond to the sulfur atom can modify the stability of the FeS_4 center.

Next, a ligand substitution reaction on the ferric-thiolate bonds in rubredoxin is investigated and discussed.

Finally, conclusions and future directions are presented

Preface

A version of chapter 2 has been published. [Zheng P.], Cao Y., Li H. (**2011**) Facile method of constructing polyproteins for single molecule force spectroscopy studies. **Langmuir** 27(10): 5713-5718. Dr. Cao Y. and Prof. Hongbin initiated the project and designed the method to build poly-GB1 protein. I and Prof. Li designed the method to build pure iron-form poly-rubredoxin protein. I engineered the rubredoxin polyprotein and performed the AFM experiments and data analysis. I worked together with Dr. Cao Y. on the experiment of protein GB1 part. I wrote the manuscript together with Prof. Li.

A version of chapter 3 has been published. [Zheng P.], Li H. (**2011**) Highly covalent ferric-thiolate bonds exhibit surprisingly low mechanical stability. **Journal of the American Chemical Society**, 133(17): 6791-6798. I designed this project with my supervisor Prof. Hongbin Li. I engineered all the proteins, performed the AFM experiments and analyzed all the experimental data. I wrote the manuscript together with Prof. Li.

A version of chapter 4 has been submitted for publication. [Zheng P.], Takayama SJ. Mauk AG. Li H. Single molecule force spectroscopy reveals that iron is released from rubredoxin by a stochastic mechanism. I designed this project together with my supervisor Prof. Hongbin Li. I engineered all the proteins, performed the AFM

experiments and analyzed all the experimental data. Takayama SJ. and I performed the CV, CD and NMR experiments. I wrote the manuscript together with Prof. Li. and Prof. Mauk.

A version of chapter 5 is in preparation for publication. [Zheng P.], Li H. Mechanical anisotropy of a simple FeS_4 metal center in rubredoxin. I designed this project together with my supervisor Prof. Hongbin Li. I engineered all the proteins. I performed the AFM experiments with the help of Chou J. and Zhou K. I analyzed all the experimental data. I wrote Chapter 5.

A version of chapter 6 has been published. [Zheng P.], Takayama SJ. Mauk AG. Li H. (**2012**) Hydrogen bond strength modulates the mechanical strength of ferric-thiolate bonds in rubredoxin, **Journal of the American Chemistry Society**, 134(9), 4124-4131. I designed the experiment together with Prof. Hongbin Li. I engineered all the proteins, performed the AFM experiments and analyzed the data. Takayama SJ. and I performed the CV and CD experiments. I wrote the manuscript with Prof. Li. and Prof. Mauk.

A version of chapter 7 is in preparation for publication. [Zheng P.], Li H. Direct measurement of thiocyanate substitution reaction on the FeS_4 center of rubredoxin. I designed this project together with my supervisor Prof. Hongbin Li. I engineered all the proteins and performed the AFM experiments with the help of Zhou, K. I analyzed the experimental data. I wrote this chapter.

Table of contents

List of figures

List of symbols and abbreviations

aa	amino acid
AFM	atomic force microscopy
$BM(PEO)_3$	1, 8-bis (maleimido)triethylene glycol
bp	base pair
CD	circular dichroism
cp	*clostridium pasteurianum*
CV	cyclic voltammetry
DNA	deoxyribonucleic acid
DDT	dithioreitol
E. coli	*Escherichia coli*
E	Extension
F	Force
GB1	the B1 binding domain of protein G
GFP	green fluorescent protein
I27	the 27^{th} Ig domain of human titin
Ig	immunoglobulin
IPTG	isopropyl β-D-1-thiogalactopyranoside
k_b	Boltzmann constant

k_0	bond dissociation rate at zero force
LB	Luria-Bertani broth
M	molar
mM	millimolar
nm	nanometer
NMR	nuclear magnetic resonance
OD	optical density
p	persistence length
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PDB	protein data bank
pf	*pyrococcus furiosus*
pN	piconewton
RD	rubredoxin
RNA	ribonucleic acid
RP<	revolutions per minute
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
STM	scanning tunneling microscopy
T	temperature
Tris	tris (hydroxymethyl) aminomethane

UHV	ultrahigh vacuum
v	pulling velocity/speed
WLC	worm like chain
WT	wild type
x	extension
ΔLc	contour length increment
ΔG	free energy
Δx	distance between bound state to transition state
3D	three-dimensional

Acknowledgements

I would like to express my sincere gratitude to all those who helped me complete this dissertation. First and foremost, my utmost thanks goes to my supervisor Dr. Hongbin Li. It is a great opportunity for me to pursue a Ph.D degree under his guidance. His support and encouragement makes this dissertation possible. His attitude and enthusiasm towards science inspire me deeply. He is a great role model for me. During the past five years in the lab, he gave me enough freedom to explore my ideas while keeping me on the right track. He taught me how to design experiments, how to write, and how to think as a scientist. Working on such an exciting research project with him will be cherished for all of my life. Hongbin is not only a good mentor but also a great personal friend. I deeply appreciate his support, care, patience and advice throughout my Ph.D study.

I would like to thank all the members of Li's group, both present and past. Your friendships made this experience more enjoyable. I am indebted to Dr. Yi Cao who taught me single molecule AFM. You gave me so much help when I first arrived in Vancouver. Thank you to Dr. Shulin Zhuang for training me in molecular biology techniques, from then on I can build my own recombinant protein. Thank you to Dr. Qing Peng for all your help when I first started my project. With all the help from you three, I am able to start my research smoothly at the very beginning. I would also like

to thank Ying Guo for her technique help on the molecule biology which saved me a lot time. Thus I can explore more on this exciting project. Thank you to Shanshan Lv, M.M. Balamuriali, Ashlee Jollymore, Eileen Wang, Tianjia Bu, Tao Shen, Chengzhi He, Devin Li, Na Kong, Jie Fang, Guillaume Lamour, Kailin Zhou, Yanyan Wang, Jinliang Li, Kai shih Er, Martha Ma and Chou Jeff for your friendship during the past five years.

I also appreciate the contributions of my collaborators, including Dr. Grant A. Mauk, Dr. Suzana K. Straus, Dr. Shin-ichi J. Takayama. Thank you to Dr. Pierre Kennepohl, Dr. Stephen G. Withers and Dr. Michael Blades for their inspiring discussions.

Thank you to Dr. Suzana K. Straus, Dr. Don Douglas and Dr. Pierre Kennepohl for your valued time and intellectual guidance as my supervisory committee. In particular, I am deeply indebted to Dr. Suzana K. Straus for her critical evaluation on the draft. I also express my utmost thanks for her insightful discussions and technical help during my Ph.D study.

Thank you to Dr. David Baker at University of Washington and Dr. Marly K. Eidsness for providing me the plasmids encoding protein GB1 and Rubredoxin, respectively.

Finally, I would like to thank my parents who raised me up, provided me with the best they can and encouraged me with their best wishes. A special thank you to my wife Xuan Ding, whose love, support, patience and understanding helped me complete this dissertation.

Dedication

To my parents and wife

Chapter 1:Introduction

The first critical step of any chemical reaction is the rupture of chemical bonds. A successful reaction requires enough energy to overcome the activation energy barrier, a process mainly achieved by heat, the subject of thermochemistry. In fact, mechanical force which is ubiquitous in nature can also facilitate chemical bond rupture and subsequent chemical reactions. This so-called mechanochemistry field is much less explored[1,2]. As early as in the 19th century, the mechanochemical reaction of silver halides such as AgCl, AgBr and AgI was investigated as they could dissociate to silver under an external compressing mechanical force[3]. Additionally, the degradation of polystyrene by milling is known as a forced bond rupture process inside the polymer molecule[4,5]. However, the direct observation of a single bond rupture event has not achieved until very recently. The measurement of the strength of a bond at the single molecule level has not been possible until the recent development of single molecule force spectroscopy techniques.

There are several different types of force spectroscopy techniques, such as magnetic tweezers, optical tweezers, atomic force microscopy, hydrodynamic methods, glass micro-needles and biomembrane force probe (BFP)[6-9,10,11]. They are able to measure force down to piconetwon (pN) and the first three techniques are widely used now. Magnetic tweezes and optical tweezes are largely used for studying

stability of DNA and RNA. However, forces that can be measured by these two techniques are below ~ 200 pN, which is less than the strength of many covalent bonds[12,13]. In comparison, the force range for AFM is between ~10 pN to tens of nanonewtons, such that it may be used to measure chemical bond strength. Consequently, we use AFM to investigate the bond strength of metal-ligand bonds in this thesis[1,14].

Atomic force microscopy (AFM) can stretch a single molecule from two holding points, thus directly applying a tiny force to the system. In doing so, the applied force sets the reaction coordinate in a single molecule along which the reaction will take place. In addition, the single molecule nature of the AFM technique greatly advances our understanding of chemical reactions by revealing many rare events and pathways which can be obscured by averaging, which is inherent in ensemble studies. Moreover, it provides a simple and reliable approach to directly measure the strength of individual chemical bonds.

This chapter provides an overview on the application of AFM to measure the strength of single bonds, focusing on two major fields of AFM studies: chemical bond rupture and protein unfolding. Both of these processes occur during the rupture of metal-ligand bonds inside metalloproteins, the overarching theme of this thesis.

I will first give an introduction to AFM, which is the most important analytical technique used in the work presented here. Selected AFM experiments measuring chemical bond strength and protein unfolding will be reviewed. Then, a very brief description about metalloproteins with a focus on the rubredoxin model system is

presented. Finally, the aim of this thesis will be stated.

1.1 Single molecule AFM

1.1.1 The principle of AFM

The atomic force microscope (AFM), invented in 1986 by Binnig, Quate and Gerber based on an earlier Nobel prize winning invention of scanning tunneling microscope (STM), is designed with the aim of investigating sample surfaces by measuring tiny forces at the atomic scale[15,16]. Compared with STM, which relies on a tunneling effect between a conducting substance and the tip, AFM is a very simple instrument. It uses a small spring-like cantilever with a sharp tip as a force sensor to measure interactions at piconewton level (10^{-12} N).

It is widely used as an imaging technique with several different modes. For example, under a constant force mode, the tip scans the surface along the *x* and *y* axes while keeping the force between them constant by adjusting the height of the tip in the *z* axis. As a result, a three dimensional image of the surface is revealed[15,17].

In this thesis, the AFM is used as a single molecule force spectroscopy tool to study mechanical stability of chemical bonds. It can directly measure the interaction force between atoms, and be used to manipulate a single molecule attached between the tip and substrate surface by mechanical force. Accurate spatial movement is achieved by moving the substrate with a piezoelectric positioner capable of moving with resolution at the Angstrom scale.

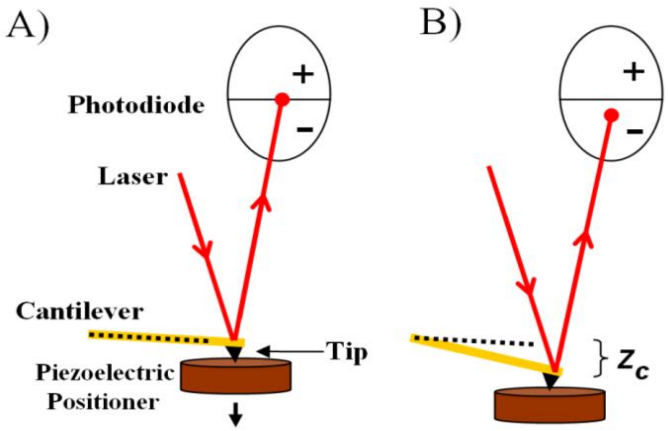

Figure 1.1 A schematic of the AFM technique A) There are three major components in the AFM: a cantilever with a sharp tip functions as a force sensor; a piezoelectric positioner moves in the z direction to stretch the molecule attached between the tip and the substrate, and a photodiode monitors the movement of a laser that shines on the back of the cantilever to indicate its deflection (z_c). B) Based on Hook's law, the resorting force can be calculated by the displacement of the cantilever monitored by the laser point on the photodiode. The dashed line indicates the original position of the cantilever.

When there are interactions between the tip and the surface, the cantilever will be deflected as a spring in the *z* direction, as shown in Figure 1.1B. At the same time, laser light bounces off the back of the cantilever; the movement of the light is recorded by a two-segment or a four-quadrant photodiode detector as an electrical signal. As a result, the deflection of cantilever (z_c) can be monitored by the change in deflection of the laser light. If the spring constant (k_c) of the cantilever is known, the force (F) exerted on the stretched molecule can be calculated by using Hook's law:

$$F = k_c z_c \tag{1.1}$$

In an AFM experiment, a commercial AFM cantilever composed of Si_3N_4 is employed, with an approximate value of the spring constant being specified by the manufactures. To obtain a more accurate spring constant, a thermal fluctuation method based on the equipartition theorem is frequently used[18,19]:

Modeled as a harmonic oscillator, the spring-like cantilever fluctuates in the *z* direction as it is driven by thermal energy. Based on the equipartition theorem, the thermal energy is equal to 1/2 k_bT. Thus, the potential energy of the cantilever equals to the thermal energy as follows:

$$\frac{1}{2}k_c < z_c{}^2 >= \frac{1}{2}k_B T \tag{1.2}$$

Here, $<z_c^2>$ is the time-average square of the displacement of the cantilever in the *z* direction. k_c is the calculated spring constant and T is the temperature. In order to separate the characteristic thermal excitation from the overall measured fluctuation which includes the environmental noise, the fluctuation of the cantilever is not analyzed in the time domain. Rather, it is converted to the frequency domain using a fast Fourier transformation.

1.1.2 Constant velocity mode of AFM to study single molecule

1.1.2.1 Working principle of constant velocity mode of AFM

There are several operational modes of AFM when it is used as a force spectroscopy. The constant velocity mode (also called constant speed mode) is widely used. In this case, the molecule is pulled by AFM under a constant velocity[13,20-28]. There are other modes such as force ramp mode and force clamp mode in which the applied force is increased linearly over time or kept constant[29].

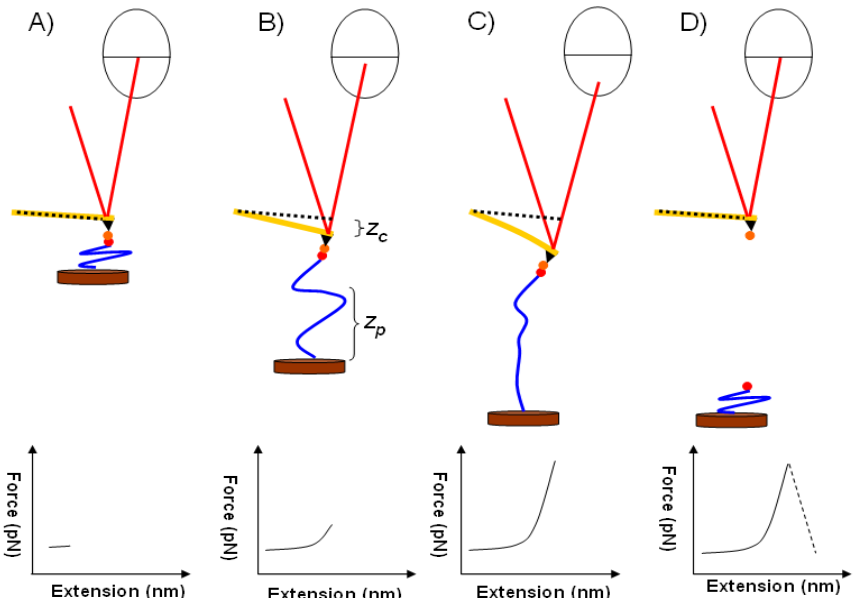

Figure 1.2 A schematic showing an AFM experiment measuring chemical bond strength by stretching a single molecule. A) The experimental sequence: a single molecule is picked up by the tip. The dashed line indicates the initial position of the cantilever. B) Stretching the molecule extends the polymer as determined by its elasticity. C) At a certain moment, the applied force triggers the rupture of the weakest chemical bond in the linkage, where the force value of the peak is recorded as the rupture force of the bond. D) After the bond breaks, the force drops immediately to nearly zero.

For example, AFM can be used to stretch a single molecule under constant velocity mode. After the substrate is loaded with sample solution, the piezoelectric positioner which holds the substrate moves towards the tip at a constant speed while the tip remains still. The two parts then come into contact with each other with a compressing force of ~1 nN and the piezoelectric positioner moves away at the same speed to its original position. This cycle is repeated thousands of times during the experiment. Approximately ~1% of the time, a single molecule will be held between the tip and the substrate (Fig. 1.2A).

The attached molecule will be then stretched and extended as the piezoelectric positioner moves away from the tip (Fig. 1.2B). At the same time, the restoring force

from the molecule to the cantilever is recorded by the deflection of the cantilever. The extension of the molecule (E) can be calculated by the difference between the movement of the positioner and the displacement of the cantilever (Z_p-Z_c) (Fig 1.2B). As a result, the relationship between the force and extension of the molecule is plotted as the so-called force-extension curve (Fig. 1.2 bottom curve). When a single polymer molecule is stretched, the force-extension curve can be described by models of polymer elasticity, such as the Freely-Joint Chain (FJC) model or the Worm-Like Chain (WLC) model[30,31].

As the polymer is stretched under an increasing force, the weakest bond in the molecular linkage will break. As shown in Figure 1.2C, a chemical bond in a polymer chain is ruptured, and the force value of the peak is recorded as the mechanical rupture force for that particular chemical bond. Consequently, the linkage between the tip and surface is broken and the force drops nearly to zero immediately (Fig. 1.2D).

1.1.2.2 Mechanical bond strength is a kinetic stability

By definition, the mechanical bond strength is the most probable force at which a chemical bond ruptures in repeated tests. Traditionally, the average rupture force with standard deviation is reported in AFM studies. As shown in Figure 1.3, the mechanical stability of a chemical bond is determined by the activation free energy barrier ($\Delta G_{N\text{-}T}$), which is the free energy difference between the bound state and the mechanical transition state. On comparison, the classic thermodynamic stability is determined by the free energy barrier ($\Delta G_{N\text{-}U}$) difference between the bound state and

the completely unbounded state. Consequently, the measured rupture force as the mechanical bond strength is a kinetic stability instead of thermodynamic stability. In addition, the mechanical bond dissociation pathway is different from the pathway of thermal assisted bond dissociation. Thus, there is no direct correlation between the mechanical bond strength and the classic chemical bond strength (energy).

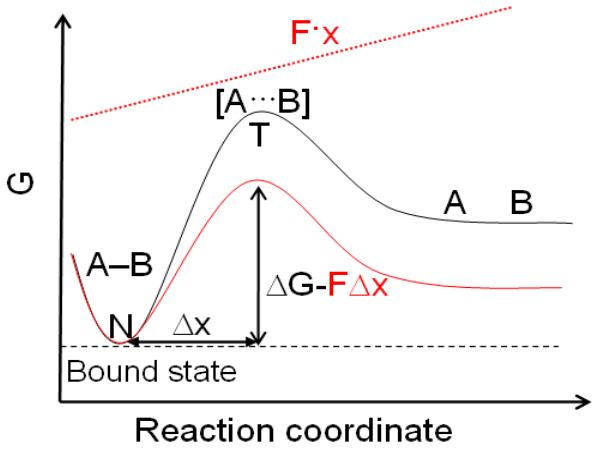

Figure 1.3 The energy landscape of a chemical bond A-B is tilted by an external force. The black line describes the energy landscape of a chemical bond A-B with three states: the native bound state (N), the transition bond dissociation state (T), and the completely dissociated bond state (U). Under an external force (*F*), the energy landscape of the bond will be tilted by the mechanical work (*F**x) as the red dash line. Consequently, the energy landscape is the new red line and the activation energy will be decreased by *F**Δx*.

1.1.2.3 Chemical bond dissociation under force

Chemical bond dissociation process under external force is far from equilibrium, as the rate of bond association under this condition can be neglected. Consequently, the bond rupture processes in AFM studies are accomplished in non-equilibrium, stochastic fashion[14,32-34][14,32-34]. To explain the bond dissociation process under mechanical force, we firstly describe the spontaneous bond dissociation process based

on Kramer's theory. For a chemical bond with single activation barrier, the spontaneous bond dissociation rate:

$$k_d = \kappa \frac{k_B T}{h} \exp(-\frac{\Delta G}{k_B T}) \tag{1.3}$$

In equation 1.3, ΔG is the bond dissociation activation energy barrier ($\Delta G_{N\text{-}T,}$ used as ΔG for simplicity in the following context), κ is the transmission coefficient, k_B is Boltzmann constant, T is the absolute temperature and h is Planck's constant.

When an external force is applied to a chemical bond, the bond dissociation process will be accelerated exponentially as is described by the Bell-Evans model[2,32]. The energy landscape of the chemical bond will be tilted by applied force (F) along the reaction coordinate (x) (Fig 1.3). As a result, the bond activation energy barrier will be decreased by $F*\Delta x$:

$$k(F) = \kappa \frac{k_B T}{h} \exp(-\frac{\Delta G - F\Delta x}{k_B T}) = k_d \exp(\frac{F\Delta x}{k_B T}) \tag{1.4}$$

Here Δx is the distance (length) between the native bound state (chemical bond length) and the bond dissociation transition state in the reaction coordinate. From equation 1.3, it is clear that the rate of bond rupture under force is exponentially dependent on the applied force. The probability of bond rupture can be described as: $dP_d=(1-P_d)*k(t)dt$, where k(t) is the mechanical bond dissociation rate at time t which can be derived from equation 1.3.

For example, if the applied force is increased linearly with a loading rate of a ($F=a*t$, t is the time), as in the case of a force ramp mode of AFM, the probability distribution (density) of bond rupture as a function of the applied force is:

$$P_d(F) = 1 - \exp^{\frac{-k_d k_B T}{a\Delta x}(\exp^{\frac{F\Delta x}{k_B T}} - 1)} \quad (1.5)$$

The most probable rupture force (F_d) at a given loading rate (a) can be calculated based on equation 1.5:

$$F_d = \frac{k_B T}{\Delta x} \ln\left(\frac{a \cdot \Delta x}{k_d \cdot k_B T}\right) \quad (1.6)$$

It shows that the most probable rupture force (F_d) is linearly dependent upon logarithm of the loading rate (a). If we plot the rupture force (F_d) versus ln(a), the spontaneous bond dissociation rate (k_d) and the distance between the bound and transition state (Δx) can be extracted from the slope and the intercept.

In practice, most experiments are performed under a constant speed mode and the rupture (unbind) force is measured as a function of pulling speed. For the simple AFM measurement of single protein-ligand or protein-protein interactions, the force applied on the complex can be assumed increasing with a constant loading rate. Thus, the loading rate (a) which is the time derivative of the applied force (dF/dt) can be calculated as the product of the pulling speed and the spring constant of the AFM cantilever ($v*k=(dl/dt)*k=dF/dt=a$)[35]. Consequently, the kinetic and energy landscape can be derived from the relationship between unbinding force and the loading rate[2,33].

To obtain more accurate value, the effective loading rate for each individual unbinding event can be directly determined by the slope of the force versus time curve just before the unbinding events[36]. It is noted that this method should be applied for studies where the protein and ligand is linked between the AFM tip and surface through a long polymer. Because the loading rate is not increased constantly at this

condition, an effective rate loading should be examined at the time of bond rupture.

This theory/prediction is verified by mechanical unbinding studies of the avidin-biotin complex and streptavidin-biotin complex at early 1990s[37-39]. These two receptor-ligand systems as non-covalent weak bonds have received tremendous attention in the studies of single molecule force spectroscopy which are among the first few experiments in the field[8,37,40,41]. By directly break the interaction and measure the unbinding force as a function of pulling speed using AFM or BFP, the linear relationship between unbinding force and loading rate is found and agreed well with the prediction. Consequently, the kinetics and energy landscape is obtained using equation 1.6.

1.1.2.4 Force-extension curve of protein unfolding

AFM is also widely used to study how a protein folds and unfolds under mechanical strain. Compared with the result of stretching a polymer, the force extension curve of protein unfolding shows a critical piece of information: the contour length increment (ΔLc)[12,42]. Similar to the chemical bond rupture process that occurs as a polymer is stretched, increasing the applied force triggers the unfolding of protein and a resultant force peak (Fig. 1.4 step 1-3). At step one, a protein molecule is picked up by the tip of cantilever. Then, the increasing force applied to the protein triggers the unfolding of the protein. A force peak from protein unfolding arises. From it, the value of the mechanical stability of the protein can be derived. Thus, the protein unfolding force can be measured directly. At step three, the peptide chain of unfolded

protein is further stretched and extended to a much longer length leading to a contour length increment (ΔLc). Finally, the protein molecule detaches from the tip or the substrate leading to a second force peak.

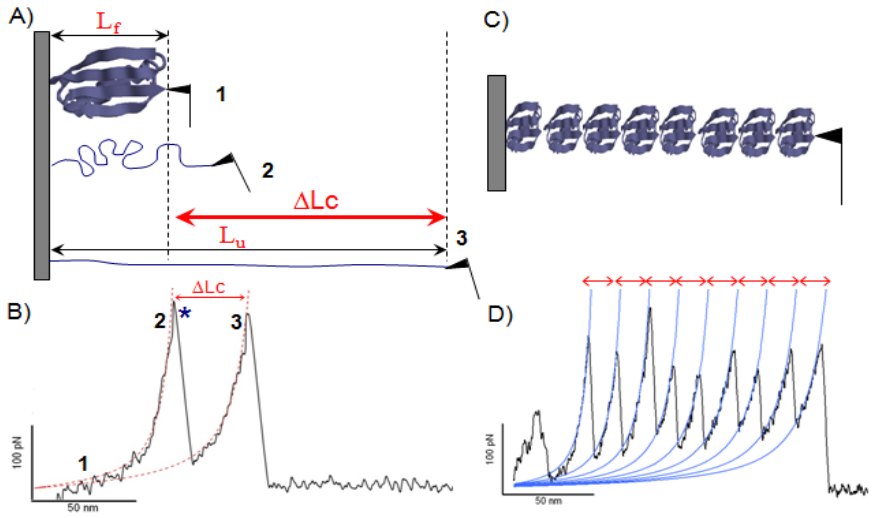

Figure 1.4 The process by which a protein is unfolded by AFM. A) The cartoon shows how single protein molecule is unfolded. The number indicates the three step of protein unfolding. The red part is the contour length increment when the protein domain unfolds. Protein GB1 (1PGA) is used for display. B) The corresponding force-extension curve shows a protein unfolding force peak (highlighted by a blue star). The red dash line is the WLC fitting to the force extension curves. C) A schematic shows a polyprotein molecule containing eight identical repeats of single protein domain stretched by AFM. D) The corresponding force-extension curve shows a characteristic saw-tooth like pattern containing eight protein unfolding peaks with equal contour length increment.

Besides the direct measurement of the force at which a protein unfolds, the contour length increment (ΔL_c) upon protein unfolding provides unique structural information about the protein (Fig. 1.4). For a folded protein stretched by AFM, the contour length is roughly the distance between the N and C terminus (L_f). When the protein is unfolded, the previously folded peptide chain will be extended, leading to a much longer contour length (L_u). By fitting the multiple unfolding events with the WLC model which describe the relationship between protein extension and the

resorting force on the protein:

$$F(x) = \frac{k_B T}{p}\left[\left(\frac{1}{4} \cdot (1 - \frac{x}{L_c})^{-2} - \frac{1}{4} + \frac{x}{L_c} \right] \right. \quad (1.7)$$

Where *F(x)* is the force at extension *x*, k_B is Boltzmann constant, *T* is the absolute temperature, *p* is the persistence length and L_c is the contour length of the polymer.

The contour length increment can be directly measured from experiment. It is of note that the value of ΔL_c is proportional to the number of amino acids (*N*) released upon protein unfolding, where the extended length of one amino acid (L_{aa}) is 0.36 nm. Thus, the number of released amino acids can be calculated as:

$$N = \Delta Lc / L_{aa} \quad (1.8)$$

If the structure of the protein is known, the number of released amino acids can be calculated and used to validate the experimental result.

In the studies of mechanical stability of proteins, polyprotein are typically used. The polyprotein molecule contains tandem repeats of an identical single protein domain. As a result, each individual protein domain will sequentially unfold, resulting in multiple unfolding force peaks with the same contour length increment (Fig. 1.4C-D). This characteristic saw-tooth like pattern of the multiple force peaks is used as an excellent single molecule fingerprint for protein unfolding studies.

However, the direct extraction of the kinetic data from polyprotein unfolding studies is challenging using the previous method for protein-ligand interaction. Because the force *(F)* applied to the protein is not increased in a linear fashion here, it follows the non-linear relationship according to the polymer elasticity of protein

which can be described by the WLC model as the equation 1.7.

As a result, the loading rate (a) during a constant velocity pulling experiment on protein unfolding is not a fixed value but varied as followed:

$$a = \frac{dF}{dt} = \frac{dF}{dx} \cdot \frac{dx}{dt} = \frac{k_B T}{pL_c}\left[(2 - \frac{2vt}{L_c})^{-3} + 1 \right] \quad (1.9)$$

Consequently, the loading rate (a) is positively dependent on the pulling velocity (v). Because of this non-linear relationship and the multiple domain nature of protein, an analytical solution for the most probable rupture force (F_d) cannot be directly derived under a constant pulling velocity mode as the equation 1.6. As a result, the extraction of kinetic information is difficult. Instead, a Monte Carlo simulation or numerical fitting is needed to extract the two kinetic parameters (k_d and Δx) from different pulling velocity experiments[43].

1.2 AFM studies on chemical bond rupture and protein unfolding

For the past two decades, single molecule AFM in force spectroscopy modehas evolved into a powerful technique applied in many interdisciplinary research approaches. Because it can apply tiny forces to manipulate single molecules, it is widely applied in the subjects of chemistry, physics and biology. For example, AFM has been used to explore molecular recognition, polymer elasticity, protein unfolding/folding, mechanoenzymatics and DNA repair[40,44-47]. Here, two AFM approaches are introduced: the mechanical rupture of individual chemical bonds and the mechanical unfolding of proteins.

1.2.1 Studies of mechanical rupture of chemical bonds

One of the most important features of AFM is that it can directly apply force on a selected target and in doing so break the molecular linkage and measure the corresponding force. Thus, AFM has been used used to measure chemical bond strength soon after it was invented[48]. One approach is to measure the interaction force between atoms on the tip and the surface. The other method is to link the target chemical bond within a polymer molecule and measure the rupture force. These methodologies will be presented below.

1.2.1.1 Measurement of interaction between AFM tip and surface

To directly measure the chemical bond strength between an AFM tip and a surface, the tip must be precisely placed at a defined position on the surface to ensure a single atom to atom interaction. However, it is difficult to achieve this at room temperature for several reasons. For one, the thermal noise affects the tip position dramatically at room temperature. In addition, the measured force needs to exclude long-range interactions (Van den Waals force and electrostatic interaction) occurring between the other atoms from the tip and the surface. To counteract these effects, experiments were performed under ultrahigh vacuum (UHV) and sometimes at very low temperature (below 10K). This approach has led to the measurement of several covalent and metallic bond strength, including Si-Si, W-Si, Au-Au and Ni-Ni. These bond rupture forces were found to be around 2-4 nN[49-52].

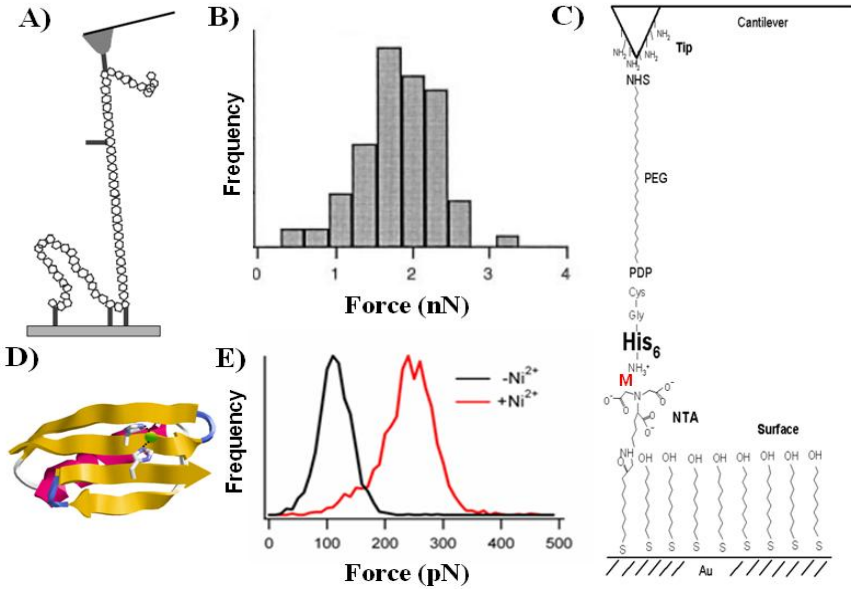

Figure 1.5 Measurement of chemical bond strength using AFM A) The schematic shows how chemical bond strength is measured by stretching a polymer. Several Si-C bonds are formed between the polymer and tip or glass surface. The short black line represents the Si-C bond. (adapted from the reference (14)). B) The histogram of measured rupture force of Si-C bond. (adapted from the reference (14)). C) The schematic shows the linkage in the AFM experiment on the measurement of NTA/Histag interaction. (adapted from the reference (63)). D) The structure of protein GB1 with a bi-histidine metal chelating motif. E) The unfolding force histogram of the GB1 mutant with and without adding Ni^{2+} (adapted from the reference (110)).

1.2.1.2 Measurement of bond strength using polymer coupling method

Another widely used approach is a polymer-facilitated chemical bond rupture experiment. In this type of experiment, a critical step is to link the polymer to a chemically modified tip and surface by a specific chemical coupling reaction. Many different chemical methods have been developed and have paved the way for this type of studies[14,53-57]. As a result, the target chemical bond is inside the polymer or at the linkage between the polymer and the tip or surface. When the polymer is picked up and stretched, it will be extended in a known fashion. At the same time, the force transmits through the polymer chain and is applied to the target chemical bond. Under

this condition, the extended polymer separates the tip and substrate surface far enough to avoid interactions between them. As a result, a tip with a low spring constant can be used to measure the force of a chemical bond rupture in solution at room temperature.

In 1999, Dr. Gaub and his colleagues reported the direct measurement of a covalent bond rupture force using AFM. The target chemical bond was linked in a polymer and was ruptured during the stretching of the polymer molecule[14]. They demonstrated a chemical bond rupture event by mechanical force at the single molecule level and reported its corresponding bond strength.

In their experimental setup, an amylose molecule was covalently anchored between an AFM tip and a glass substrate using carbodiimide chemistry. As a result, several Si-C bonds were formed between the polymer, the tip and the glass surface (Fig. 1.5A). When a single molecule was picked up, it was stretched leading to the rupture of the molecular linkage. The last force peak was determined to be the chemical bond rupture event and showed an average rupture force of 2.0 ± 0.3 nN (Fig. 1.5B). This force value was assigned to the weakest Si-C bond according to theoretical calculations of the bond strength.

This pioneering experiment clearly demonstrated that a single covalent bond can be ruptured by mechanical force and that bond strengths can be identified and measured using AFM with the help of theoretical calculations. It is now generally accepted that the force needed to rupture a covalent bond is above 1 nN[14,58]. This experiment paved the way for the so-called covalent mechochemistry studies. A similar combination of experiment and theoretical calculation have been carried out to

determine the Si-O, Au-S and Au-Au bond strengths[28,52,59,60].

1.2.1.3 Interactions between metal and ligand coordination bond system

Besides the covalent bond, there are several interesting AFM measurements on metal-ligand bond interactions. For example, the NTA/Histag (N-nitrilo-triacetic acid/Histidine tag) system with metal-histidine interactions has been studied intensively.

In 2000, three groups independently reported details on the NTA/His-tag interaction using AFM[61-63]. In this system, a metal ion (such as Ni^{2+}) with six coordination sites is used to bind a target protein. Four sites are occupied by the NTA through the nitrogen and oxygen atoms (Fig. 1.5C), and the other two sites used to bind two additional histidines in the Histag (six histidine residues in series) present within the sequence of the target protein. All three experiments used a long molecule to attach the histag or the NTA system and to serve as the linkage between tip and surface.

In one of these studies, Dr. Hinterdofer and co-workers used a long hetero-bifunctional polyethylene glycol (PEG) derivative containing an amine and a thiol reactive end as a bridge to couple the His_6 portion in a peptide to an amine-functionalized Si_3N_4 tip. A self assembled monolayer containing NTA was attached to the gold coated surface via gold-thiol chemistry (Fig. 1.5C)[54,63]. As a result, the NTA-His_6 system was successfully coupled between the AFM tip and substrate surface. After adding nickel ion into the solution, they found that the rupture force of

the interaction between Ni^{2+} and NTA was ~150 pN. In addition, the lifetime of this interaction was revealed to be on the order of milliseconds. As a result, they showed that the NTA-His_6 interaction system is of considerable mechanical stability and potentially as a general tool to anchor ligands to AFM tips. Indeed, they later successfully demonstrated such a application in molecular recognition experiments[64]. The NTA system containing other metal ion such as Co^{2+}, Cu^{2+} and Zn^{2+} were also studied[62].

From our perspective of singe chemical bond strength measurement, these studies provide valuable information about metal-ligand bond interaction strength in the NTA/Histag system. Because of the complexity of the system, it is difficult to identify which bond is ruptured in the experiment. The bond ruptured from interaction between histidine and the metal ion or the oxalate and the metal ion are both possible. In addition, it is not known whether the measured force is from a single bond or multiple bonds.

A range of other metal-ligand bonds systems have been studied by AFM, such as rubidium terpyridine bonds, the β-cyclodextrin to ferrocene bonds, which demonstrated rupture forces below 100 pN[65,66].

1.2.1.4 Measurement of electromechanical property of a single molecule

The break junction method, where molecular junctions are repeatedly formed and broken using STM, is widely used to measure single molecule conductance[67-69]. Recently, a conducting AFM has been developed to simultaneously measure

conductance and rupture force[67,69,70]. The principle of this instrument is based on a gold-coated AFM cantilever and substrate to capture a molecule, which allows an electric circuit to form. By applying a voltage as well as stretching the molecule, electromechanical properties of the single molecule can be measured. Frequently, a gold-ligand bond is tested in this type of experiments. For example, the rupture force of an Au-N bond between a gold atom and bipyridine molecule is found to be ~0.8 nN with a conductance of 1.0×10^{-4} G_0[67,71]. From the perspective of chemical bond strength measurements, this technique provides additional information concerning the electrical properties of the molecule and can potentially be used to study electron transfer processes in chemical and biological systems.

1.2.1.5 The interaction between metal and ligand bond in a protein

Another interesting metal-ligand interaction study has been reported by our group[23,72]. With the aim of increasing the mechanical strength of a particular protein, Dr. Li and co-workers engineered a bi-histidine metal-chelating site into the model protein GB1 (Fig.1.5D)[110]. Adding nickel ion to the protein solution dramatically increases the unfolding force of the protein, from ~100 pN to ~250 pN (Fig.1.5D). It is clear that this enhanced stability derives from the added interactions between nickel and histidine in the protein, giving insight into the degree of metal-ligand interaction in the protein. Furthermore, the unfolding of this protein shows a clear fingerprint for the metal-ligand bond rupture event. However, a pure measurement of the bond strength is not available as the measured rupture force arises from a combination of

simultaneous protein unfolding as well as metal ligand bond rupture.

1.2.2 Protein mechanical unfolding studies

In addition to studies on the mechanochemistry of chemical bond rupture processes, single molecule AFM is also extensively used to study the mechanical properties of proteins[20,74]. In particular, both the force and spatial resolution (in the pN and nm range) of the AFM make it very suitable for protein unfolding studies.

Mechanical force is ubiquitous in nature and many proteins are responsible for mechanical process in biological systems. For example, the giant muscle protein titin is one of the mostly studied proteins in this family[75]. Titin is essential for the passive elasticity of muscle. Under high stretching forces, it was found that titin can be extended to several times its original length without breaking.

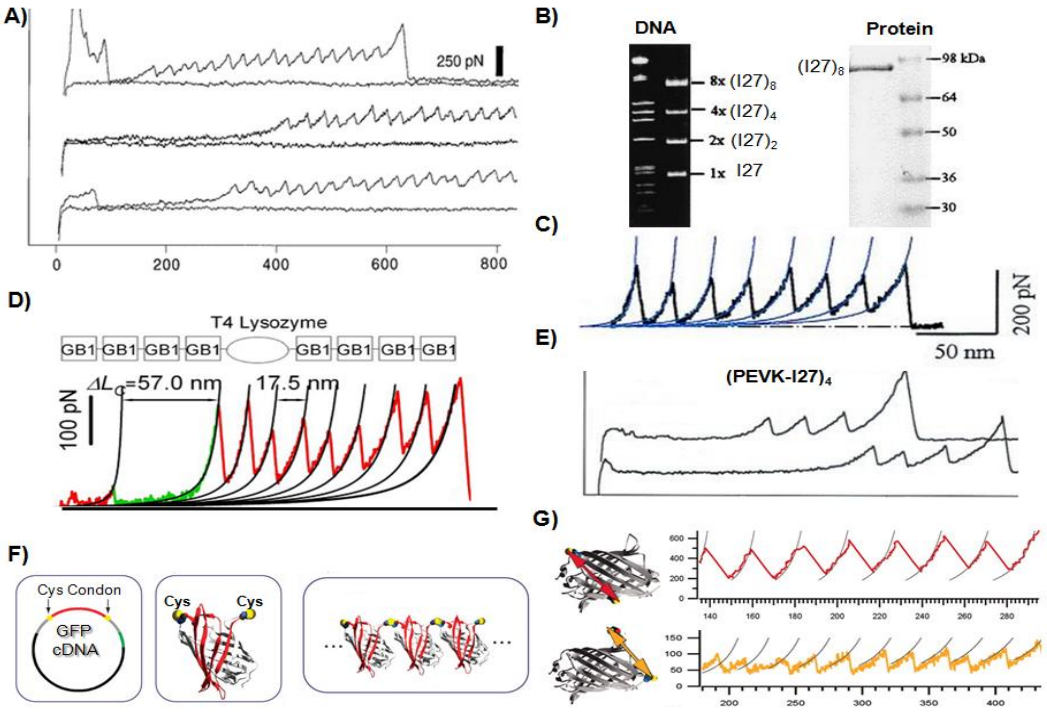

Figure 1.6 AFM experiments on protein unfolding. A) The mechanical unfolding of

titin resulted in a sawtooth-like force extension curves. (adapted from the reference (76)). B) The figure showed the successive construction of the gene coding for polyprotein $(I27)_8$ and band of protein $(I27)_8$. (adapted from the reference (42)). C) The unfolding of polyprotein $(I27)_8$ showed a sawtooth like force-extension pattern. (adapted from the reference (42)). D) The T4 lysozyme was sandwiched by four GB1 domains in an engineered hetero-polyprotein. The trace showed the unfolding event of T4 lysozyme (in green) and GB1 (in red). (adapted from the reference (26)). E) The hetero-polyprotein $(I27\text{- }PEVK)_3$ was constructed to study the mechanical stability of PEVK. The initial curves showed a low force corresponding to the extension of PEVK. Adapted from (92). F) The construction of poly-GFP. The cysteine residue was added at two positions in GFP, leading to the expression of a GFP monomer with two cysteines. A GFP polyprotein was constructed by forming an intermolecular disulfide bond. (adapted from the reference (98)). G) By stretching GFP from two different locations, the mechanical anisotropic deformation of GFP was found. (adapted from the reference (99)).

1.2.2.1 Mechanical unfolding of titin

It is not surprising that titin, with its natural mechanical functionality, was the first protein to be probed by AFM. In 1997, mechanical unfolding experiments on titin were first reported[7,76,77]. Using AFM, a sawtooth-like pattern showing multiple peaks in the force-extension curve with an unfolding force of ~200 pN was found (Fig.1.6A). The individual peaks were assigned to the unfolding of each Ig domain of titin, with a representative contour length increment of 27-29 nm. Many other proteins with mechanical properties have been studied by AFM, including fibronectin, tenascin, ankyrin and spectrin[14,30,78-83].

1.2.2.2 Using the polyprotein method for AFM studies

The characteristic sawtooth-like force-extension curves observed in titin unfolding experiment not only provide a wealth of information regarding the protein's stability, but also serve as a natural fingerprint for a single molecule stretching event,

originating from the unfolding of multiple domains within the protein. Dr. Fernandez and co-workers developed a polyprotein approach mimicking titin's structure for AFM studies. They built a so-called 'polyprotein' consisting of multiple repeats of an identical protein domain[42]. For example, they chose the I27 domain and built the octomeric gene sequence $(I27)_8$ in a stepwise fashion (Fig. 1.6B). The expressed polyprotein $(I27)_8$, containing eight identical I27 domains, can be used directly for AFM experiment. The unfolding of this engineered polyprotein gives the representative sawtooth-like force-extension curve similar to titin. Furthermore, each individual force peak in the curve can be unambiguously assigned to the unfolding event of an I27 domain (Fig. 1.6B) and a characteristic unfolding force can be also be assigned. This polyprotein methodology has been widely used as an efficient and accurate way of studying protein mechanical stability. Many other proteins have been studied in this fashion, such as ubiquitin, protein L,GB1 and spectrin[84-91] .

1.2.2.3 Hetero-polyprotein approach to study protein mechanics.

Based on the polyprotein strategy, a hetero-polyprotein approach is often used to study complex protein unfolding pathways. The target protein is sandwiched by a well-characterized protein domain that servers as a single molecule fingerprint. For example, Dr. Li and co-workers studied the parallel unfolding mechanism of T4 lysozyme using a construct of $(GB1)_4$-T4L-$(GB1)_4$[26]. The GB1 domain served as a fingerprint to help identify the unfolding behavior of lysozyme (Fig. 1.5D)[26].

This approach is also used to investigate proteins with low mechanical stability.

For example, Dr. Fernandez and co-workers proved that the PEVK sequence was mechanically labile in titin using a $(I27\text{-}PEVK)_3$ polyprotein construct[92]. During the unfolding experiment, they found a sawtooth like pattern which is from unfolding of the I27 domains only after a long initial spacer (Fig. 1.6E). Consequently, they concluded that the initial trace was from the extension of the PEVK, revealing mechanically labile properties of the PEVK. This approach has been widely used to study mechanically uncharacterized proteins, such as scaffold and maltose binding proteins[93-96].

1.2.2.4 Polyprotein engineering by disulfide bond at the protein level

The polyproteins constructed previously are mostly designed at the DNA level in which a plasmid encoding the entire polyprotein is built. Recently, a disulfide bond crosslink strategy was developed to construct polyproteins based on single protein molecule[97,98]. As shown in Figure 1.6F, Dr. Rief and co-workers engineered a single GFP (Green Fluorescent Protein) with a pair of cysteines, building a poly-GFP via an intermolecular disulfide bond[98,99]. Compared with the more traditional method, where the protein linkage is restricted to the N and C terminus, this method allows the protein to be stretched in selected directions based on the location of the engineered cysteine residues. It can be used to explore the mechanical anisotropy of proteins[85,99,100]. For example, this method was used to investigate the anisotropic nature of GFP (Fig. 1.6G). Similar studies were carried out on ankyrin, E2lip3, src SH3 protein domains and GB1[100-103].

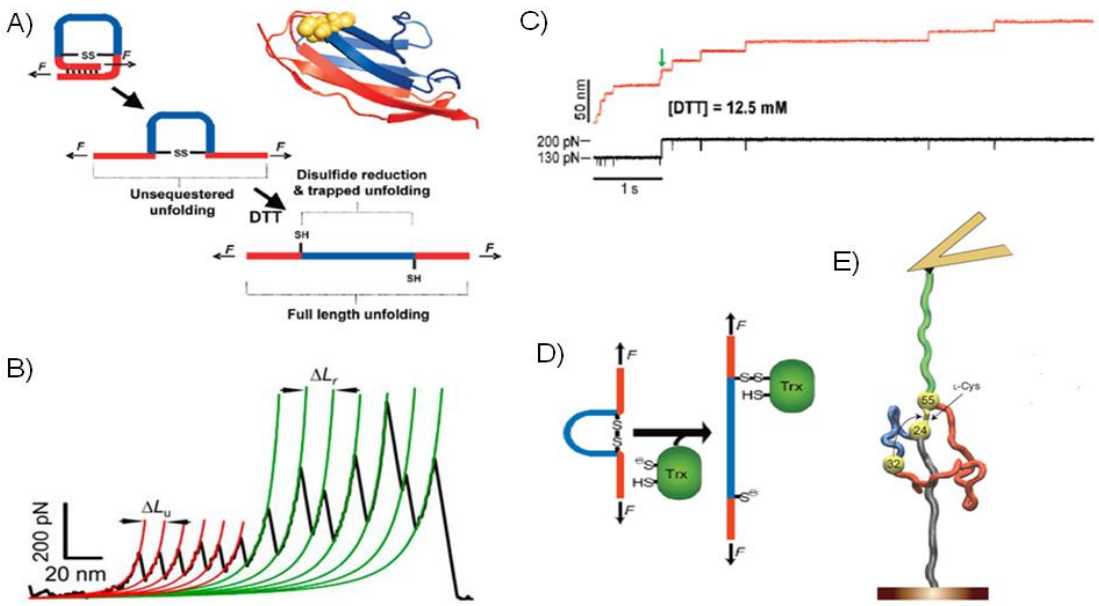

Figure 1.7 Mechanical reduction of a disulfide bond inside a protein. A) A schematic showing two-step unfolding of I27 with a disulfide bond. The disulfide bond is reduced under force leading to the extension of previous trapped residues by the bond. (adapted from the reference (99)). B) The corresponding force-extension curve of the two-step unfolding of the polyprotein using constant speed mode. The red initial portion is from the first step unfolding of I27, and the green part is due to the rupture of the disulfide bond. (adapted from the reference (102)). C) The mechanical unfolding trace of the protein using a force-clamp mode AFM. The stairs are from the unfolding of the protein as the force peaks at constant speed mode. (adapted from the reference (99)). D) The mechanical rupture of a disulfide bond by the enzyme thioredoxin. Adapted from (98). E) A schematic of studying the disulfide bond isomerization event probed by AFM using an I27 protein with multiple disulfide bonds. (adapted from the reference (58)).

1.2.2.5 Disulfide bond reduction in a protein under force

In addition to many protein unfolding experiments, a series of pioneering studies on the mechanical reduction of disulfide bonds inside a protein framework have been performed recently[58,104-107]. Here protein structure only serves as a sensitive probe for identifying the single bond rupture event. Disulfide bonds are covalent, with a mechanical stability above 1 nN. Under a reducing environment, it can be cleaved at relatively low force. Dr. Fernandez and co-workers engineered an artificial disulfide

bond inside the I27 protein domain, and studied the reduction of this disulfide bond in the presence of mechanical force[108]. The disulfide bond blocks the protein into two parts. As shown in Fig. 1.7A, two types of unfolding peak patterns are observed: one is from the partial unfolding of protein structure which is outside the disulfide bond, and the second is from the force-activated disulfide bond reduction (Fig. 1.7B). This I27 protein framework, which can contain one or multiple disulfide bonds was used to study interesting features of the disulfide bond reduction process, such as the mechanism of disulfide bond reduction in the presence of different reduction reagents and biological enzymes, and the disulfide isomerization process. Because the extension of previously trapped residues upon disulfide bond cleavage serves as a clear fingerprint, the results clearly illustrate many novel aspects of the disulfide bond.

1.2.3 Other processes studied under mechanical force

Although this thesis focuses on mechanical metal-ligand bond rupture processes in metalloproteins, it is worth mentioning many other interesting studies under mechanical force.

Investigations of protein-ligand and protein-protein interactions are pioneering studies in the field of single molecule force spectroscopy as early as 1990s and are still of great attention[8,44,54,109-118]. To understand the mechanical determinant of a mechanically stable protein is another popular research field and thus to rationally design proteins with enhanced mechanical stabilities is of great interest[21,23,119-125]. In

addition, studies on the environmental effects on protein stability have also been conducted for decades. Moreover, several theoretical models are proposed to explain the mechanical protein unfolding experiment. Many molecular dynamics simulations and quantum chemical calculations are performed which provide critical insight to mechanical process and significantly increases our understanding of experimental results.[37,60,126-136]. Recently, several novel studies focus on the mechanical stability of membrane protein and intrinsically disordered protein which increase the application of AFM in protein studies[137-142]. In addition, the mechanical role on other systems are also explored, such as polymers, DNA molecules, and living cells[143-153].

1.2.4 Why metalloproteins are chosen to study chemical bond strength?

Aspects of AFM studies discussed so far have been described in terms of either chemical bond rupture or protein unfolding. The chemical bond rupture experiment suffered from a lack of a clear fingerprint when single bond rupture events occur, while the protein unfolding experiment exhibits a clear signature from the characteristic contour length increment exhibited upon unfolding. Interestingly, the mechanical reduction of artificial disulfide bonds inside a protein demonstrated a promising example where the unfolding of the protein can be a fingerprint for the mechanical bond reduction process. However, the disulfide bond typically cannot be ruptured at low force without a reducing agent using AFM, with the focus so far being on engineered systems.

This thesis utilizes AFM to investigate rupture processes in chemical bonds in

naturally occurring biological systems, with a clear signature from a coupled protein unfolding event. This type of metalloproteins is ubiquitous in nature, with ~30% of proteins known thus far containing metal co-factors[154]. The unfolding of metalloproteins frequently includes a step in which the metal is lost so that the metal-ligand bonds dissociate. Consequently, the multiple metal-ligand bonds present in the metal center of metalloproteins are a great candidate for AFM studies. Despite this, there have been few studies to date on the metal-ligand bond rupture process occurring inside metalloproteins by AFM. Additionally, very little information is available regarding metal-ligand bond strength at the single bond level.

1.3 Metalloproteins

1.3.1 General introduction

Metalloproteins are a ubiquitous type of biomolecule in nature[154]. The incorporation of various types of metal ions into the protein structure greatly expands the function and stability of this class of proteins[155]. Considering the variety present within this class of proteins, metal-ligand bonds are the focus of this introduction.

Among more than tens of different metal ions in proteins, iron and copper are the two most common transition metal ions in metalloproteins[156]. These types of proteins function mainly as oxygen carriers, electron transfer proteins and enzymes in nature. Some ions such as zinc can provide structural stability and promote protein assembly and folding. Frequently, these metal ions are bonded by an endogenous ligand from the protein residues forming multiple metal-ligand bonds as an active or structural site.

The nitrogen atom from histidine, the sulfur atom from cysteine and the oxygen atom from aspartate are the three most common residues in metalloproteins to coordinate metals[157].

One important feature of metal-ligand bonds in metalloproteins is the affinity between the metal ion and the ligand (the dissociation constant)[158]. These measurements are mainly performed using the equilibrium competition method such as metal or ligand competition. This approach provides important information on metal selection and speciation in the protein. In addition, unfolding experiments on metalloproteins also involve the rupture of metal-ligand bonds. The dissociation of the metal-ligand bonds which results in change of a spectroscopic signal can be easily identified as a single step or coupled with protein unfolding using thermal or chemical denature methods.

1.3.2 Rubredoxin

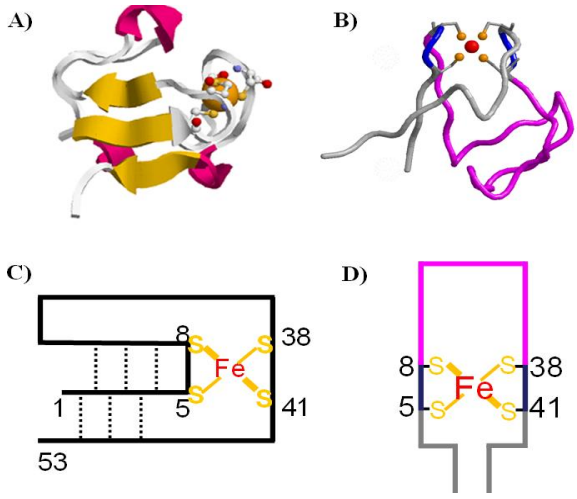

Figure 1.8 Structure of rubredoxin with a FeS_4 center including four ferric-thiolate bonds. A) Side view of the rubredoxin structure (PDB: 1BRF) shows a three-stranded β sheet and the FeS_4 center. B) The structure of rubredoxin shows the residues

enclosed by the FeS_4 center (in red), the CXXC motifs (color in blue) and the residues outside the center (in grey). C-D) simplified schematics of rubredoxin corresponds to figure A-B.

This work focuses on the ferric-thiolate bonds present in the small metalloprotein rubredoxin which belongs to the iron-sulfur protein family. Iron-sulfur protein, cytochromes and blue copper proteins are the three major metal-containing electron transfer proteins[154,159]. For example, rubredoxin participates in the electron transfer process of alkane hydroxylation and oxygen detoxication in some bacteria[160]. It is widely present in sulfur-metallolizing bacteria and archaea and proposed to be one of the most ancient proteins in nature. Rubredoxin is the simplest type of iron-sulfur protein with only one ferric ion. It was first discovered at the early 1960s and given the name rubredoxin because of its characteristic red color when in solution[161]. Rubredoxin shows a characteristic UV-Vis spectrum with adsorption maxima at 280, 390 and 494 nm. Visible absorption arises from the ligand to metal charge transfer of the ferric-thiolate bonds. This work mostly utilizes rubredoxin from *Pyrococcus furiosus* (abbreviated as *pf*RD or RD) as a model system. *P. furiosus* is an organism that grows at ~100 ℃[162]. Rubredoxin from *Clostridium pasteurianum* (*cp*RD) is also used.

As the first protein revealed with a high resolution structure by X-ray crystallography, the structural information of ferric-thiolate bonds in rubredoxin is abundant in the literature, with more than twenty crystal structures available[163-165]. For example, *pf*RD shows a typical $\alpha+\beta$ structure, with a three-stranded β-sheet at one side of the protein and a FeS_4 metal center in which a ferric ion is coordinated by four

sulfur atoms from cysteine residues in a tetrahedral geometry (Fig.1.8 A-D). The FeS_4 center divides the protein into two parts: the residues from 1 to 4 and 42 to 53 are outside the center, while the residues between 5 and 41 are trapped by the center[162,163,166].

The four cysteines in the FeS_4 center are arranged in a two CXXC binding motif (C5XXC8 and C38XXC41). The four corresponding ferric-thiolate bonds are present in two bond lengths: the bond between the ferric ion and the two exterior cysteines (Cys5 and C38) is longer: ~2.31 Å than the bonds with the two interior cysteines (Cys8 and Cys41) at ~2.25 Å. There is no other cysteine residue besides these four iron-coordinated ones within rubredoxin.

The high bond covalency which describes the electron mixing between the two bonding atoms is an important feature of the ferric-thiolate bond in rubredoxin. The sulfur K-edge experiments showed a high degree of electron sharing between the *d* orbital of iron and *p* orbital of sulfur[167,168]. Consequently, the ferric-thiolate bond in rubredoxin is often regarded as a highly covalent bond, which accounts for the excellent electron transfer properties of the protein. Interestingly, the same Fe-thiolate bonds in different types of rubredoxin are of different covalency: *pf*RD is ~125%, while *cp*RD, a homologue of *pf*RD, shows a value of ~135%, and an inorganic analog of the FeS_4 center in rubredoxin $[Et_4N][Fe(o\text{-}C_6H_4(CH_2S)_2)_2]$ is ~150%[169,170]. In addition, it is proposed that the protein environment around the FeS_4 center can modify the Fe-thiolate bond through the amide hydrogen bond. Here, the covalency reflects the degree of mixing present in the bonding orbital, and may related to the

bond strength[167].

The breaking of the FeS_4 center including the rupture of Fe-thiolate bond is an important and independent step during thermal and chemical denaturation of rubredoxin. By monitoring the decrease in the visible absorption spectrum of rubredoxin due to the ligand to metal charge transfer from the ferric-thiolate bonds, the rupture of the metal center event can be easily identified at an ensemble level. However, it cannot reveal how many ferric-thiolate bonds are indeed ruptured and the detailed bond rupture sequence. In addition, the *pf*RD is an extremely thermal stable protein with melting temperature of ~100 ℃. Thus, conclusive thermal or chemical denaturation experiments are very difficult[171-173].

In summary, rubredoxin is a small soluble protein with abundant structural information available. It has a simple FeS_4 center with four corresponding ferric-thiolate bonds. Thus, it represents an optimal model protein to investigate mechanical rupture (activation) of metal-ligand bonds in metalloproteins using AFM.

1.4 The aim of this thesis

To our knowledge, there are no force spectroscopy studies on a naturally occurring metalloprotein focusing on the rupture processes of metal-ligand bonds inside. To unambiguously identify the rupture events of these bonds, the development of a general method to build pure metal-form poly-metalloprotein is a necessary step. Demonstrating that the metal-ligand bonds in the metal center of metalloprotein can be ruptured under mechanical force is the next and most critical step in the study

presented here. Considering the multiple metal-ligand bonds structure present in metalloproteins, it is quite interesting to study the mechanism and sequence of bond rupture. Furthermore, the relationship between protein structure and metal-ligand bond strength, which is a unique feature of chemical bonds inside a protein structure, is of great interest. Finally, the chemical reactivity of the ferric-thiolate bonds in rubredoxin is an attractive aspect to investigate.

Chapter 2: A facile method for constructing polyproteins for single molecule force spectroscopy studies[1]

2.1 Synopsis

To study mechanical rupture processes of metal-ligand bonds in metalloproteins, the construction of poly-metalloproteins is a necessary step for AFM studies. However, it is challenging to make such polyprotein using the classic recombinant polyprotein building method, which is at the DNA-level. Compared with simple proteins, over-expression and purification of metalloproteins often results in the apo-form or non-native metal substituted form of the metalloprotein. Consequently, poly-metalloproteins built at the DNA level exhibit several different types of metal forms in the same molecule. This makes the assignment of a given rupture event to a specific metal-ligand bond extremely difficult. In addition, the recombinant DNA method is tedious and time-consuming. A method of building polyproteins at the protein level is therefore important for studying metalloproteins.

---

[1] A chapter of this chapter has been published as "[Zheng P.], Cao Y. and Li H. (2011) Facile method of constructing polyproteins for single molecule force spectroscopy studies. Langmuir 27(10): 5713-5718

In this chapter, we developed a facile maleimide-thiol coupling method to build poly-metalloprotein molecules at the protein level. At first, a well-studied non-metalloprotein GB1 was built to polyprotein $(GB1)_n$ using this method to demonstrate its feasibility and to optimize the reaction conditions. The AFM results on $(GB1)_n$ were identical to the results on $(GB1)_8$ built using the recombinant DNA method.

This method was then applied to rubredoxin to successfully produce pure iron-form poly-rubredoxin molecules suitable for AFM studies. The rubredoxin monomer protein with two cysteines was expressed first. However, a non-native zinc ion substituted rubredoxin was also expressed at the same time. The protein solution with a mixture of Fe and Zn form were purified and separated by a FPLC purification step using anion-exchange chromatography. The resultant pure Fe-form rubredoxin was then chemically coupled to form poly-rubredoxin through the maleimide-thiol reaction. Moreover, compared with disulfide bond, the resultant thioether bond between individual domains in the polyprotein is resistant to reducing agents. Thus, the mechanical stability of rubredoxin at different oxidation states can be probed.

2.2 Introduction

Protein mechanics plays an important role in a wide variety of biological processes[34,92,174]. Many proteins are subject to mechanical stretching force under their biological conditions and play various structural and mechanical roles. Elastomeric proteins are one representative class of such mechanical proteins, where many of

them are tandem modular proteins made of multiple individually folded domains[76,79]. They function as molecular springs under their biological settings to provide tissues with elasticity, extensibility and strength, and also play important regulatory roles in a wide variety of biological processes[175]. The development of single molecule force spectroscopy techniques, in particular atomic force microscopy (AFM)-based force spectroscopy, over the last two decades has made it possible to measure the mechanical properties of elastomeric proteins at the single molecule level, and AFN has provided tremendous insights into the mechanical design of elastomeric proteins[22,25,76,176,177]. Moreover, single molecule AFM has evolved into a general method to characterize the mechanical unfolding/folding dynamics of proteins along a well-defined reaction coordinate at the single molecule level.

In single molecule AFM studies, it is critical to unambiguously distinguish single molecule stretching events of the protein of interest from non-specific interactions between the AFM tip and sample as well as from non-single molecule events[20,178]. Constructing polyproteins made of multiple identical tandem repeats of the protein of interest provides unambiguous fingerprints for identifying single molecule events and has thus become a gold standard in single molecule AFM studies of protein mechanics. Since polyproteins are made of identical tandem repeats of individual domains, the resultant force-extension curves will exhibit a characteristic sawtooth-like force-extension pattern with identical contour length increments between consecutive sawtooth peaks. Such repetitive patterns allow one to readily and unambiguously identify single molecule stretching events from a myriad of non-specific interactions

and complex multi-molecule stretching events.

Construction of polyproteins is typically accomplished using a recombinant DNA technology-based DNA concatamerization method[42]. This method provides precise control of the molecular structure and composition of the resultant polyprotein, and has been the most widely used method[42,179,180]. However, this method requires stepwise repetitive cloning, making it laborious and expensive. To facilitate faster construction of polyproteins, alternative methods have been pursued. A cysteine engineering based method was developed for this purpose. This method was first successfully accomplished in a solid-state synthesis of polyproteins based on the specific crystallographic arrangement of proteins in their crystals[97] and was later generalized to proteins in solution[98,99]. In this method, a pair of cysteines was introduced into the protein of interest at chosen locations. The subsequent oxidation of cysteines results in the polymerization of the protein via the formation of intermolecular disulfide bonds. This method provides an efficient alternative to the recombinant DNA technology based polyprotein engineering approach and also allows the polyproteins to be stretched in any defined pulling direction. This method has therefore found many applications in a variety of proteins[24,181]. However, the oxidation of cysteines is a slow process and the disulfide linkage in the resultant polyproteins limits the use of this method for studying redox-dependent mechanical properties of proteins. Here, we report a facile method based on maleimide-thiol coupling chemistry to construct polyproteins in order to complement these existing polyprotein engineering methods.

2.3 Results

2.3.1 General principle

Maleimide-thiol coupling-based bioconjugation methods have been used extensively in biochemistry, due to the rapid and specific reaction of maleimide with sulfhydryl groups under mild experimental conditions[182,183]. Here we employ this maleimide-thiol coupling chemistry to develop an efficient method for constructing polyproteins for single molecule AFM studies.

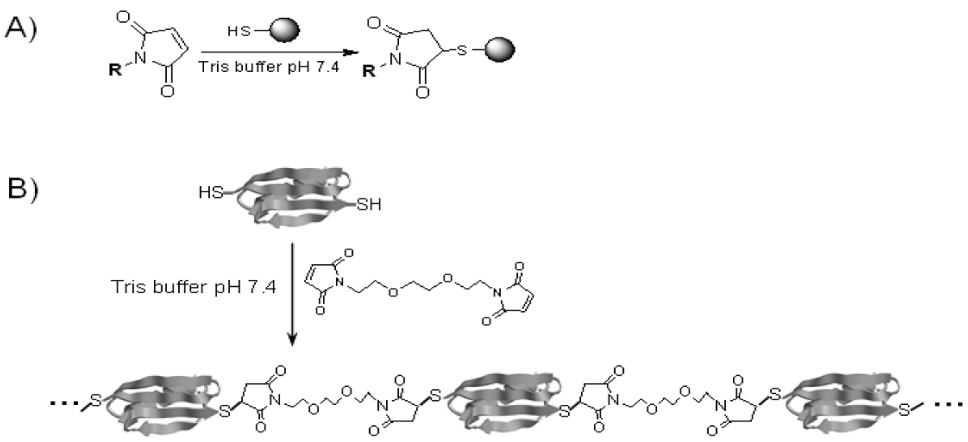

Figure 2.1 Construction of a polyprotein using the maleimide-thiol coupling reaction. A) The maleimide group can specifically react with the sulfhydryl group forming a covalent thioether bond that is un-cleavable under reducing conditions. B) $BM(PEO)_3$ can link two GB1s together by reacting sequentially with the sulfhydryl group on the protein via a maleimide group at its both end to form a stable thioether bond, and further extension leading to coupling-built polyprotein $(GB1)_n$.

The basic experimental design is schematically depicted in Fig. 2.1. First, two cysteine residues will be introduced into the protein of interest via site-directed mutagenesis at well-defined, solvent-exposed positions to provide sulfhydryl groups. Then, the protein will react with the bi-functional maleimide compound $BM(PEO)_3$ at a molar ratio of 1:1, leading to the polymerization of the protein of interest via the

formation of intermolecular thioether bonds. In the bi-functional maleimide compound $BM(PEO)_3$, PEO serves as a flexible linker with increased solubility. $(PEO)_3$ has a contour length of ~1.5 nm and can effectively minimize domain-domain interactions due to the formation of thioether bonds[183,184].

2.3.2 Synthesis and characterization of polyprotein $(GB1)_n$

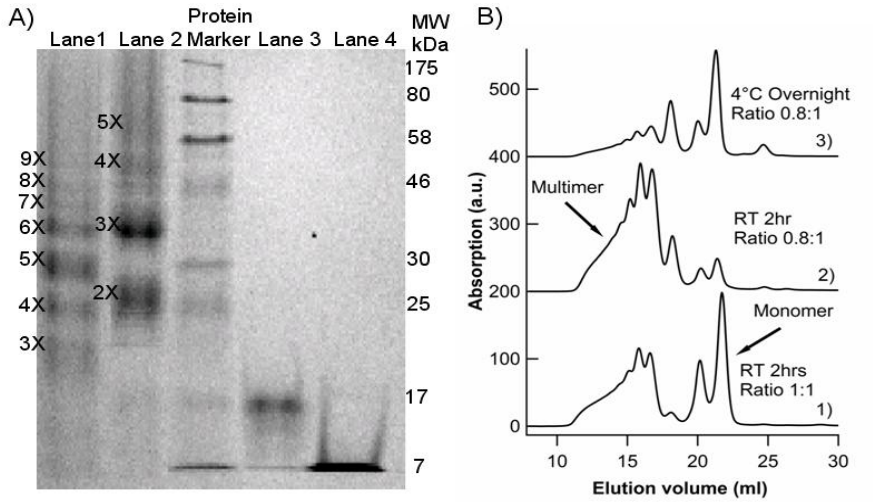

Figure 2.2 Characterization of polyproteins constructed using maleimide-thiol coupling chemistry. A) SDS-PAGE analysis of polymerization degree of polyproteins $(GB1)_n$ (lane 1) and $(GB1\text{-}RD)_n$ (lane 2) constructed using maleimide-thiol coupling chemistry. Monomeric protein GB1 (lane 4), RD-GB1 (lane 3) and protein makers are shown. B) Analysis of polymerization efficiency under different conditions by size-exclusion chromatography. Curve 1 shows a chromatograph of products with the reaction being carried out at room temperature for 2 hours with a molar ratio of 1:1; curve 2 corresponds to the reaction performed with a molar ratio of 0.8:1 for GB1 to $BM(PEO)_3$, and curve 3 is for the reaction carried out at 4 °C overnight.

To validate the design of this maleimide-thiol coupling-based polyprotein engineering method, we first constructed a model polyprotein of GB1which is an excellent elastomeric protein model domain[73,84]. Polyprotein $(GB1)_8$ has been constructed using a recombinant DNA based strategy and its mechanical properties

have been characterized in great detail in our previous single molecule AFM studies[73,84]. Thus, constructing GB1 polyproteins provides a good test for the proposed maleimide-thiol coupling approach.

By reacting Cys-GB1-Cys, which was pre-treated with β-mecaptoethanol at 4 °C overnight to prevent the formation of intermolecular disulfide bonds, with $BM(PEO)_3$ at a molar ratio of 1:1 in PBS buffer (pH 6.5) at room temperature for 2 hours, we found that polyprotein $(GB1)_n$ readily formed. From the SDS-PAGE gel (Fig. 2.2A), it is evident that the vast majority of monomeric GB1 has been converted to a GB1 multimer and the formation of multimeric GB1 (up to 9) can be readily identified. At higher molecular weight, the protein band becomes smeared, indicating the formation of a mixture of higher order multimers of GB1.

The reaction of Cys-GB1-Cys and $BM(PEO)_3$ is similar to the 2+2 condensation polymerization. The molecular weight of the resultant polymer depends on the stoichiometry of the two functional groups[185]. To obtain higher multimers, it is important to ensure that the two functional groups are present at a 1:1 molar ratio. However, due to the air oxidation of cysteine residues into disulfide bonds, the molar ratio of GB1 versus $BM(PEO)_3$ may not be necessarily equal to that of sulfhydryl groups versus maleimide groups. Thus, fine-tuning the molar ratio of GB1 and $BM(PEO)_3$ is necessary. As shown in Fig. 2.2B, due to the incomplete reduction of disulfide bonds, a 1:1 molar ratio of GB1 versus $BM(PEO)_3$ led to a large fraction of monomeric GB1 remaining in the reaction mixture, as monitored by analytical size exclusion chromatography (Fig. 2.2B, curve 1). By adjusting the ratio of

GB1:BM(PEO)$_3$ to 1:0.8, the fraction of remaining monomeric GB1 decreased after polymerization and the yield of multimeric GB1 increased significantly (Fig. 2.2B, curve 2).

We also found that the reaction temperature is an important parameter affecting the formation of multimeric GB1. It can be seen that incubating the reaction mixture at 4 °C overnight only led to the formation of a very small amount of oligomeric GB1, and that a large amount of GB1 remained monomeric (Fig. 2.2B, curve 3). However, reaction at room temperature for 2 hr led to efficient polymerization of GB1. Thus, reacting Cys-GB1-Cys with BM(PEO)$_3$ at room temperature provides an efficient alternative method for constructing polyproteins for single molecule AFM experiments.

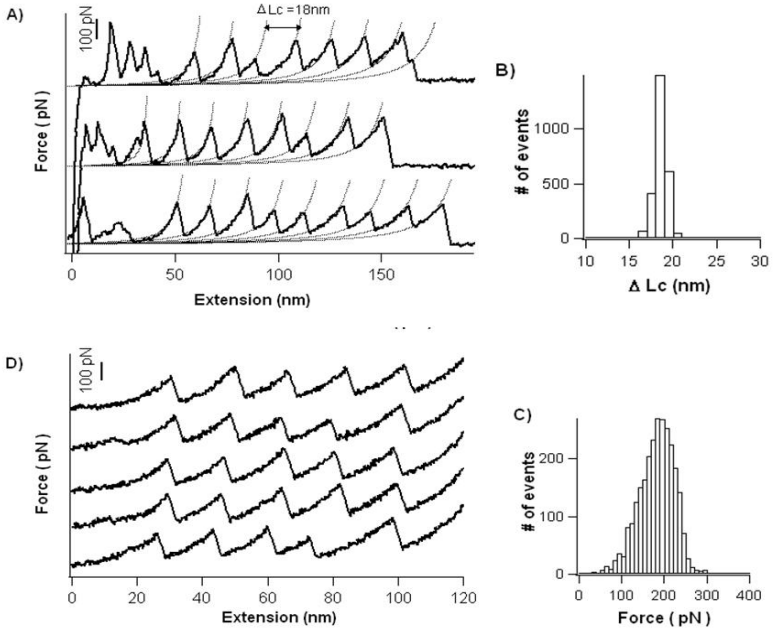

Figure 2.3 Mechanical unfolding and refolding of the polyprotein (GB1)$_n$ constructed using maleimide-thiol coupling chemistry and recombinant-DNA techniques. A) Typical force-extension curves from the unfolding of (GB1)$_n$ shows the characteristic contour length increment of ~18 nm which is identical to that of the recombinantly

produced polyprotein $(GB1)_8$. B) The histogram of contour length increment of $(GB1)_n$ shows 18.0 ± 0.7 nm. C) The histogram of unfolding force shows an average force of 182 ± 42 pN (n=2628). D) Five consecutive force-extension curves of polyprotein $(GB1)_n$ from a repeated stretching and relaxation experiment that lasted 100 cycles.

2.3.3 $(GB1)_n$ shows the same mechanical properties as $(GB1)_8$

The formation of the $(GB1)_n$ polyprotein was further confirmed by AFM experiments. Stretching the $(GB1)_n$ polyprotein results in characteristic sawtooth-like force-extension curves, where each individual peak corresponds to the mechanical unfolding of individual GB1 domains in the polyprotein chain (Fig. 2.3A). The force-extension curve shows as many as 17 unfolding events, suggesting the polyprotein contained 17 or more GB1 domains (the degree of polymerization is 17 or higher). Fitting force-extension curves using the worm-like-chain (WLC) model of polymer elasticity[31] revealed that the contour length increment ΔLc of GB1 is ~18 nm, identical to that of polyprotein $(GB1)_8$ constructed using a recombinant approach. The average unfolding force of GB1 is ~180 pN, again close to that of $(GB1)_8$ (Fig. 2.3B). These results suggest that the thioether linkage between GB1 domains does not affect the structural integrity and mechanical properties of GB1. In addition, the polymerized $(GB1)_n$ is also amenable for cyclic stretching-relaxation experiments, in which the folding kinetics of GB1 domains can be readily measured. Fig. 2.3 shows four consecutive stretching-relaxation curves from a $(GB1)_n$ polyprotein molecule that was repeatedly stretched and relaxed for more than 100 times (Fig. 2.3D). These results clearly indicate that the polyprotein $(GB1)_n$ constructed using maleimide-thiol

coupling chemistry is the same as $(GB1)_8$, suggesting that this method is suitable for constructing polyproteins for single molecule force spectroscopy studies.

2.3.4 Construction of the pure iron form poly-rubredoxin molecule

Many proteins require co-factors to carry out their biological functions. Metalloproteins are one representative class of such proteins[154]. Interactions between proteins and their co-factors can facilitate the folding of proteins and increase their thermodynamic stability. However, the existence of holo- and apo-forms of proteins leads to a mixture of proteins in different conformations. In addition, non-endogenous ligands can compete with endogenous ones, creating additional complexity when considering the composition of the protein[186]. For example, the metalloprotein rubredoxin can accommodate different metal ions in its tetra-cysteine metal binding center. When expressed in *E. coli*, two different forms, wild type Fe-rubredoxin and Zn-substituted rubredoxin, co-express naturally, leading to a mixture of the two different metal-containing rubredoxins[161,164,165]. These issues make it difficult, if not impossible, to use the recombinant DNA methodology to construct polyproteins where all individual domains are in the same form for single molecule AFM studies. Instead, the maleimide-thiol coupling method provides a means of overcoming this technical hurdle. Figure 2.5 shows one example of using the maleimide-thiol coupling method to construct a metalloprotein rubredoxin polyprotein from *Clostridium pasteurianum* (cpRD) for single molecule AFM experiments.

Rubredoxin is the simplest iron-sulfur protein containing one Fe(III) bound by

four cysteines[163,167] and has been used in our group as a model system for investigating the mechanical unfolding/folding dynamics of metalloproteins. To construct pure Fe(III)-cpRD for single molecule AFM experiments, we first constructed a Cys-GB1-cpRD-Cys chimera as a building block for the construction of polyprotein $(GB1\text{-}cpRD)_n$. We used anion exchange chromatography to separate Zn(II) and Fe(III) forms of cpRD and obtained Fe(III)-cpRD. Then, the purified Cys-GB1-Fe(III)-cpRD-Cys was used to react with $BM(PEO)_3$ (molar ratio 1:1) at room temperature for 2 hours to produce polyprotein $(GB1\text{-}cpRD)_n$, which contains Fe(III)-cpRD only. The formation of the polyprotein was verified by SDS-PAGE (Fig. 2.2A). In addition, the maleimide-thiol coupling reaction did not alter the characteristic UV-Vis spectrum of Fe(III)-cpRD (Fig. 2.4A), suggesting that the polymerization process does not affect the Fe-S center. This result also shows that the $BM(PEO)_3$ reacted specifically with the cysteine residues at the N- and the C-termini and the Fe(III)-coordinating cysteines did not react with maleimide groups.

Thus, we can now use single molecule AFM to investigate the mechanical unfolding process of Fe(III)-cpRD in a well-controlled manner using polyprotein $(GB1\text{-}cpRD)_n$, where the GB1 domains serve as single molecular fingerprints for identifying single molecule stretching events as well as being an internal force caliber[187]. Stretching polyprotein $(GB1\text{-}Fe(III)\text{-}cpRD)_n$ results in characteristic sawtooth-like force-extension curves, where the sawtooth peaks correspond to the mechanical unfolding of GB1 and Fe(III)-cpRD domains (Fig. 2.4C, curves 1-2). Unfolding events of GB1 are characterized by ΔLc of ~18nm (colored in black), while

unfolding events of ΔLc of ~13nm are due to the mechanical unfolding of Fe(III)-cpRD domains (colored in red), providing the characteristic signatures of the unfolding of cpRD.

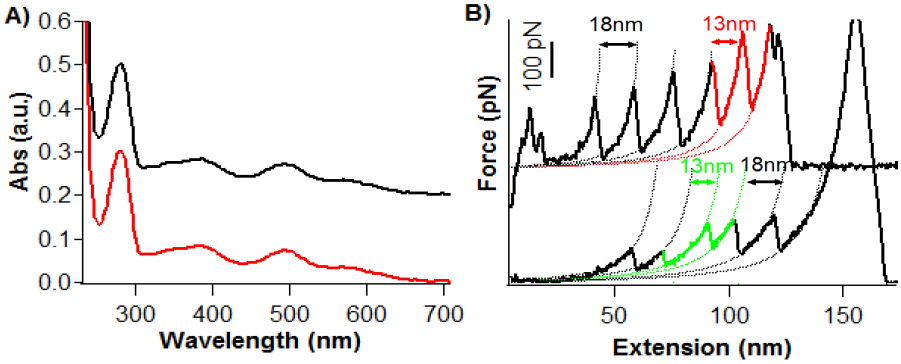

Figure 2.4 UV-Vis spectrum and AFM result of $(RD\text{-}GB1)_n$. A) UV-Vis spectra of RD-GB1 before (top panel) and after (bottom panel) reacting with $BM(PEO)_3$. For clarity, the two spectra are offset with each other. B) Mechanical unfolding of $(RD\text{-}GB1)_n$ under oxidized and reduced environment show the polyprotein is well-stable under reduced condition. They both show characteristic ΔLc of ~13 nm which is from rupture of Fe-thiolate center in Rubredoxin and a much lower unfolding force is observed in Fe(II)-RD-GB1(colored in green).

2.4 Discussion

In this chapter, we demonstrate a maleimide-thiol coupling-based facile method for constructing polyproteins for use with single molecule AFM studies. This new method makes it possible to construct polyproteins in a much faster fashion compared with the traditional recombinant DNA method. This approach allows for the construction of polyproteins consisting of pure forms of individual proteins to investigate the mechanical properties of each individual protein in an unambiguous way. In addition, it is possible to use this method to construct polyproteins to investigate the anisotropy of mechanical stability of proteins. Compared with

traditional recombinant DNA methods, the new approach does not have a precise control over the molecular weight of the polyprotein or the orientation of each individual domain in the polyprotein chain. In the polyprotein constructed using maleimide-thiol coupling method, the molecular weight of the polyprotein is polydispersed and the orientation of the individual protein domains are linked in either head-to-tail or head-to-head (tail-to-tail) fashion. This is in contrast to the monodispersed molecular weight and the uniform head-to-tail orientation (N to C) in polyproteins constructed using recombinant methods. However, these drawbacks do not affect the measurement of the mechanical properties of the protein of interest[98,99]

Although the method reported here is similar to the disulfide bond-based polymerization method in many ways, it does have some unique features. First, the maleimide-thiol-coupling chemistry is much more efficient. In two hours at room temperature, high molecular weight multimeric proteins can readily form. In contrast, the disulfide bond formation is much slower and the formation of polyproteins required longer periods of incubation[98]. Second, the PEG-linker sequence between the two maleimide groups provides a space between the two domains and may help prevent the unfavorable domain-domain interactions in the resultant polyprotein; Third and probably the most important, the resulting thioether bond is no longer sensitive to redox condition, as evidenced by fact that the polyprotein $(GB1)_n$ constructed using maleimide-thiol coupling reaction remains intact in the reducing SDS-PAGE in Fig. 2.2A. This makes it possible to investigate the mechanical properties of the polyproteins under reducing as well as oxidizing conditions[21,106]. Fig.

4B shows one such experiment in which the mechanical unfolding of Fe(II)-rubredoxin was investigated using single molecular AFM. Under reducing conditions, Fe(III)-cpRD was reduced to Fe(II)-cpRD. Stretching $(RD\text{-}GB1)_n$ resulted in similar sawtooth-like force-extension with ΔLc of 18 nm and 13 nm, indicating the polyprotein was still linked under reducing conditions (Fig. 2.4B Curves 3-4). It is evident that the unfolding of Fe(II)-rubredoxin occurs at much lower force as compared with that of Fe(III)-rubredoxin. Thus, the ferrous-thiolate bond is successfully measured and clearly shows the polyprotein is robust for studies in reducing environments.

In summary, we report a novel facile chemical coupling strategy to prepare polyproteins for single molecule force spectroscopy studies. This maleimide-thiol coupling-based strategy offers a new alternative method that not only complements the existing strategies, but also offers some unique advantages that will facilitate some special single molecule AFM experiments that are otherwise difficult to study using current polyprotein construction strategies. We anticipate that continuous development of polyprotein engineering will greatly facilitate protein mechanics studies and help to develop single molecule AFM into a general biophysical tool for protein characterization.

2.5 Materials and methods

2.5.1 Protein engineering

The model proteins we used in this study are GB1, the B1 IgG binding domain of

protein G from *Streptococcus*, and Fe(III)-rubredoxin from *Clostridium Pasteurianum* (cpRD). The gene coding cys-GB1-cys was amplified by using the polymerase chain reaction (PCR) and subcloned into expression vector pQE80L (Qiagen). The gene encoding cys-cpRD-GB1-cys was constructed using standard molecular biology techniques and cloned into expression vector pQE80L between *BamH*I and *Kpn*I sites. The sequences of both genes were verified by direct DNA sequencing. Both constructs were over-expressed in *Escherichia coli* strain DH5α and purified using Co^{2+}-affinity chromatography using TALON His-Tag purification resin (Clontech.). The purified protein was kept in PBS (Phosphate Buffered Saline) buffer at 4 ℃ with a concentration of ~ 2 mg/mL.

The pure Fe(III)-form of Cys-cpRD-GB1-Cys was obtained by further purification using anionic exchange chromatography. First, using a 9K MWCO pierce concentrator (Thermo Scientific), the protein chimera RD-GB1 solution was buffer exchanged into a 40mM Tris (Tris (hydroxymethyl) aminomethane) buffer at pH 8.5 and concentrated to ~8 mg/mL. The Fe form of cpRD-GB1 protein was eluted first at around 120 mM NaCl using a linear gradient elution (0-300 mM NaCl in 40mM Tris buffer and 1mM TCEP (Tris(2-carboxyethyl)phosphine) at pH 8.5) with a Mono-Q 5/50GL anion exchange column (GE Healthcare) in AKTA FPLC system (GE Healthcare) at a flow rate of 2 ml/min (Fig. 2.4A). The UV-Vis spectrum of purified Fe(III)cpRD-GB1 was recorded on a NanoDrop ND-1000 Absorption Spectrometer and the purity of Fe(III)-cpRD-GB1 was estimated to be >90% (based on its molar extinction coefficient 8.2 mM^{-1} cm^{-1} at 494nm)[164]. The protein was kept in Tris buffer

at pH 7.4.

In a typical experiment, 1 mL freshly purified Cys-GB1-Cys (or Cys-cpRD-GB1-Cys; at a concentration of 2 mg/mL) was reacted immediately with 40 μL of 10 mM concentrated stock solution of $BM(PEO)_3$ (1, 8-bis (maleimido)triethylene glycol, Molecular Biosciences) in Tris buffer at pH 7.4. The solution was incubated at room temperature for ~2 hours, and then the resultant polyprotein solution was used directly in AFM experiments.

SDS-PAGE was performed to examine the molecular weight of the resultant polyproteins. The polymerization efficiency of GB1 was also analyzed using gel filtration chromatography using a superdex 75 10/300GL column in an ATKA FPLC system (GE Healthcare). The buffer contains 50 mM sodium phospate plus 100mM sodium chloride at pH 6.5, and the flow rate was 0.5 mL/min.

2.5.2 Single molecule AFM experiments

Single-molecule AFM experiments were carried out on a custom-built AFM as described previously[46]. Each Si_3N_4 cantilever (Veeco Corp.) was calibrated in solution using the equipartition theorem before each experiment to obtain the spring constant (typically around 60 pN/nm). All experiments were done in Tris buffer at pH 7.4 at room temperature.

In a typical experiment, 2 μL polyprotein sample (at a concentration of ~2 mg/mL) was added onto a clean glass coverslip covered by ~50 μL Tris buffer. The protein was allowed to absorb for ~5 minutes before starting the AFM measurements.

During the experiment, the cantilever was brought into contact with the substrate at a contact force of ~1 nN to pick up proteins. The pulling speed was 400 nm/s.

For experiments on the Fe(II)-cpRD-GB1 polyprotein, 10 μL of 200mM dithioreitol (DTT) was added to the solution after the $(Fe(II)\text{-}cpRD\text{-}GB1)_n$ protein was absorbed on the glass cover slip. AFM experiments started after incubation for ~20 minutes. The same amount of DTT solution was added every hour thereafter to ensure a reduced environment.

Chapter 3: The Fe-thiolate bond in rubredoxin shows a surprisingly low mechanical strength[2]

3.1 Synopsis

After the successful construction of poly-rubredoxin molecules in pure iron-form, the mechanical rupture processes of metal-ligand bonds in a metalloprotein were investigated. We found that the ferric-thiolate bonds in rubredoxin can be ruptured under mechanical force, showing a characteristic contour length increment upon breaking of the FeS_4 center. This is the first example in which metal-ligand bonds in a biological system have been cleaved under mechanical force. Surprisingly, the measured rupture force of these bonds is very low ~ 200 pN. This is one magnitude lower when compared to the bond strength of a typical covalent bond. In addition, we also indirectly measured the bond strength of a single ferric-thiolate bond and found that this value is dependent on bond length. Finally, based on the rupture force of the FeS_4 center in different types of rubredoxins, we propose that the mechanical

[2] A version of this chapter has been published as "[Zheng P.], Li H. (2011) Highly covalent ferric-thiolate bonds exhibit surprisingly low mechanical stability. Journal of the American Chemical Society, 133(17): 6791-6798"

ferric-thiolate bond strength is related to the bond covalency which describes the degree of electron sharing between ferric ion and sulfur atoms in the protein.

3.2 Introduction

Chemical bonding, which describes the degree of interactions between atoms, is a fundamental concept in chemistry. Depending on their nature, different chemical bonds display vastly different strength and stability, with covalent bonds showing the highest degree of stability. Single molecule force spectroscopy studies[1,14] have shown that the mechanical stability of chemical bonds, defined as the most probable force at which the bond ruptures, of chemical bonds also follow this hierarchy: the rupture force of non-covalent bonds such as hydrogen bonds ranges from a few pN to a few tens of pN, while covalent bonds rupture at forces that are orders of magnitude higher, ranging from 1.4 to 3 nN[1,14,37,53,61,188,189]. The mechanical strength of chemical bonds provides new information about the characteristics of chemical bonds that are complementary to the classical thermodynamic descriptions[1].

Fe-S bonds are ubiquitous in nature and an essential component of a myriad of proteins[159]. These bonds are highly covalent[167,168], making Fe-S-containing proteins suitable electron transfer proteins[154,159,190] as well as facilitating protein folding and maintaining their overall three-dimensional structures[154]. Due to their highly covalent nature, it is believed that these Fe-S bonds are mechanically stable. However, no experimental study is thus far available to prove it. To understand the mechanical nature of such high covalent Fe-S bonds, we combined single molecule AFM and

protein engineering techniques to carry out the first direct experimental measurement on the mechanical strength of these ferric-thiolate bonds in rubredoxins. For this purpose, rubredoxin from *P. furiosus* was used as a model system. Rubredoxin is the simplest iron-sulfur protein, which consist of one iron atom bound by four cysteinyl sulfur atoms (Fig. 3.1A)[162,166]. Fe(III)-thiolate bonds are highly covalent with a total covalency of ~130% as measured by sulfur K-Edge X-ray absorption spectroscopy (S K-edge XAS)[169]. Upon reduction, these Fe-S bonds lengthen by an average of 0.033 Å accompanied by a reduction in their covalency[163,191]. The four coordinating cysteine residues are grouped into two CXXC chelating motif (C5XXC8 and C38XXC41) and are highly conserved in different types of rubredoxins[162,191]. Thus, rubredoxin provides an ideal model system to study the mechanical strength of highly covalent Fe-S bonds.

The AFM-based single molecule force spectroscopy has evolved into a powerful tool to investigate the mechanical activation (bond rupture) of chemical bonds, ranging from noncovalent bonds (such as hydrogen bonds) to covalent bonds (such as C-Si bonds), as well as elucidating the influence of the stretching force on chemical reactions[1,14,49,76,106]. Here we used single molecule AFM to directly measure the mechanical bond strength of Fe-thiolate bonds in rubredoxin and investigate the nature of their mechanical activation.

3.3 Results

In single molecule AFM experiments, the construction of polyproteins or

polyprotein chimeras is necessary for identifying mechanical unfolding signatures of proteins in an unambiguous fashion[42]. Since Zn-substituted RD and Fe-RD are co-expressed in *E. coli*, the conventional DNA-concatamerization method for polyprotein construction will lead to the production of mixed metal-containing rubredoxin[162]. To overcome this hurdle, we developed a novel chemical coupling method based on maleimide-thiol chemistry to construct polyprotein chimera $(Fe(III)RD\text{-}GB1)_n$, in which rubredoxin exists solely as Fe(III)-RD[192], as demonstrated in Chapter 2. In $(Fe(III)RD\text{-}GB1)_n$, Fe(III)-RD alternates with the well-characterized GB1 domain, which was used as the internal force caliber and fingerprint for identifying single molecule stretching events for $(Fe(III)RD\text{-}GB1)_n$[73,187].

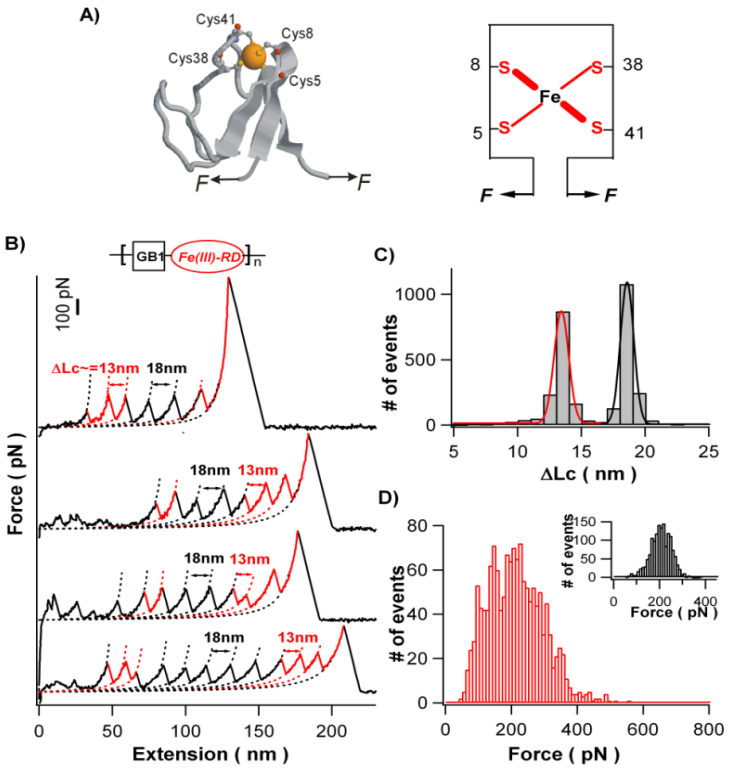

Figure 3.1 Mechanical unfolding experiments on the $(Fe(III)\text{-}RD\text{-}GB1)_n$ polyprotein

revealed that the Fe(III)S_4 center ruptures at low forces. A) 3D structure of Fe(III)-*pf*RD. Right panel schematically shows the pulling geometry of Fe(III)-RD in single molecule AFM experiments. B) Typical force-extension curves of $(Fe(III)\text{-}RD\text{-}GB1)_n$ characterized by two groups of unfolding events. The unfolding events of ΔLc of ~13 nm (colored in red) are attributed to the unfolding of Fe(III)-RD domains. C-D) Mechanical unfolding signatures of Fe(III)-RD. The ΔLc of Fe(III)-RD shows a distribution (shown in C)) with an average of 12.6 ± 1.3 nm (n = 1421 The ΔLc histogram for GB1 is shown in black in C) with an average of 18.2 ± 0.8 nm (n=1534). The histogram of the rupture forces of FeS_4 (shown in D)) is characterized by a very broad distribution, with an average of 211 ±86 pN (n=1421).

3.3.1 The FeS_4 center in rubredoxin ruptures at ~200 pN

Stretching the $(GB1\text{-}RD)_n$ polyprotein results in force-extension curves with a characteristic sawtooth-like appearance (Fig. 3.1B), in which individual force peaks correspond to the mechanical unfolding of individual domains and the last peak corresponds to the stretching of the fully unfolded polypeptide chain and its subsequent detachment from either the AFM tip or glass substrate. The force-extension curves of $(RD\text{-}GB1)_n$ are characterized by two populations of unfolding force peaks. Fitting the Worm-like chain model of polymer elasticity[31] to consecutive force peaks revealed that one group of force peaks displayed contour length increments ΔLc of 18.2 ± 0.8 nm (average±standard deviation) with number of events (n) of 1534, which is the unfolding signatures of the well-characterized GB1 domains[73,187], while the other group (colored in red) showed ΔLc of 12.6 ± 1.3 nm (n=1421) (Fig. 3.1C). Since rubredoxin alternates with GB1 in $(GB1\text{-}RD)_n$, the unfolding events of ΔLc of ~13 nm can thus be attributed to the mechanical unfolding of Fe(III)-RD. Single molecule AFM experiments on the polyprotein $(I27\text{-}RD)_4$, where a different fingerprint domain I27 was used, showed the same unfolding events

of ΔLc of ~13 nm and further validate their origin to the mechanical unfolding of Fe(III)-RD.

The contour length increment ΔLc upon domain unfolding is an important structural parameter that can provide detailed information about the unfolding mechanism of a protein[106]. If the structure of the protein is known, ΔLc can also be calculated according to: ΔLc=Lc(unfolded)-Lc(folded), where Lc(unfolded) is the length of the unfolded and fully extended polypeptide chain and Lc(folded) is the distance between the N- and C-termini in the folded structure. If there is a strong bond (such as a disulfide bond) in the protein structure linking two parts of the protein, ΔLc will be affected, as the sequence between the strong bond will be sequestered and shielded from the stretching force, as demonstrated in the disulfide bond mutants of I27[21,105,106]. Rubredoxin contains 53 residues. The highly covalent Fe(III)-thiolate bonds could serve as a strong bond such that the rubredoxin sequence enclosed in the FeS_4 center (residues 5 to 41) is sequestered and shielded from the stretching force until the FeS_4 center ruptures. Thus, when the FeS_4 center is intact, rubredoxin can only partially unravel, leading to the unfolding and stretching of the polypeptide sequence from residues 1 to 5 and 41 to 53. Such a partial unfolding would result in unfolding events of ΔLc of 5.3 nm ((5+13)aa*0.36nm/aa-1.2nm, where 1.2 nm is the distance between the N- and C-termini of rubredoxin). However, we did not observe such unfolding events, suggesting that residues 1 to 5 and 41 to 53 unfold at low forces that are below our AFM detection limit (~20 pN). Instead, we observed unfolding events with a ΔLc of ~13 nm. A ΔLc of ~13 nm corresponds to the exposing

of a polypeptide of 36 aa (36*0.36 nm/aa) to a stretching force during the mechanical unraveling of rubredoxin. In rubredoxin, there are only 18 aa (residues 1-5, 41-53) outside the FeS_4 center. During the unfolding of Fe(III)-RD, if Fe-thiolate bonds were not broken, it would be impossible to obtain ΔLc of as long as 13 nm. Therefore, the unfolding of Fe(III)-RD must involve the breaking of the Fe(III)-thiolate bonds. In fact, based on the three-dimensional structure of rubredoxin, it is expected that the unraveling of the FeS_4 center and subsequent unfolding of the remainder of rubredoxin (residues 5 to 41) would result in unfolding events of ΔLc of 12.4 nm (37aa*0.36 nm/aa-0.9 nm). This expected value is in close agreement with our experimentally determined ΔLc of rubredoxin, strongly indicating that the observed unfolding events of rubredoxin correspond to the rupturing of FeS_4 center and the subsequent unfolding and extending of rubredoxin.

Based on these results, the mechanical unfolding of Fe(III)-RD occurs in two steps: the first step is the mechanical unraveling of protein structure outside the FeS_4 center (which occurs at low forces); and the second step is the mechanical rupture of the FeS_4 center followed by the unfolding and extension of the remaining structure (residues 5 to 41), leading to the observed unfolding events with a ΔLc of ~13 nm. It is important to note that in order to rupture the FeS_4 center, at least two Fe(III)-thiolate bonds from the same side of FeS_4 center need to be broken completely (Fe-C5/Fe-C8 or Fe-C38/Fe-C41). However, we do not know exactly how many Fe(III)-thiolate bonds are broken during this mechanical unfolding process and whether Fe(III) is still attached to rubredoxin after the FeS_4 center has been

mechanically ruptured. In principle, depending on the number of Fe-S bonds that are broken during unfolding, slightly different ΔLc value should be observed (where the difference is ~1 nm). However, the length resolution of our measurements is not sufficient to allow us to unequivocally determine the number of Fe-thiolate bonds that break during the AFM experiments. To address this issue, we developed a loop insertion rubredoxin mutant to increase the length increment upon single ferric-thiolate bond rupture which is fully described in the next chapter.

Having confirmed that the unfolding events with a ΔLc of ~13 nm correspond to the mechanical rupture of the FeS_4 center, we then measured the rupture force of FeS_4 center. Fig. 3.1D shows the rupture force histogram of FeS_4 center at a pulling speed of 400 nm/s. The rupture force showed a very broad distribution from 100 pN to 500 pN, with an average rupture force of 211 ± 86 pN (n = 1421). This result suggested that the rupture of the FeS_4 center formed by four highly covalent Fe(III)-thiolate bonds occurred at forces of ~200 pN, which are surprisingly low compared to typical covalent bond strength[1,14].

To ensure that we did not miss unfolding events of rubredoxin occurring at forces that are higher than the detachment force, we measured the ratio of the number of unfolding events for rubredoxin versus GB1. Since GB1 alternates with rubredoxin in the polyprotein, the number of rubredoxin unfolding events should be roughly equal to that of GB1. Indeed, the force-extension curves shown in Fig. 3.1B contained similar number of unfolding events of rubredoxin and GB1. The overall ratio between rubredoxin and GB1 unfolding events is 0.93:1, which is close to the theoretical ratio

of 1. This result indicates that the rupture force histogram for the FeS_4 center genuinely reflects the true mechanical strength of the highly covalent ferric-thiolate bonds. And the broad distribution of the rupture force is clearly beyond the experimental errors and reflects the intrinsic short distance from the bound state to the mechanical dissociation transition state for ferric-thiolate bonds[37].

3.3.2 Unfolding of apo-rubredoxin shows no detectable rupture force

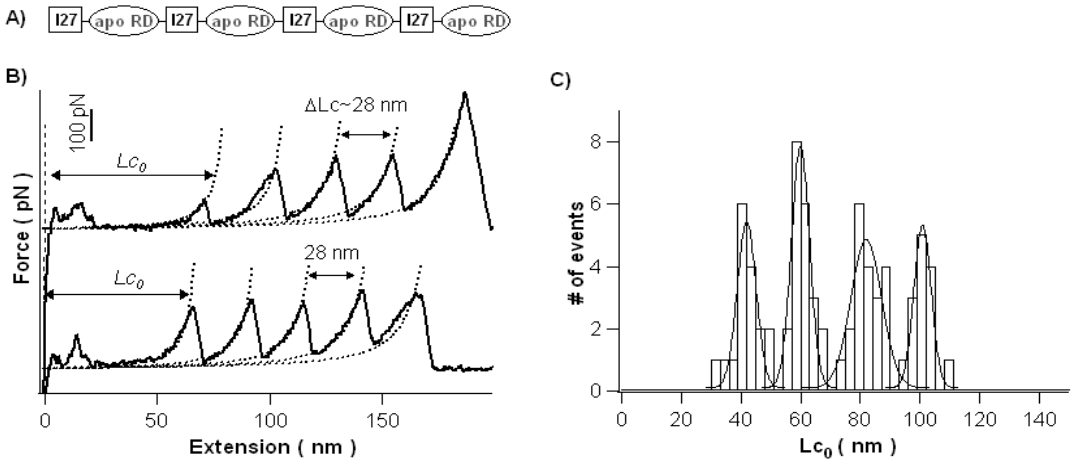

Figure 3.2 Mechanical unfolding of the $(I27\text{-}apoRD)_4$ polyprotein indicates that the unfolding of the RD protein structure does not contribute to the rupture force of Fe(III)-RD. A) A schematic of the polyprotein $(I27\text{-}apoRD)_4$. B) The force-extension curves of $(I27\text{-}apoRD)_4$ are characterized by featureless spacer followed by the unfolding events of I27 domains, which are characterized by ΔLc of ~28 nm. Lc_0 measures the length of featureless region of the polyprotein, which largely correspond to the length of unfolded and fully stretching apo-RD proteins in the polyprotein construct. C) The distribution of Lc_0 shows four clearly separated peaks (n = 72), corresponding to Lc_0 of different length of polyprotein $(I27\text{-apo-RD})_4$ fragments.

To confirm that the unfolding of the secondary structures of rubredoxin does not contribute to the rupture force histogram of the ferric-thiolate bond shown in Fig. 3.1D, we measured the unfolding force of a pseudo apo-rubredoxin (apo-RD). This pseudo apo-RD, a computationally designed rubredoxin mutant in which four

iron-coordinating cysteines were mutated to alanine, lysine and threonines, foled into the same three-dimensional structure as wild type Fe(III)-RD, but does not have a FeS_4 center[193]. If apo-RD is mechanically stable and unfolds in a two-state fashion, the unfolding of apo-RD will result in unfolding events with ΔLc of ~17 nm (53aa*0.36nm/aa-1.2nm, where 1.2 is the distance between the N, C-termini of apo-RD). However, this contour length increment is similar to that GB1, making the identification of apo-RD unfolding events difficult. To avoid this potential complication, we used the well characterized I27 domain as the fingerprint domain[42], as the unfolding of I27 leads to ΔLc of ~ 28 nm. We constructed a $(I27\text{-}apoRD)_4$ polyprotein for single molecule AFM experiments (Fig. 3.2A). In the vast majority of force-extension curves, we observed that stretching $(I27\text{-}apoRD)_4$ results in sawtooth-like force-extension curves characterized by a long featureless spacer followed by unfolding events with a ΔLc of ~28 nm corresponding to the unfolding of I27 domains (Fig. 3.2B). Hence, the long featureless spacer originates from the unfolding and stretching of the pseudo apo-RD, suggesting that apo-RD is mechanically labile and unfolds at forces that are below the detection limit of our AFM. The distribution of the length of the featureless spacer shows four clearly separated peaks, corresponding to the stretching of a different number of I27-apo-RD repeating units. The measured length is ~20 nm per apo-RD, close to the expected contour length of fully extended apo-RD (~19 nm). This result corroborated that the rupture force histogram of wt-RD largely results from the rupture of theFeS_4 center.

3.3.3 The distance to the transition state of FeS_4 is ~0.11 nm

The broad distribution of rupture force for FeS_4 is clearly beyond the experimental error in our single molecule AFM experiments and reflects the intrinsic energy landscape underlying the mechanical activation process. The width of the distribution is related to the distance from the bound state to the mechanical dissociation transition state (Δx_u)[32,33]. A broader distribution implies a shorter Δx_u. To quantitatively measure the Δx_u during the mechanical rupture of FeS_4, we carried out single molecule AFM experiments at different pulling speeds (Fig. 3.3). Based on the Bell-Evans model, we carried out Monte Carlo simulations to reproduce speed-dependence results[32,33,43]. The data can be reproduced using a Δx_u of 0.11 nm and a spontaneous dissociation rate of 0.15 s^{-1}. The distance to the transition state is about half of the bond length of the Fe(III)-thiolate bond[162,166] and the distance is smaller than that for typical protein unfolding (~0.2 nm) but larger than that of disulfide bond (~0.2 Å). These differences reflect the unique nature of Fe(III)-thiolate bonds. It is of note that data obtained here likely reflects the average properties of the FeS_4 center.

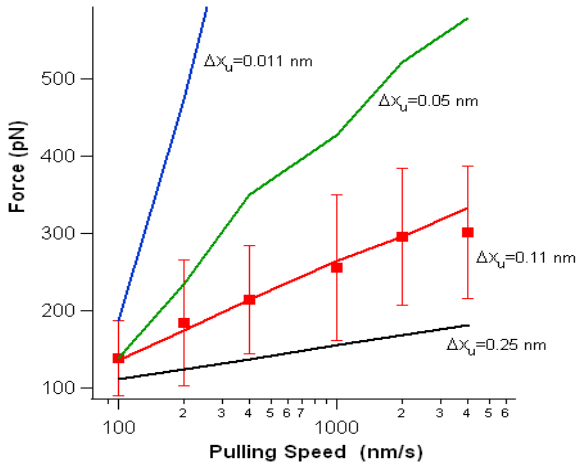

Figure 3.3 The rupture force of FeS_4 center of rubredoxin depends on the pulling velocity. The red line corresponds to the Monte Carlo simulation using a Δx_u of 0.11 nm and a spontaneous dissociation rate of 0.15 s^{-1}. For comparison, the simulated pulling velocity dependence of the unfolding force of I27 (Δx_u of 0.25 nm, k_0 = 3.3×10^{-4} s^{-1})[42] and rupture force of disulfide bond in the presence of ~0.25mM hydroxide anions (Δx_u = 0.011 nm, k_0 = 0.13 s^{-1})[106] are shown. In addition, simulated pulling velocity dependence of the rupture force of a hypothetical bond (Δx_u of 0.05 nm, k_0 = 0.13 s^{-1}) is also shown. For clarity, the data is offset with each other.

3.3.4 The two types of ferric-thiolate bonds display different stabilities.

From the geometry of the FeS_4 center, it is clear that in order to fully unfold and extend rubredoxin, at least two Fe(III)-thiolate bonds in the same CXXC chelating motif should be ruptured (Fig. 3.1A). To prove this point, we engineered a double histidine mutant of rubredoxin C38,41H-RD, in which both Cys38 and Cys41 in the same chelating motif were substituted by histidines. Since the interactions between the Fe and N atom from histidine are much weaker than Fe(III)-thiolate bonds, we anticipated that the mutation C38HC41H would significantly weaken rubredoxin. Indeed, force-extension curves of $(GB1C38,41H\text{-}RD\text{-}GB1)_n$ showed that the unfolding of the majority of C38,41H-RD occurred at very low forces and behaved

like as long featureless spacers (Fig. 3.4A). Only a small fraction of C38,41H-RD domains showed clear unfolding events of ΔLc of ~13nm with a rupture force close to ~100 pN, leading to the observation that the number of C38,41H-RD unfolding events is only 17% of the GB1 unfolding events (93 versus 540) (Fig. 3.4B). Mutating Cys5 and Cys8 to histidines led to a similar destabilization effect. These results clearly indicated that removing the two Fe(III)-thiolate bonds from the same CXXC chelating motif significantly weakens rubredoxin, emphasizing the important roles of the two CXXC chelating motif in the mechanical stability of rubredoxin. Therefore, the two Fe-thiolate bonds in the same CXXC loop are required to provide the mechanical stability for rubredoxin.

Since the two types of Fe(III)-thiolate bonds are of different bond length, it is possible that they display different mechanical stability[163]. To experimentally test this hypothesis and dissect the difference in the mechanical strength of these two types of Fe(III)-thiolate bonds, we engineered two rubredoxin mutants C38H-RD and C41H-RD. Since Fe(III)-N bond is more mechanically labile than a Fe(III)-thiolate bond (Fig. 3.4B), selectively mutating one of the two cysteines in the same CXXC motif with histidine should allow us to determine the mechanical rupture force of the two different Fe(III)-thiolate bonds.

Stretching polyprotein $(GB1\text{-}C38H\text{-}RD)_n$ resulted in sawtooth-like force-extension curves which are similar to those of wt-rubredoxin. The unfolding events of C38H-RD, thus the rupture of Fe(III)-thiolate bonds, are characterized by ΔLc of 12.0 ± 1.8 nm and an average unfolding force of 203 ± 92 pN (n = 332) (Fig.

3.4C). In addition, the number of unfolding events of C38H-RD is ~72% of that of GB1 (332 versus 460). These results suggested that the mutation C38H does not change the mechanical stability of the Fe-S center significantly. Since the Fe(III)-thiolate bond formed by Cys41 remains the force-bearing bond, our results suggested that the rupture force measured on C38H-RD likely reflects the mechanical stability of the Fe(III)-S41 bond. It is of note that C38H-RD mutant is not as stable as wt rubredoxin, as we have observed that C38H-RD can lose its iron ion over time. Thus, the 28% missing events for C38H-RD are likely due to the unfolding of C38H-RD at low forces.

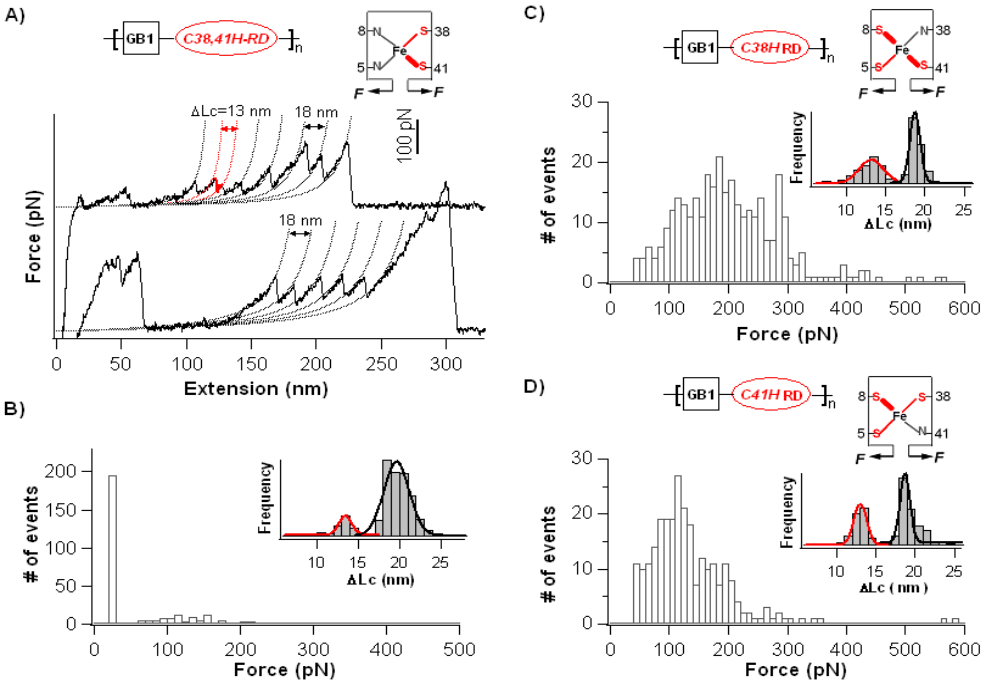

Figure 3.4 Dissection of the mechanical stability of the two types of ferric-thiolate bonds in the FeS_4 center. A) AFM experiments revealed that mechanical unfolding of C38,C41H-RD occurs at very low forces. B) Unfolding force histogram of C38,41H-RD. The majority of C38,41H-RD occur at forces below 20 pN. C) The mechanical unfolding force of Fe(III)-C38H-RD remains largely unchanged. The

average unfolding force of Fe(III)-C38H-RD is 203±92 pN (n=332). D) Fe(III)-C41H-RD unfolds at significantly reduced forces (121±74 pN (n = 262)).

In contrast to C38H-RD, the unfolding events of C41H-RD were observed to occur at much lower forces 121 ± 74 pN (n=262) with ΔLc of 12.6 ± 0.8 nm (Fig. 3.4D). In addition, the number of C41H-RD unfolding events is clearly fewer than that of GB1 (262 versus 472), suggesting that some "missing" C41H-RD domains may unfold at forces below 20 pN. These results indicate that mutation C41H significantly weakens the mechanical stability of rubredoxin. Due to the mutation of C41H, the force-bearing Fe(III)-thiolate bond shifted to the Fe(III)-S38 bond. Therefore, the measured rupture force for C41H-RD likely reflects the mechanical stability of the Fe(III)-S38 bond. It is of note that the mechanical strength for Fe(III)-S38 and Fe(III)-S41 bonds is fittingly correlated with the bond length of these two types of Fe(III)-thiolate bonds. Thus, our results indicated that the two types of Fe(III)-thiolate bonds are not equivalent in terms of their mechanical strength. The shorter Fe(III)-thiolate bonds (Fe(III)-S8 and Fe(III)-S41) are mechanically stronger than the longer Fe(III)-thiolate bonds (Fe(III)-S5 and Fe(III)-S38). To our best knowledge, this is the first direct experimental evidence that the bond strength of the two types of Fe-thiolate bonds is different. It is of note that although histidine mutation may slightly alter the structure of the Fe-S center in rubredoxin, the difference in bond length of the two types of Fe(III)-thiolate bonds should remain similar: Fe-S(5) and Fe-S(38) bonds should be longer while Fe-S(8) and Fe-S(41) bonds be shorter. This trend has been observed experimentally in Cys to Ser mutants of rubredoxin in X-ray crystallographic studies[194]. Thus, our results on histidine

mutants of rubredoxin can provide a reasonable estimate of the bond strength of individual Fe-thiolate bonds in rubredoxin.

3.4 Discussion

Our results clearly demonstrate that the mechanical rupture force of the highly covalent Fe(III)-thiolate bonds is ~200 pN (with the shorter ferric-thiolate bonds being stronger than the longer ones), which is significantly lower than what one would expect for a highly covalent bond. For example, Si-Si ruptures at 2.1nN, C-Si at 2.0nN and Au-S at above 2.5nN[1,14,49,188]. In addition, the rupture force of Fe-thiolate bond is also significantly lower than that of a disulfide bond (~1.4 nN), which is generally perceived as a weaker covalent bond[14,21]. In contrast, the mechanical strength of ferric-thiolate bonds is more comparable to the mechanical strength of non-covalent bonds, such as hydrogen bonds. For example, the unbinding force of the avidin-biotin complex is ~200 pN[37,40]. The mechanical unfolding force of mechanically stable proteins ranges from ~50 pN to 300 pN, with a few proteins unfolding at more than 500 pN[99,195]. The unfolding of such non-metalloproteins corresponds to the rupturing of a network of non-covalent bonds, including hydrogen bonds and hydrophobic interactions. Thus, the finding that the highly covalent ferric-thiolate bonds rupture at ~200 pN is surprising, suggesting that such Fe-thiolate bonds are mechanically labile and display features clearly distinguishing themselves from those of typical covalent bonds.

Chemical bonding reflects the degree of interactions between two atoms and can

be described by covalency, the amount of covalent mixing of their atomic orbital[196]. The highly covalent nature of ferric-thiolate bonds originates from the high degree of mixing of the p orbital of a sulfur atom with the d orbital of a iron atom[159,167,196]. S-K edge XAS has been used extensively to probe the covalency of metal-thiolate bonds[168]. It was discovered that the covalency of ferric-thiolate bonds depends on the chemical environment in which the ferric-thiolate bonds are located. Due to the formation of hydrogen bonds in rubredoxin, the covalency of the ferric-thiolate bond in rubredoxin is significantly lower than that of its inorganic analogues $[Et_4N][Fe(o\text{-}C_6H_4(CH_2S)_2)_2]$ (~150%)[169,170]. These observations raise interesting questions about the relationship between covalency and mechanical strength of Fe-thiolate bonds, and the strength of the ferric-thiolate bond in the inorganic analogue. To address these issues, we measured the mechanical stability of Fe(III)-thiolate bonds in *clostridium pasteurianum* rubredoxin (*cp*RD) and the Fe(II)-thiolate bond in *pf*RD. *cp*RD is a homologue of *pf*RD with an identical FeS_4 center, but shows a higher covalency (~135%) than *pf*RD (~125%)[169]. Moreover, reduction of Fe(III) to Fe(II) leads to a significant decrease in the Fe-S covalency to ~80%[196]. These proteins provide ideal model systems to investigate the relationship between covalency and mechanical strength.

Our single molecule AFM experiments revealed that the rupture force of Fe(III)-thiolate bonds in *cp*RD shows a broad distribution with an average force of 258 ± 122 pN (N=686) at a pulling speed of 400 nm/s (Fig. 3.5A&C), significantly higher than that of *pf*RD. In contrast, the unfolding force for Fe(II)-thiolate is 152 ±

62 pN (n=579) (Fig. 3.5B&D), which is ~30% lower than the rupture force of Fe(III)-thiolate bonds in *pf*RD. It is evident that there is a positive correlation between the covalency and mechanical strength of the Fe-thiolate bond: the higher the covalency, the higher the rupture force (Fig. 3.5E). Clearly the mechanical strength of Fe(III)-thiolate bond in the inorganic analogue $[Et_4N][Fe(o\text{-}C_6H_4(CH_2S)_2)_2]$ will be stronger than that in rubredoxin, possibly as high as 350 to 500 pN. Although the exact value is yet to be experimentally determined, the mechanical strength of the Fe(III)-thiolate bond in $[Et_4N][Fe(o\text{-}C_6H_4(CH_2S)_2)_2]$ will be still significantly lower than that of a typical covalent bond. Therefore, from a mechanical perspective, the metal-thiolate bond itself is distinctly different from covalent bonds and can thus be only considered as labile highly covalent bonds. Moreover, the protein environment may also play important roles in determining the mechanical strength of Fe-thiolate bonds. It is well known that in rubredoxin the hydrogen atom of a nearby amide can form backbone hydrogen bonds with the cysteine sulfur atoms. The formation of these hydrogen bonds decreases the ability of the S atom binding with iron and lead to the reduction in covalency[1,167]. Hence, the reduced mechanical stability of the Fe-thiolate bond in rubredoxin is also likely due to the unique environment in rubredoxin. This hypothesis is experimentally tested in Chapter 6.

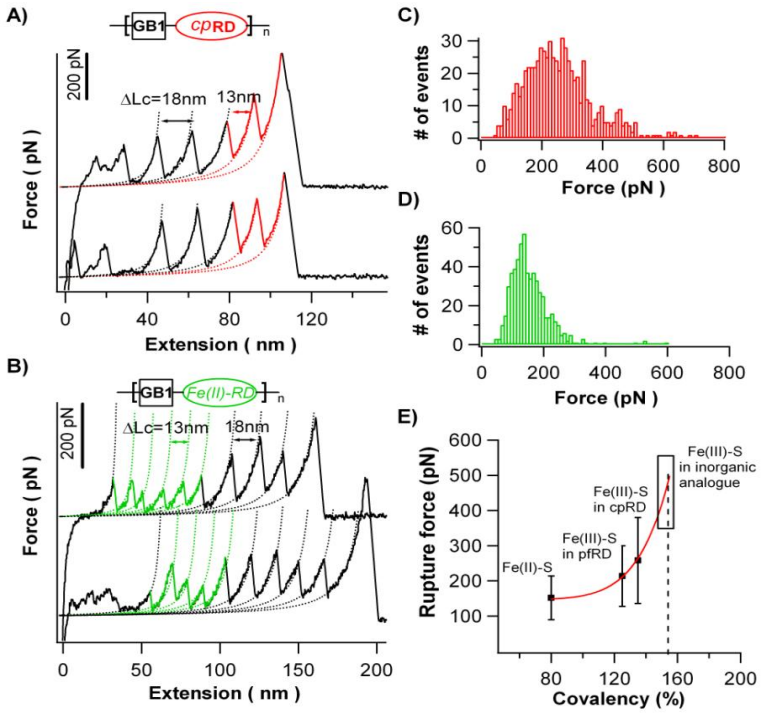

Figure 3.5 The rupture forces of FeS_4 center correlate with its covalency. A) Force-extension curves of (GB1-Fe(III)-*cp*RD)$_n$. The unfolding events of *cp*RD tend to occur after GB1 domains, suggesting that the rupture force of FeS_4 center is higher than that of GB1. B) Force-extension curves of (GB1-Fe(II)-*pf*RD)$_n$. The rupture force of Fe(II)-thiolate bonds is clearly lower than the unfolding force of GB1. C) Histogram of the rupture force of Fe(III)-thiolate bonds in *cp*RD. The average rupture force is 258 ± 122 pN (n=686). D) Histogram of the rupture force of Fe(II)-thiolate bonds in *pf*RD. The average rupture force is 152 ± 62pN (n=579). E) Rupture force correlates with the covalency of the Fe-thiolate bonds.

Furthermore, our work also raised interesting questions on the mechanism of mechanical activation of Fe-thiolate bonds as well as the nature and reactivity of the broken Fe(III)-S bond[197]. Our preliminary results indicated that the broken Fe-S bond can reform within rubredoxin upon relaxation of the rubredoxin chain as evidenced by the ability of rubredoxin to recover its mechanical stability. However, the mechanism of mechanical activation remains unknown. During the mechanical activation, it is possible that the breaking of Fe-thiolate bonds may be a heterolytic process involving competition with a proton, just like in traditional thermodynamic activation processes

of Fe-thiolate bonds[198]. Similar heterolytic processes have been experimentally observed for the mechanical unfolding of disulfide bonds in the presence of reducing agents[21,106] and predicted for the mechanical unfolding of polyethylene glycol[197]. It is also possible that the mechanical breaking of Fe-thiolate bonds is a homolytic process, as suggested for the breaking of C-Si covalent bonds[14,189]. These two different scenarios may lead to different products (sulfhydryl versus free radical) and possibly different reactivity. This important question needs thorough mechanistic investigation by combining experiments with computational chemistry methods. Future work along this direction will be crucial to elucidate the detailed roles of covalent and ionic characters in the mechanical activation (bond rupture) process of metal-thiolate bonds as well as their reactivity after rupture. These studies will likely provide new insights into the nature of this important class of chemical bonds.

3.5 Materials and methods

3.5.1 Protein engineering

The gene of protein chimera Cys-RD-GB1-Cys was constructed in expression vector pQE80L using well-established standard molecular biology techniques[42]. Cys-RD-GB1-Cys was over-expressed in *E. coli* strain DH5α and purified by Co^{2+}-affinity chromatography using TALON resins (Clontech.). The protein was kept in Tris buffer in pH 7.4 at a concentration of ~ 2 mg/mL. Cys-*cp*RD-GB1-Cys was constructed in a similar fashion. Histidine mutants of rubredoxin were generated via standard site-directed mutagenesis methods using the RD gene as the template.

Pseudo-apo-RD (apoRD), which contains four mutations Cys5Lys, Cys8Thr, Cys38Ala and Cys41Thr, was generated via the megaprimer approach using the wt-RD as the template[193].

Since Zn-RD and Fe-RD are co-expressed in *E. coli*[162], we used ion-exchange chromatography to produce pure Fe(III)-RD-GB1 proteins. First, the protein chimera RD-GB1 was concentrated and buffer exchanged into a 10 mM Tris buffer at pH 8.5 using a 9K MWCO pierce concentrator (Thermo Scientific). The Fe and Zn form RD were separated using Mono-Q 5/50GL anion exchange column (GE Healthcare), and eluted using a linear gradient elution (0-300mM NaCl in 10mM and 1mM TCEP at pH 8.5) in AKTA FPLC system (GE Healthcare) at a flow rate of 2 mL/min. The Fe-RD-GB1 was eluted first at around 100 mM NaCl. The purity of Fe-form RD was confirmed by UV-Vis Absorption Spectroscopy based on its characteristic absorption at 494nm. Using the extinction coefficient of 9.22 mM^{-1} cm^{-1} at 494 nm, the concentration of Fe(III)-RD was calculated, and the overall protein concentration was obtained by measuring the absorption at 280 nm[162]. The purify of Fe(III)-RD was estimated to be >90% after 2 to 3 times ion-exchange chromatography purification. The Fe-Form of the rubredoxin mutants were purified using the same method.

3.5.2 Engineering polyproteins for single molecule AFM experiments

The polyprotein gene $(I27\text{-}apoRD)_4$ and $(I27\text{-}pfRD)_4$ were engineered using the well-established stepwise DNA-concatamerization method based on the identity of the sticky ends generated by *BamH*I and *Bgl*II restriction digestion[42]. The polyprotein

$(I27\text{-}apoRD)_4$ and $(I27\text{-}pfRD)_4$ were over-expressed in DH5α and purified by Co^{2+}-affinity chromatography using TALON resins.

Since Zn-RD and Fe-RD are co-expressed in *E. coli*[162], the conventional method for polyprotein construction cannot be used to construct polyproteins of rubredoxin, as it will lead to the production of mixed metal-containing rubredoxin[162]. To overcome this hurdle, we developed a novel chemical coulping method based on maleimde-thiol coupling chemistry to construct polyprotein chimera $(Fe(III)\text{-}RD\text{-}GB1Fe(III)RD\text{-}GB1)_n$, in which rubredoxin exists solely as Fe(III)-RD, as demonstrated in Chapter 2. In a typical reaction, 40 uL 10mM $BM(PEO)_3$ (1, 11-bis (maleimido)triethylene glycol, Thermo Scientific) solution was added to 1ml Fe(III)-RD-GB1 solution (at a concentration of 2 mg/ml in Tris buffer under pH 7.4), which was purified via ion-exchange chromatography to remove Zn-RD, at a molar ratio of 1 to 1. The solution was incubated for crosslinking for two hours, and the protein solution was used directly in AFM experiments.

3.5.3 Single molecule AFM experiments

Single molecule AFM experiments were carried out on a custom-built AFM as described[73,187]. Each individual cantilever was calibrated in solution using the equipartition theorem before each experiment to obtain the spring constant (typically around 40 pN/nm). All experiments were done at room temperature in Tris buffer at pH 7.4 at a pulling speed of 400 nm/s unless otherwise indicated.

For experiments on Fe(II)-RD, after the oxidized $(Fe(III)\text{-}RD\text{-}GB1)_n$ protein was

absorbed onto the glass cover slip, 10 μL of 200 mM dithioreitol (DTT) was added to the solution. After 15 minutes of incubation, the protein was subject to AFM experiments. To ensure the reducing environment, the same amount of DTT solution was added every hour thereafter.

Chapter 4:Single molecule AFM reveals the iron release from rubredoxin through a stochastic mechanism

4.1 Synopsis

After demonstrating that the FeS_4 center in rubredoxin can be mechanically ruptured, the mechanism of this mechanical activation was explored. There are four ferric-thiolate bonds in the FeS_4 center. Thus, how these multiple bonds ruptured over time is a novel and important question.

Multiple metal-ligand bonds in a metal site is a common feature in metalloproteins, where the loss of metal ions from metalloproteins can have significant biological consequences, such as protein unfolding and loss of function. By definition, metal release requires disruption of multiple metal-ligand bonds. Thus, understanding the bond activation mechanism of multiple ferric-thiolate bonds is of great importance. Detailed activation pathways have been difficult to elucidate using classical ensemble techniques. Using single molecule force spectroscopy techniques, we investigated the mechanical activation mechanism of the iron center in rubredoxin. Our results show that the release of the iron center in rubredoxin is stochastic and follows complex and multiple pathways, including cooperative rupture of multiple ferric-thiolate bonds as well as stepwise rupture of ferric-thiolate bonds that leads to the formation of intermediate species. Our results reveal the complexity in the

activation process even within the seemingly simple iron center in rubredoxin, and providing the first unambiguous experimental evidence for the detailed mechanical activation mechanism of a metal center in its native protein environment in aqueous solution.

4.2 Introduction

The mechanical activation of chemical bonds is increasingly recognized as an important means by which chemical reactions can be induced and as a complement to the classical thermochemical, electrochemical and photochemical activation processes[1,14]. The mechanical approach provides new perspectives on the activation of chemical bonds and may reveal mechanistic information that is otherwise difficult to obtain by other methods[106,197,199]. The development of single molecule force spectroscopy techniques has enabled the investigation of the mechanical activation of chemical bonds at the single molecule level. Pioneering studies have provided new insights into some important chemical reactions[14,58,200-202].

Metalloproteins are ubiquitous in nature and play essential roles in a wide range of biological processes. Through the formation of multiple metal-ligand coordination bonds, metal centers are incorporated into metalloproteins to serve as structural and active sites that greatly expand the stability and functionality of proteins[154,156]. The mechanical activation of metal-ligand bonds can lead to the disruption of metal-ligand bonds and the dissociation of metal ions from metalloproteins and may ultimately lead to the unfolding of metalloproteins. Understanding the mechanisms by which

metal-ligand bonds in metalloproteins are disrupted is not only of critical importance for the understanding of the functional properties of such proteins but may also offer new insights into the mechanisms by which metalloproteins fold[203,204].

Spectroscopic signatures of metal-ligand bonds show the activation and dissociation of metal-ligand bonds at the ensemble level[156,205-207]. However, such classical methods cannot directly reveal either the dissociation of an individual metal-ligand bond or the sequence of events leading to the dissociation of multiple metal-ligand bonds. At present, the mechanistic complexity of metal center activation remains largely uncharacterized. Activation and dissociation of many metal centers in metalloproteins have been shown to exhibit first-order kinetics, suggesting that the multiple metal-ligand bonds at these metal centers dissociate concurrently in a cooperative fashion.

Here, we use single molecule force spectroscopy techniques to investigate the mechanical activation process for the dissociation of the $Fe(SCys)_4$ center in rubredoxin. Our results reveal that mechanical release of iron from the active site of this protein is complex and involves several pathways, both cooperative and sequential, that occur in parallel. This discovery that disruption of even the simplest of metal ion binding sites occurs by a stochastic mechanism provides new insight into the functional properties of such sites and the forces that stabilize them.

4.3 Results

4.3.1 Construction of a loop elongation variant of rubredoxin for detection of single-bond rupture events

We still use our model system *P. furiosus* rubredoxin with a $Fe(SCys)_4$ center as to study the mechanical activation mechanism of the metal center. As a simple electron transfer protein with only 53 residues, rubredoxin contains two classic CXXC loops (C5XXC8 and C38XXC41) that bind a ferric ion and form four ferric-thiolate bonds (Fig. 4.1A colored in blue)[159,162,167]. Previously, we showed that these highly covalent ferric-thiolate bonds in rubredoxin can be mechanically activated and ruptured by single molecule AFM.

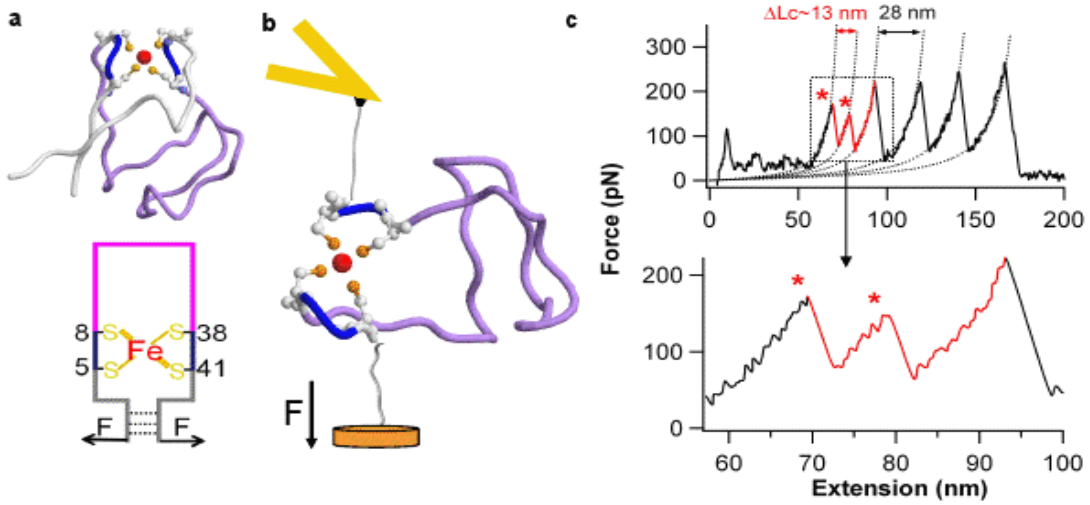

Figure 4.1 Investigation of the mechanical activation of the $Fe(SCys)_4$ center ferric-thiolate bonds in rubredoxin by single molecule AFM. A) 3D structure of wild-type rubredoxin. The two iron chelating loops are colored in blue. Residues 8 to 41 (in pink) are sequestered by the iron-center. The bottom panel shows a simplified schematic of the rubredoxin structure. B) Schematic of the mechanical activation of the $Fe(SCys)_4$ center by AFM. The $Fe(SCys)_4$ center is directly subject to the stretching force after residues 1-5 and 41-53 have been ruptured and extended. C) Typical mechanical activation events of the $Fe(SCys)_4$ center in wt rubredoxin. The

force-extension curve is from the stretching of polyprotein chimera (wt-RD-I27)$_n$. It appears that rupture of the Fe(SCys)$_4$ center occurs in one step, as shown in the enlarged figure of the two rupture events at the bottom, resulting in a single force peak and a length increment of ΔLc ~13 nm.

It showed that the mechanical unfolding of rubredoxin follows a two-step process[201,202]. Residues 1-5, and 41-56 outside the metal center are unraveled first by the stretching force, while the residues sequestered by the metal center (residues 5-41) are shielded from the stretching force. Further stretching leads to the mechanical activation of the Fe(SCys)$_4$ center, rupture of the ferric-thiolate bonds, and complete unfolding of rubredoxin. Mechanical activation of the iron center appears to be a cooperative process, as the rupture of the metal center led to a one-step extension of the protein by ~13 nm, which agrees well with the extension of residues 5-41 (Fig. 4.1C). However, the two chelating loops C5XXC8 and C38XXC41 are very short. If the mechanical activation of Fe(SCys)$_4$ were to occur in a sequential fashion, the rupture of a single ferric-thiolate bond formed by Fe and C5 or C41 would lead to an length increment of 0.72 nm resulting from the extension of the two XX residues. Because of the soft polymer chain, this length gain is even shorter when the bond is ruptured under a relatively low force, and cannot be detected by current single molecule AFM instruments. As a result, the mechanical rupture event observed in the force-extension curve may be dominated by the subsequent rupture event of the additional ferric-thiolate bond, and a sequential mechanical activation process cannot be resolved.

To overcome this technical challenge and gain detailed insight into the mechanical activation mechanism of the iron center in rubredoxin, we designed a loop

insertion variant of rubredoxin (RDβ2), in which a long loop is inserted into the C38XXC41 chelating loop. Insertion of a long loop into the CXXC chelating loop should permit any potential stepwise mechanical activation events to be captured unambiguously.

To construct the loop elongation rubredoxin variant RDβ2 while maintaining the structure and functional properties of wild-type rubredoxin, we chose the second β hairpin of GB1 as the inserted loop[209]. It is well known that the second β hairpin of GB1 is stable and can exist in isolation and that the N, C-termini of this structural element are close to each other (0.4 nm apart)[210,211]. In addition, the C38XXC41 loop is present in a largely unstructured region of the rubredoxin. Thus, the insertion of this β hairpin into C38XXC41 should not affect the structure of rubredoxin or the iron center. Furthermore, this hairpin is comprised of 19 residues and stepwise rupture of the $Fe(SCys)_4$ center would lead to the extension of the hairpin with an increase in length of ~ 6.8 nm (19*0.36nm).

To evaluate the consequences of the β hairpin insertion on the structure of the $Fe(SCys)_4$ center and the overall structure of the protein, we undertook spectroscopic characterizations of the variant protein (RDβ2) (Fig. 4.2C-E). The electronic absorption spectrum of RDβ2 exhibits maxima at 390, 495, and 570 nm resulting from ligand-to-metal charge-transfer transitions that are identical to those of oxidized, wild-type rubredoxin[162,167]. The reduction potential of RDβ2 is similar to that of wild-type RD. These results strongly suggest that the coordination environment of the $Fe(SCys)_4$ center is not affected by the insertion of the β hairpin and that the iron

center retains electron transfer properties similar to the wild-type protein[162,209]. In addition, the far-UV CD spectrum of RDβ2 is similar to that of wt rubredoxin, suggesting that the three-stranded antiparallel β-sheet structure remains largely intact. This conclusion is further supported by $^{1}H^{15}N$-HSQC experiments. In sum, these results provide strong evidence that the structural and functional properties of rubredoxin are minimally perturbed by this loop insertion.

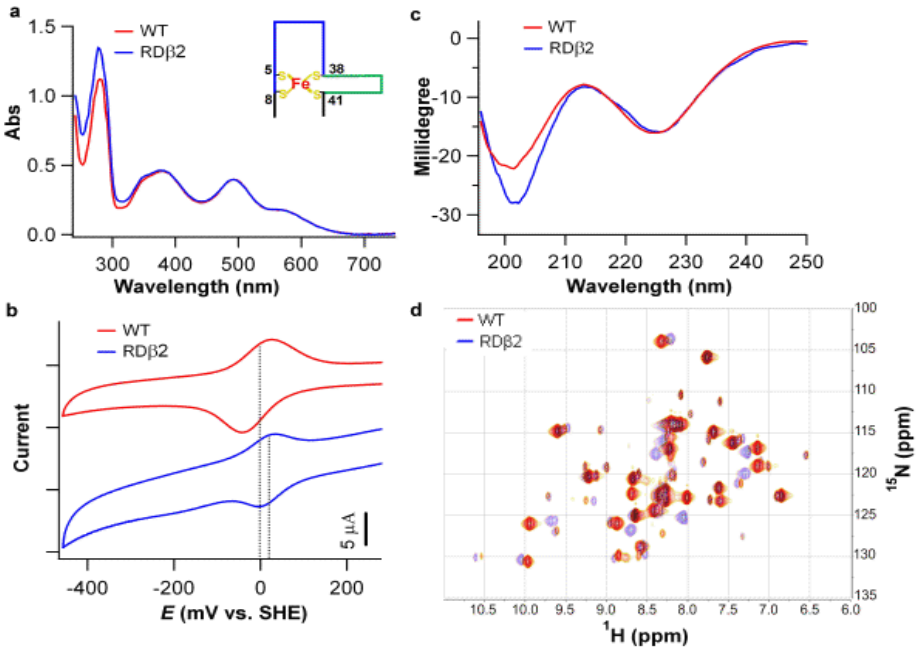

Figure 4.2 The iron center and overall three-dimensional structure are minimally perturbed in the loop insertion variant RD. A) The UV-Vis spectrum of RDβ2 (color in blue) exhibits the characteristic features of the wild-type protein (in red) and the same A_{484}/ A_{380} ratio. The inset shows a simplified schematic of RDβ2. B) CV of wild-type rubredoxin and the RDβ2 variant. The reduction potential of RDβ2 shows slight shift (15 mV) compared with that of wt-RD (-4 mV). C) Far-UV CD spectrum of wild-type rubredoxin and the RDβ2 variant. D) Overlap of the HSQC spectra of RDβ2 and WT-RD (Fe(III)). Signals without affect from the paramagnetic ferric ion are present in the figure. Signals in wt-RD spectrum were also observed in RDβ2 spectrum at the same or slightly shifted (< 0.5 ppm in 1H) positions, strongly suggesting that the backbone structure of RDβ2 is very similar to that of wt-RD.

With this loop elongation variant, we are now able to distinguish whether the

bonds between the ferric ion and the new bi-cysteine binding motif rupture in a cooperative or stepwise fashion upon mechanical activation. If the two ferric-thiolate bonds Fe-Cys38 and Fe-Cys41 rupture cooperatively, the mechanical rupture of the $Fe(SCys)_4$ center should result in a single unfolding step with a ΔLc of ~18.7 nm (6.8+11.9 nm), which corresponds to the extension of the inserted β-hairpin together with the residues sequestered between residues C5 and C41. In the case of a stepwise mechanical activation mechanism, two rupture events should be observed. The first event (ΔLc = 6.8 nm) corresponds to the extension of the inserted β-hairpin, and the second event (ΔLc = 12 nm) corresponds to the extension of residues 5-38. Consequently, both the cooperative and step-wise mechanisms for Fe-ligand bond rupture can be identified unambiguously from unique signatures in the force-extension curves.

4.3.2 Direct observation of multiple pathways for iron release

To monitor the mechanical activation pathways of the $Fe(SCys)_4$ center in rubredoxin directly, we constructed a polyprotein chimera $(RD\beta2\text{-}I27)_n$ for single molecule force spectroscopy measurements. This chimera incorporates the well-characterized I27 domain from the giant muscle protein titin to provide an unambiguous fingerprint for identification of single molecule stretching events[130,212]. The mechanical unfolding of the I27 domain is characterized by ΔLc of 28 nm and an unfolding force of ~200 pN at a pulling speed of 400 nm/s.

Representative force-extension curves for this polyprotein chimera $(RD\beta2\text{-}I27)_n$

are shown in Fig. 4.3A. The unfolding events with ΔLc of 28 nm can be assigned readily to the unfolding of the I27 domains while the other unfolding events can be assigned to the unfolding of RDβ2 domains. The complete unfolding of RDβ2 leads to a contour length increment (ΔLc) of ~19 nm. In the force-extension curves, the unfolding events of RDβ2 display multiple appearances: some unfolding events display a single unfolding step with a ΔLc of ~19 nm while others display two unfolding steps with the sum of ΔLc1 and ΔLc2 being 19 nm. In most of the two-step unfolding cases, the first step leads to a ΔLc of ~7 nm followed by a second step of a ΔLc of ~12 nm. In rare cases, the sequence of the two steps reverses so that the first peak exhibits a ΔLc of ~12 nm that is followed by an unfolding step of with ΔLc of ~7 nm. The histogram of ΔLc values (Fig. 4.3B) clearly exhibits three dominant peaks with an average ΔLc1 of 7.2 ± 1.6 nm (n = 121), ΔLc2 of 12.3 ± 1.6 nm (n = 121), and ΔLc3 of 19.1 ± 2.0 nm (n = 547). The forth peak, with an average of 28.2 ± 1.0 nm (n = 327), corresponds to the ΔLc of the unfolding of I27 domains.

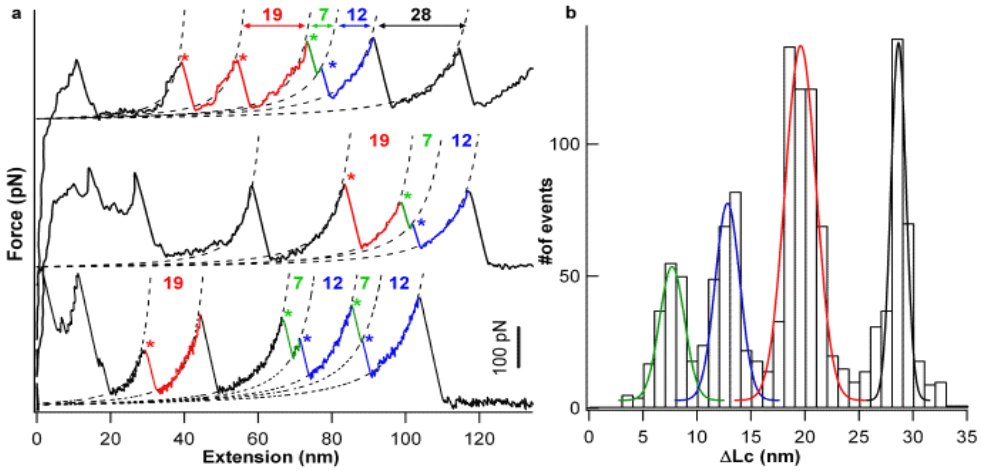

Figure 4.3 The mechanical unfolding experiments on rubredoxin loop variant RDβ2 demonstrate a complex rupture pattern of the $Fe(SCys)_4$ center. A) Typical

force-extension curves observed during stretching the polyprotein $(RD\beta2\text{-}I27)_n$. In addition to the unfolding force peaks from the I27 domain (ΔLc ~28nm), three distinct rupture force peaks resulting from unfolding of the RDβ2 are observed. The dominant component (~80%) is a single force peak (red star) with ΔLc of 19 nm (colored in red). The remaining events (~20%) involve two sequential force peaks with a ΔLc of 7 nm and 12 nm, respectively (colored in green and blue). B) The ΔLc histogram on the polyprotein $(RD\beta2\text{-}I27)_n$ clearly shows the four different types of peak: the average ΔLc values for these four components are 7.2 ± 1.6 nm (n = 121); 12.3 ± 1.6 nm (n = 121); 19.1 ± 2.0 nm (n = 547) and 28.2 ± 1.0 nm (n = 327).

The observation of multiple unfolding patterns for RDβ2 demonstrates clearly that the mechanical rupture of the $Fe(SCys)_4$ center in rubredoxin occurs through multiple pathways, including both single step rupture as well as stepwise rupture. This finding establishes the complexity of the mechanical activation of the $Fe(SCys)_4$ center in rubredoxin and demonstrates the rich information about the mechanical activation mechanism of the metal center that cannot be discerned from ensemble studies.

4.3.3 Cooperative rupture of the $Fe(SCys)_4$ center is dominant

From typical force-extension traces of the polyprotein chimera $(RD\beta2\text{-}I27)_n$, we observed that a large population of RDβ2 (~81%) unfolds in a single step that results in unfolding events of ΔLc of ~19 nm (Fig. 4.4A&B). The average of the measured ΔLc (18.6 nm (n = 547)) is in excellent agreement with the expected contour length increment that should result from the extension of residues sequestered by the metal center (residues 5 to 38) as well as the inserted β hairpin (11.6 + 6.9 = 18.7 nm), suggesting that the two ferric-thiolate bonds from the C38XXC41 chelating loop must have been ruptured concurrently by the applied stretching force during the mechanical

activation (Fig. 4.4A). This result is strong evidence for a cooperative mechanical activation process for the $Fe(SCys)_4$ center, despite the insertion of the 19 residue long β hairpin. The average force required to rupture the two ferric-thiolate bonds simultaneously is 194 ± 92 pN (n = 547). To confirm that the single step unfolding is not an artifact of limited time resolution, we carried out the single molecule AFM experiments at higher data sampling rate (Fig. 4.4B, bottom trace). It is clear that with a sampling frequency of 50 kHz, the unfolding step occurs in a single step, suggesting that the mechanical activation of the $Fe(SCys)_4$ center occurs simultaneously within a time window of 20 μs. This result highlights the cooperativity of the $Fe(SCys)_4$ center in rubredoxin as observed previously in ensemble studies[171,206].

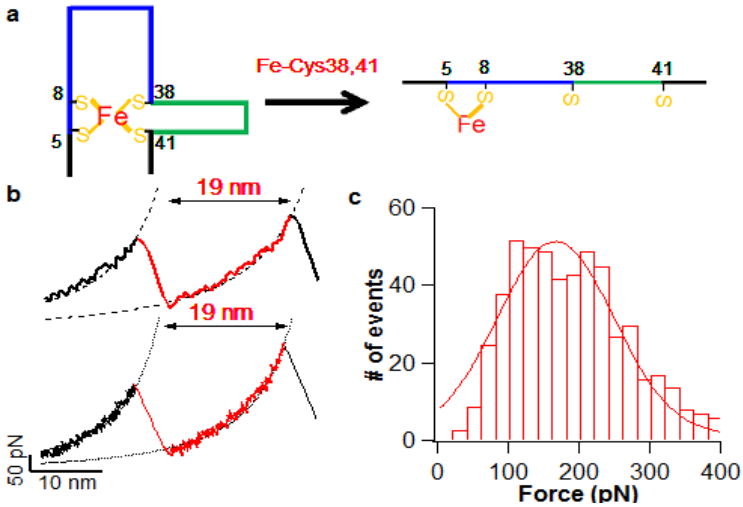

Figure 4.4 The cooperative mechanism for mechanical disruption of the rubredoxin $Fe(SCys)_4$ center. A) A schematic of the cooperative activation pathway of the $Fe(SCys)_4$ center in RDβ2. Upon stretching, two ferric-thiolate bonds (Fe-S_{Cys38} and Fe-S_{Cys41}) ruptured concurrently, leading to a single force peak. B) Detailed view of the single step rupture event of the $Fe(SCys)_4$ center at two sampling rates. The top trace, 1 kHz; bottom trace, 50 kHz. It is clear that the rupture event with ΔLc ~19 nm is a single step. C) The rupture force histogram for the cooperative rupture of the $Fe(SCys)_4$ center. The average rupture force is 194 ± 92 pN (n = 547).

4.3.4 Sequential mechanism for iron release and measurement of individual ferric-thiolate bond strength

In addition to the dominant cooperative mechanical activation, we also observed that ~20% of RDβ2 is ruptured in a stepwise fashion by the applied stretching force. As shown in Figure 4.5A&B, the mechanical rupture of RDβ2 domains resulted in pairs of unfolding force peaks. The first peak (ΔLc ~7 nm (colored in green, 7.2 ± 1.6 nm)) occurs at a slightly elevated force relative to the second peak (ΔLc ~12 nm (colored in blue, 12.3 ± 1.6 nm)) (Fig. 4.6A). The ΔLc of the first peak agrees well with that for the extension of the inserted loop initiated by the rupture of the Fe-Cys41 bond (0.36*20-0.3 = 6.9 nm) while the value for the second peak agrees well with the value for the complete rupture of the partially disrupted $Fe(SCys)_3$ center (34*0.36-0.6 = 11.6 nm). These results strongly suggest that the RDβ2 domains rupture in a stepwise, sequential fashion such that the Fe-Cys41 and Fe-Cys38 bonds rupture sequentially.

Moreover, from the pair of force peaks resulting from these sequential rupture events, we can attribute the rupture force value of each step to the mechanical strength of single ferric-thiolate bonds (Fe-Cys41 or Fe-Cys38). The average rupture force of the first peak (176 ± 62 pN, n = 99, Fig. 4.5C) corresponds to the mechanical strength of the Fe-Cys41 bond while the rupture force of the second peak (127 ± 48 pN, n = 99, Fig. 4.5D) corresponds to the strength of the Fe-Cys38 bond. From the force-extension curve as well as the force histogram, it is evident that the mechanical strength of the Fe-Cys41 bond is greater than that of Fe-Cys38 bond. It is consistent

with the bond lengths of the two distinct ferric-thiolate bonds insofar as the shorter bond (Fe-Cys41) is stronger than the longer bond (Fe-Cys38). It is interesting to note that in this stepwise mechanical activation pathway the mechanically stronger bond Fe-C41 ruptures first following the mechanical rupture of the weaker bond Fe-C38, suggesting that the Fe-C41 bond protects the mechanical integrity of the Fe-Cys38 bond by means of a gating mechanism. A similar gating mechanism has been reported for both designed proteins and naturally occurring proteins, such as titin kinase in muscles[213,214].

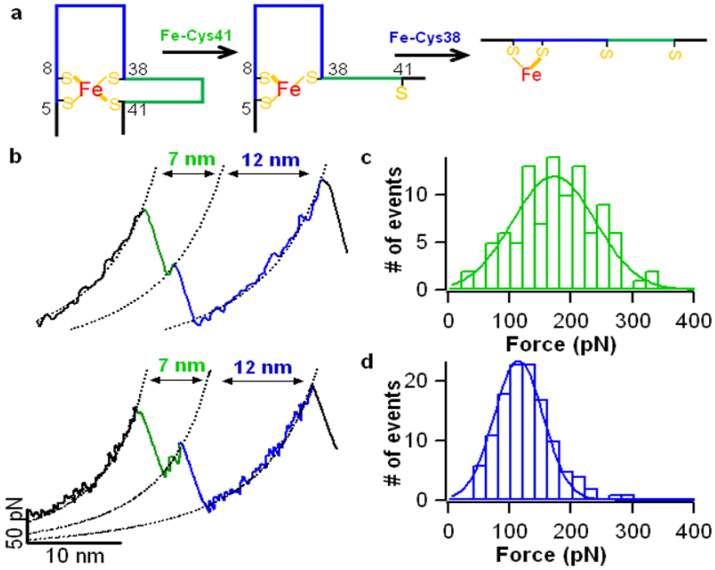

Figure 4.5 The sequential mechanism for disruption of the rubredoxin $Fe(SCys)_4$ center. A) A schematic of a sequential activation pathway of the $Fe(SCys)_4$ center in RD where the two ferric-thiolate bonds (Fe-Cys38 and Fe-Cys41) rupture sequentially. The Fe-Cys41 bond ruptures first, leading to an intermediate $Fe(SCys)_3$ and the extension of the inserted β hairpin (color in green). Then the rupture of the Fe-Cys38 bond leads to the complete unfolding of rubredoxin and extension of the sequestered residues (color in blue). B) A detailed view of the sequential activation of the $Fe(SCys)_4$ center. C) The rupture force histogram of the ferric thiolate bond Fe-Cys41. The average force is 176 ± 63 pN (n = 99). D) The rupture force histogram of the ferric-thiolate bond Fe-Cys38. The average force is 127 ± 48 pN (n = 99).

Remarkably, we also detected two-step rupture events in which the order of

Fe-Cys bond disruption was reversed so that the rupture event with a ΔLc of ~12 nm occurred first, followed by the rupture event with ΔLc of 7 nm. This sequence of rupture events suggests that the rupture of the $Fe(SCys)_4$ center can also be initiated from the side of the C5XXC8 chelating loop, indicating that the mechanical rupture of the $Fe(SCys)_4$ center in rubredoxin is stochastic in nature and can be initiated from either CXXC chelating loop.

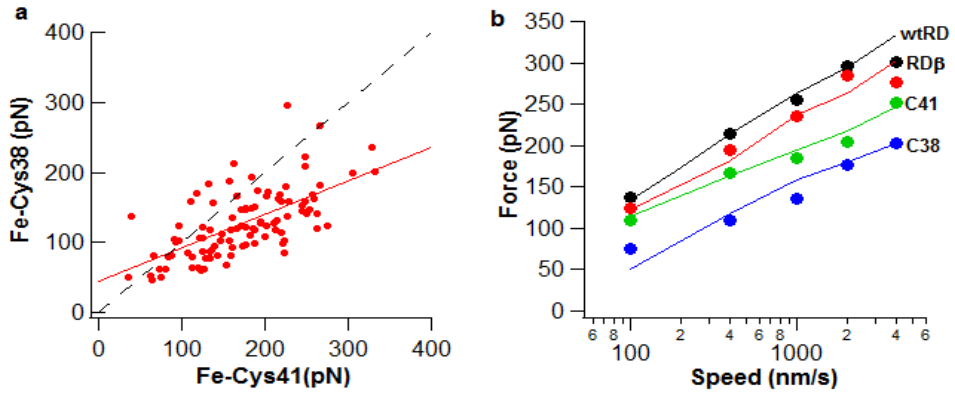

Figure 4.6 Strength and kinetics of the ferric-thiolate bond. A) The relationship between the bond strength of the Fe-Cys41 and Fe-Cys38 bonds in the same pair of the sequential rupture process. The dashed line is a straight line with a slope of 1. The solid line represents a linear regression of the data with a slope of 0.48. These results clearly demonstrate that the strength of Fe-Cys41 bond is greater than that the Fe-Cys38 bond. B) The cooperative rupture process (red), rupture of the Fe-Cys41 bond (green) and rupture of the Fe-Cys38 bond (blue) are shown. Results obtained for the wild-type protein (black) are shown for comparison. The lines represent fits to the data obtained by Monte Carlo simulation.

Having observed multiple mechanical activation pathways for disruption of the rubredoxin active site, we carried out pulling experiments as a function of pulling velocity to characterize bond rupture kinetics (Fig. 4.6B). As expected, the rupture forces were observed to increase with the increase of the pulling velocity. Using well-established Monte Carlo simulation protocols, we estimated the spontaneous dissociation rate and the distance between the bound state and the mechanical

transition state. For the cooperative rupture pathway, the rate constant for spontaneous dissociation at zero force is 0.17 s^{-1} and the distance between the bound state and the mechanical rupture transition state, Δx_u, is 0.11 nm. These values are quite similar to those observed for wt-RD (0.15 s^{-1} and 0.11 nm)[201]. In contrast, the kinetics for single bond rupture events is significantly different. The Δx_u is 0.13 nm for Fe-Cys41 and Fe-Cys38 with α_0 of 0.21 s^{-1} and 1.4 s^{-1}, respectively. The difference in the kinetics data of the two distinct mechanical activation pathways also demonstrates that these two rupture mechanisms are different in nature.

4.3.5 Stochastic iron release mechanism also observed for RD1β

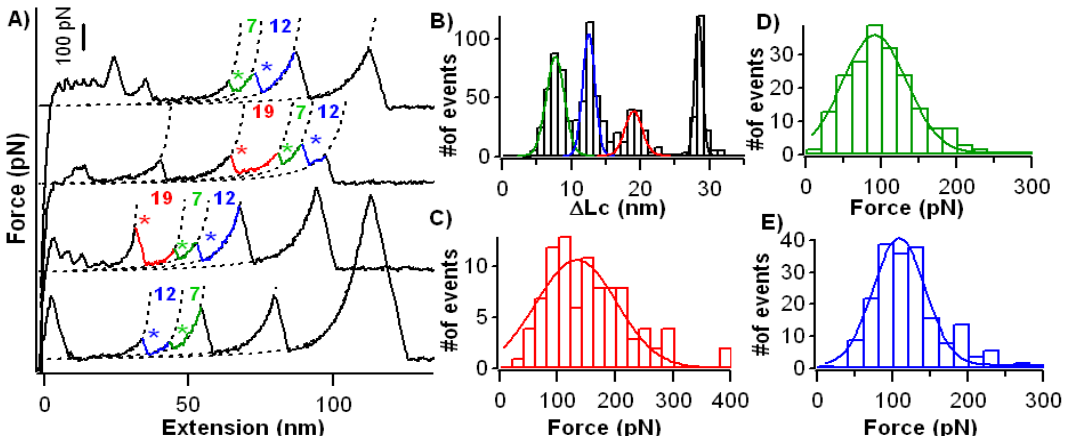

Figure 4.7 Mechanical rupture of $Fe(CysS)_4$ center in RD1β. A) Typical force-extension of (RD1β-I27) shows two different rupture scenarios of FeS_4. B) Histogram of contour length increment from stretching (RD1β-I27) polyprotein. C) The rupture force histogram for the cooperative rupture of the $Fe(SCys)_4$ center. The average rupture force is 154 ± 72 pN (n = 93). D) Histogram of rupture force of Fe-Cys5 bond, the average force is 107 ± 60 (n=202). E) Histogram of rupture force of Fe-Cys8 bond, the average force is 129 ± 64 pN (n=202).

To investigate the iron release mechanism on the other C5XXC8 loop, we constructed a rubredoxin variant (RD1β) in which the longer loop is inserted between

Cys5 and Cys8. Similar AFM experiments were carried out on the $(RD1\beta–I27)_n$ polyprotein. The AFM results are highly similar as that of RD2β. Stretching the polyprotein results in two types of rupture scenario of FeS_4 center: single force peak with ΔLc of 18.5 ± 1.8 nm (n = 93) as the cooperative rupture of the metal center, and force peak with an intermediate step (ΔLc1=7.2 ± 1.8 nm; ΔLc2=12.1 ± 1.2; n=202) corresponding to a sequential rupture of the two individual ferric-thiolate bonds (Fig 4.7). Consequently, a stochastic iron release mechanism is also observed for the first CXXC loop. However, the ratio of cooperative rupture even in RD1β is much less than that of RD2β (~30% versus ~80%). Additionally, in the sequential rupture scenario, force peak with ΔLc of ~7 nm shows first, followed by force peak with ΔLc of ~12 nm is dominant compared with the opposite pathway (n=202 versus n=8). It indicates that the rupture of the FeS_4 center starts from C5XXC8 in this variant which different than that of RD2b. It suggests that the inserted loop may destabilize the CXXC metal binding loop and bias the rupture pathway to the inserted loop. Nevertheless, both pathways are present in two variants. Similarly, the cooperative rupture force and individual ferric-thiolate bond strength are measure: 154 ± 72 pN, 107 ± 60 pN (Fe-Cys5) and 129 ± 64 pN (Fe-Cys8), respectively.

4.4 Discussion

4.4.1 Stochastic nature of iron release from the $Fe(SCys)_4$ center

Metal centers with multiple metal-ligand bonds are a common feature in metalloprotein.[154,156,167]. The loss of metal ions in metalloproteins can have significant

biological consequences and understanding the activation mechanism of such metal centers in metalloproteins is thus of great importance[156,204,215,216]. The release of metal ions from metalloproteins involves the dissociation of multiple metal-ligand bonds and detailed activation pathways have been difficult to elucidate with classical ensemble techniques because such methods are typically unable to detect short-lived and/or less populated intermediate species in which just one or two metal-ligand bonds have been ruptured. The single molecule force spectroscopy techniques used here provide the first unambiguous experimental evidence that the mechanical activation of even a simple iron center such as that in the iron-sulfur protein rubredoxin is stochastic and follows multiple, complex pathways that include cooperative rupture of multiple ferric-thiolate bonds as well as stepwise rupture of ferric-thiolate bonds that leads to the formation of intermediate species.

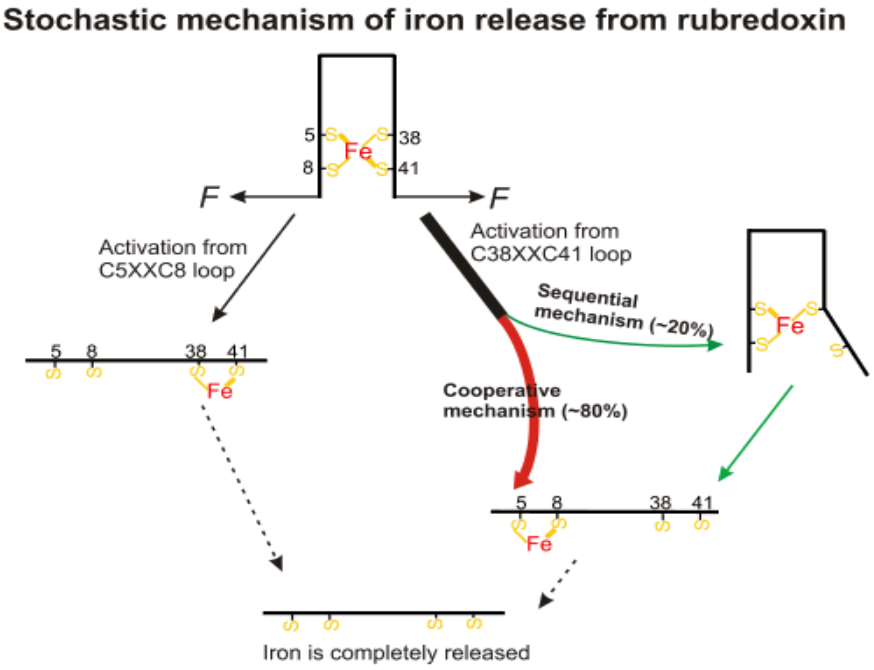

Figure 4.8 Stochastic mechanical disruption of the rubredoxin $Fe(SCys)_4$ center

derived from single molecule force spectroscopy experiments. The rupture of the $Fe(SCys)_4$ can be initiated at either CXXC loop. For the pathway initiated at the C38XXC41loop, 80% of rupture events occur in a single step in which the ferric-thiolate bonds behave cooperatively and rupture of Fe-Cys bonds is concurrent. In addition, 20% of Fe-Cys bond rupture events occur in a stepwise fashion in which the Fe-Cys41 bond ruptures first followed by the rupture of the Fe-Cys38 bond. For the pathway initiated through the C5XXC8 loop, it is very likely both cooperative and sequential rupture processes occur in a manner similar to that observed for the C38XXC41 loop.

Based on these results, we propose the mechanical activation mechanism for disruption of the $Fe(SCys)_4$ center in rubredoxin that is shown in Fig. 4.7. Mechanical activation can initiate from either side of the metal center and involves both cooperative and sequential rupture of the metal center. It mainly based on result from RD2β due to the well-maintained structure. On comparison, the structure of RD1β is distorted significantly. The NMR spectrum shows a large variation compared with that of WT-RD. Nevertheless, the stochastic iron release mechanism from rubredoxin is still valid. It is important to note that the pathway for mechanical activation of wt-rubredoxin disruption is likely to be more complex because it should involve a combination of rupture events from both CXXC chelation loops. For example, it is possible that the mechanical rupture of the wt rubredoxin starts from the breaking of the Fe-Cys5 bond followed by the cooperative rupture of the C38XXC41 chelation loop.

4.4.2 Cooperative versus sequential bond rupture

With the length increment as an unambiguous fingerprint, we identified the co-existence of the cooperative and sequential mechanical activation pathways for the

$Fe(SCys)_4$ center in rubredoxin. The observation of cooperative bond activation of multiple ferric-thiolate bonds even in the presence of an inserted β hairpin clearly highlights the cooperative character of the $Fe(SCys)_4$ center disruption, a feature that has been proposed on the basis of ensemble measurements[171].

On the other hand, the discovery of the stepwise activation pathway sheds new light on the $Fe(SCys)_4$ center in rubredoxin. Direct detection of partially or completely unfolded metalloprotein with metal ion still bound to the protein could in principle be identified by NMR spectroscopy[216-218]. However, metal centers with a few original metal-ligand bonds lost should occur as an intermediates during the disruption of metal centers in metalloproteins, has rarely been observed. The stepwise activation pathway demonstrated in the present study clearly indicates that the intermediate states such as $Fe(SCys)_3$ are kinetically stable and can exist in aqueous solutions of rubredoxin. The formation of $Fe(SCys)_3$ can result from cleavage of the Fe-Cys41 bond or the Fe-Cys5 bond. To the best of our knowledge, this is the first experimental evidence that an intermediate iron center can occur during the release of iron from a metalloprotein. This finding implies that the $Fe(SCys)_4$ center in rubredoxin can be more dynamic than previously believed. The existence of a stable metal center intermediate also supports the hypothesis that iron priming is an important step in the folding of rubredoxin and that incorporation of iron into rubredoxin involves the sequential formation of ferric-thiolate bonds[173,194].

4.4.3 Why is iron release stochastic?

Although rubredoxin is the simplest iron-sulfur protein, the mechanical activation of $Fe(SCys)_4$ center disruption exhibits a complex mechanism. Understanding how a stretching force activates the $Fe(SCys)_4$ center in rubredoxin provides fundamental insight that will be essential to understanding the mechanism of iron release from more complex iron-sulfur proteins. The $Fe(SCys)_4$ center in rubredoxin is a five-atom, four-bond system in which the ferric ion is at the center of a pseudo-tetrahedron and is bonded to four sulfur atoms provided by the four chelating cysteinyl residues. Following the unraveling and extension of residues 1-5 and 41-56, the $Fe(SCys)_4$ center of rubredoxin is subject to the mechanical stretching force directly through residues C5 and C41. To activate the $Fe(SCys)_4$ center mechanically, at least two ferric-thiolate bonds from the same side of the CXXC chelation loop must be ruptured.

Compared with classic chemical or thermal activation processes that impose a global perturbation on the metal center, the mechanical activation process stretches the metal center along a well-defined reaction coordinate that is set by the applied force. Upon stretching, the two outer ferric-thiolate bonds (Cys5 and Cys41) are likely to be subject to greater force at the beginning of the process than are the two inner bonds formed by Cys8 and Cys38. Thus, the greater mechanistic resolution provided by single molecule atomic force spectroscopy has allowed us to observe stepwise mechanisms of iron release from the active site of rubredoxin that could not be

detected by conventional methods. Nevertheless, the dominant pathway for iron release is the cooperative, simultaneous rupture of two thiolate bonds, an observation that highlights the cooperative nature of the iron center in rubredoxin. We interpret this finding as indicating that the stretching force propagates to the entire metal center through the bonding network as predicted for the propagation of the applied force through an entire protein upon stretching[219,220]. We propose that propagation of the stretching force through the active site in this manner is responsible for the cooperative, pairwise disruption of ferric-thiolate bonds. The structural origin of this cooperativity between the ferric-thiolate bonds in rubredoxin remains to be elucidated. Combined use of molecular dynamics simulations with quantum mechanical calculation of the iron center provides one plausible means of addressing this issue[60,197].

In addition to discovering and partially characterizing the mechanistic pathways for iron release from rubredoxin, this study provides a quantitative characterization of the mechanical strength of individual ferric-thiolate bonds in rubredoxin and demonstrates that these nominally identical metal-ligand bonds are energetically inequivalent. Although X-ray crystallographic structure determination has established that the Fe-C41 bond is shorter than the Fe-C38 bond[211], the physical properties of these minimally different chemical bonds have been challenging to discern. To the best of our knowledge, no comparison of the mechanical strength of nearly identical chemical bonds involved in metal ion binding at the active site of a metalloprotein or in a biomimetic metal cluster has been reported.[1,14]. Our characterization of loop

insertionvariant of rubredoxin has determined that the bond strengths of Fe-Cys41 and Fe-Cys38 in this protein are ~170 pN and ~130 pN, respectively, a result that agrees well with our previous indirect measurements of rubredoxin variants in which the coordinating cysteinyl residues were replaced with histidinyl residues[192,201]. Moreover, this difference in bond strength agrees well with the bond lengths of the ferric-thiolate bonds noted above, supporting the long established correlation of longer bond length with lower bond strength. We expect that this strategy can be applied to other metalloproteins to resolve small energetic differences of chemically identical metal-ligand bonds of inequivalent length.

4.5 Materials and methods

4.5.1 Protein engineering

The genes encoding the rubredoxin variant with a β-hairpin fragment inserted in the C38XXC41 metal ion binding loop were constructed as follows. First, an oligodeoxyribonucleotide encoding the second β-hairpin of GB1 (15 aa EWTYDDATKTFIVTE plus two 2 amino acid LG at each end that result from the restriction site) with additional *Ava*I restriction site at the N and C termini was generated by standard PCR (Polymerase Chain Reaction) methods with the GB1 gene as the template. The resulting gene was digested with the enzyme *Ava*I to produce the insert for subsequent ligation into the rubredoxin expression vector pQE80L, which carries an N-terminal His_6-purification tag. Standard site-directed mutagenesis methods were used to introduce an *Ava*I restriction site (CTC GGG) in rubredoxin to

replace the codons CCC ATC that encode residues Pro39 and Ile40 (the residues between Cys38 and Cys41) and the product was digested. The oligoDNA encoding the β-hairpin fragment was then ligated into the rubredoxin gene between the Cys38 and Cys41 through the *Ava*I site. This construct introduces 19 residues (15 plus 4 from the two *Ava*I sites) in the bi-cysteine loop to produce the rubredoxin loop extension variant with the sequence:

AKWVCKICGYIYDEDAGDPDNGISPGTKFEELPDDWVC*LGEWTYDDATK TFIVTELG*CGAPKSEFEKLED (italics represent the inserted peptide loop).

The gene encoding RD1β is constructed in the same fashion with the sequence as follow:

AKWVC*LGEWTYDDATKTFIVTELG*CGYIYDEDAGDPDNGISPGTKFEELP DDWVCPICGAPKSEFEKLED

For the expression of rubredoxin with an I27 domain as a fingerprint region for single molecule AFM studies, the gene encoding Cys-wt-RD-I27-Cys was used as the template for construction of the loop insertion variant using methods similar to those described above. In addition, the template possesses an N and C-terminal cysteinyl residue to facilitate the further construction of polyproteins by means of maleimide-thiol chemistry[192,201]. The DNA sequences were confirmed by direct DNA sequencing at the NAPS of University of British Columbia.

All proteins used in this work were over-expressed in *Escherichia coli* strain DH5α (Qiagen, Valencia, California, USA) and purified by Co^{2+}-affinity chromatography using TALON His-Tag purification resins (Clontech, Mountain View,

California, USA). The protein solution was exchanged into Tris buffer (pH 8.5, 10 mM) by centrifugal ultrafiltration (Amicon 3K MWCO filter, Millipore, Billerica, MA, USA). Zn-substituted rubredoxin produced during expression of both RDβ2 and the RDβ2-I27 chimera was removed by ion-exchange chromatography using Mono Q 5/50 GL column (GE Healthcare Bioscience, Pittsburgh, PA, USA). Polyproteins $(RD\beta\text{-}I27)_n$, were prepared by a maleimide-thiol coupling reaction in which the monomer chimera protein was reacted with $BM(PEO)_3$ (1, 8-bis (maleimido)triethylene glycol, Molecular Biosciences, Boulder, CO, USA) as described previously[192].

4.5.2 Single molecule AFM

Single molecule AFM experiments were performed with a custom-built AFM as reported[20]. Prior to each experiment, the spring constant (~40 pN/nm) of each MLCT Si_3N_4 cantilever (Bruker, Camrillo, CA) was calibrated in solution using the equipartition theorem. In a typical experiment, the polyprotein $(RD\beta\text{-}I27)_n$ (2 μL, 2 mg/mL) was deposited onto a clean glass coverslip covered by Tris buffer (~50 μL, pH 7.4, 100 mM). The protein was allowed to absorb onto the coverslip for ~5 minutes before the AFM experiment. The experiments were carried out at a pulling speed of 400nm/s unless otherwise indicated.

The mechanical rupture of single ferric-thiolate bonds and the cooperative rupture of the bonds formed by the coordinated ferric ion and the C41XXC38 bi-cysteine ligand pair can both be modeled as a two-state dissociation process as a

function of the force-dependent rate constants:

$$\alpha_{(F)} = \alpha_0 \exp(F\Delta x_u / k_B T) \quad (4.1)$$

The $\alpha_{(F)}$ is the rate constant for dissociation under a stretching force F, α_0 is the rate constant for spontaneous dissociation in the absence of force, Δx_u is the distance between the bound and transition states, T is the absolute temperature and k_B is the Boltzmann constant. The rate constant for dissociation at zero force, α_0, and Δx_u were estimated by Monte Carlo simulations as described previously[43].

4.5.3 Nucleic magnetic resonance experiments

$^{15}N/^{1}H$ heteronuclear single-quantum coherence (HSQC) spectra were recorded on a Bruker Avance-500 FT NMR spectrometer operating at a 1H frequency of 500 MHz using a standard pulse sequence. Spectra were zero-filled to give a final matrix of 2048 x 256 data points and apodized with a 90° shifted sine-bell window function in both dimensions. ^{1}H and ^{15}N chemical shifts were calibrated against the ^{1}H shift of sodium 2,2-dimethyl-2-silapentane-5-sulfonate.

Chapter 5: Mechanical anisotropy of the FeS_4 center in rubredoxin

5.1 Synopsis

Mechanical anisotropy is an important feature of mechanical force acting on three dimensional objects. Because force is a vector, the effect of mechanical perturbation on a 3D structure can vary if the force is applied in different directions. Here, we investigate these effects on a very simple FeS_4 center in rubredoxin, finding that this tiny structure is still anisotropic under force at the single molecule level.

Considering the small scale of proteins, selection and assurance of the direction of applied force is most challenging. The FeS_4 center in rubredoxin is a simple structure consisting of only four ferric-thiolate bonds. The mechanical force applied on the metal center will be ultimately transmitted through these four ferric-thiolate bonds. Consequently, these four bonds are naturally selected as the direction of applied force. In the experiments, all possible combinations involving force applied to the two ferric-thiolate bonds were selected as the stretching sites and their corresponding stability were measured. The results demonstrated that even this molecular structure still possesses mechanical anisotropy. In addition, the characteristic rubredoxin unfolding results validate the direction of the applied force due to the unique contour length increment associated with how the FeS_4 center

ruptured.

Besides the mechanical anisotropy of the metal center, we also observed a stochastic rupture mechanism of the FeS_4 center in the wild type rubredoxin when the protein is stretched from the Fe-Cys5 and Fe-Cys38 direction. The result agrees well with that of the previous chapter in this work, which utilized a loop mutant. Additionally, this work also demonstrates that the pulling direction can affect the rupture mechanism.

5.2 Introduction

Mechanical force is ubiquitous in nature and participates in every aspect of life[34,221]. An important feature of force is that each applied force is associated with a defined direction or vector. Consequently, the effect of mechanical perturbation on a three-dimensional object can be varied if the force is applied in different directions. This so-called mechanical anisotropy has been found for many macroscopic objects, which arises from their anisotropic structure[222]. Recently, it has been found the stability of a single protein molecule under force can be varied using single molecule force spectroscopy[85,99-101,103,223]. However, a single protein molecule is a macromolecule containing hundreds of chemical bonds and a wide range of interactions. Thus, it is quite interesting and largely unknown as to the scale at which the mechanical anisotropy property can be observed.

Here, we investigate the mechanical anisotropy property on an extremely small object: a natural FeS_4 metal center which consists of only five atoms and four

ferric-thiolate bonds in the small metalloprotein rubredoxin[162,163]. The force applied on the FeS_4 center with a well-defined direction is achieved by stretching the single rubredoxin molecule from different sites. Because the metal center is covalently linked within the protein structure, the force will propagate through the peptide chain and ultimately be applied on the metal center in a define direction from certain ferric-thiolate bonds. Moreover, protein unfolding beginning from different sites shows unique unfolding force peak patterns with different contour length increments. Because how the metal center ruptures is directly related to the pathway of rubredoxin unfolding, the direction of force applied on the FeS_4 center is guaranteed. Consequently, the stability of FeS_4 center under mechanical force from various directions can be unambiguously measured and a mechanical anisotropy is found on such a small metal center. In addition, different rupture mechanisms are observed when the metal center is stretched in various directions.

We continue to use *P. furiosus* rubredoxin as the model system here. It is a very simple three-dimensional object with four ferric-thiolate bonds (Fig. 5.1A, abbreviated as Fe-Cys5; Fe-Cys8; Fe-Cys38 and Fe-Cys41bond). The FeS_4 center divides rubredoxin into two parts: Residues 1-5 and 42-53 are outside the center, while residues 5-41 are trapped inside the center (Fig. 5.1B). Combined with single molecule AFM and protein engineering, we stretched the protein from different sites to identify the mechanical anisotropy of the metal center.

5.3 Results

5.3.1 Design principle to probe mechanical anisotropy of FeS_4 center

The first step to explore the mechanical anisotropy is to choose the direction in which the force should be applied on the FeS_4 center. When the rubredoxin is stretched, the force will be propagated through the peptide chain and then be applied to the four ferric-thiolate bonds. Thus, the mechanical behavior of these bonds under force ultimately determines the stability of the metal center. Consequently, the selection of different combinations of two ferric-thiolate bonds as stretching directions can fully explore the anisotropic mechanical response of the FeS_4 center. Based on the structure of FeS_4 center in the protein, four different combinations are chosen: Fe-Cys5 and Fe-Cys8 bonds, Fe-Cys8 and Fe-Cys41 bonds, Fe-Cys8 and Fe-Cys38 bonds and Fe-Cys5 and Fe-Cys38 bonds.

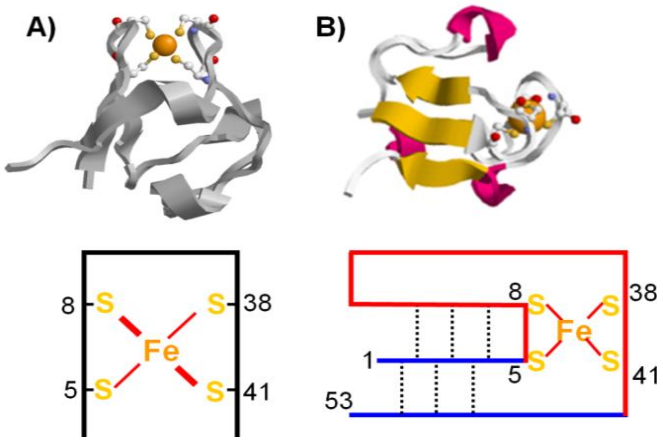

Figure 5.1 Schematic of the rubredoxin structure. A) The 3D structure of rubredoxin showing a FeS_4 center in which a ferric ion is coordinated by four sulfur atoms from cysteine residues. The bottom is a simple schematic showing the metal site in the protein with four ferric-thiolate bond. B) The side-view of rubredoxin. The bottom is a simple schematic showing the overall structure of rubredoxin.

To fulfill this goal, we design four different bi-cysteine rubredoxin variants. Single rubredoxin molecules are then chemically linked through the two engineered cysteines using a maleimide-thiol coupling method[98,192]. It forms a polyprotein molecule $(RD_{variant})_n$ containing identical repeats of rubredoxin which is a typical strategy for AFM studies (Fig. 5.2A). As a result, the two introduced cysteine residues serve as anchoring sites from which the protein will be stretched. Beginning from these two sites, the force will propagate along the peptide chain and ultimately be functioned onto the metal center. It is noted that the ferric-thiolate bond closest to the cysteine in the protein sequence will be stretched initially and thus determines the direction of applied force on the FeS_4 center. By carefully selecting two cysteines as stretching sites for protein unfolding, desired direction of force applied on the metal center through certain ferric-thiolate bonds is achieved. Four rubredoxin variants are designed as follows: RD1,49; RD15,49; RD15,35 and RD1,35 (the number indicates which residue is substituted by cysteine) and the corresponded poly-rubredoxin $(RD_{variant})_n$ were constructed for AFM studies using maleimide-thiol coupling reaction, respectively[192]. All these rubredoxin variants are red color which show identical UV-Vis spectrum as that of WT-RD indicating the intact of FeS_4 center in these variants.

5.3.2 Mechanical anisotropy of the FeS_4 center at single bond level

We first stretched the FeS_4 center from direction of Fe-Cys5 and Fe-Cys41 bonds using RD1,49 variant. Stretching the polyprotein $(RD1,49)_n$ results in the

characteristic sawtooth like force-extension curves from unfolding of rubredoxin. Fitting the elasticity of the force trace by WLC (worm-like chain) model reveals force peaks with two different contour length increments (ΔLc): ~4.5 nm and ~13.0 nm.

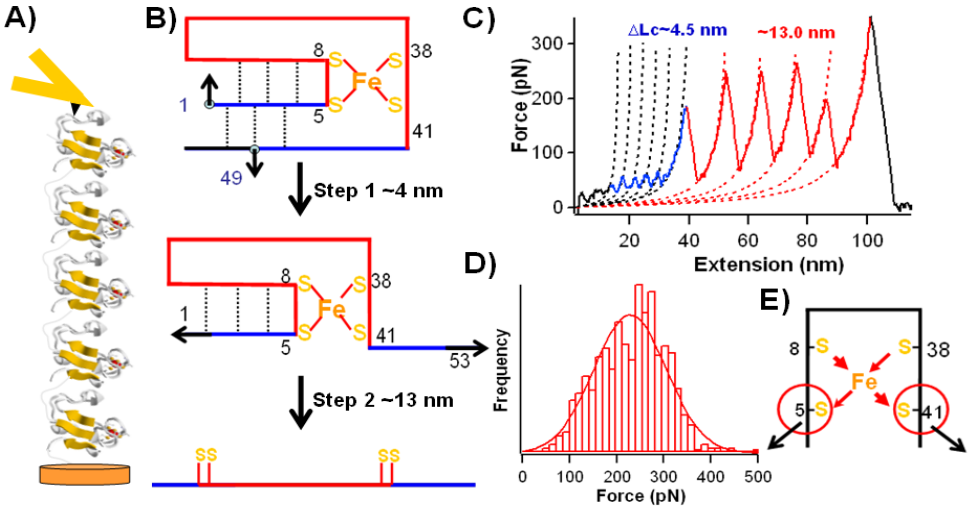

Figure 5.2 Mechanical stability of FeS_4 stretching from Fe-Cys5 and Fe-Cys41. A) Schematic of the mechanical unfolding experiment on RD1,49 by AFM. B) The two-step unfolding scenario of rubredoxin stretched from residue 1 and 49. C) Typical force extension curve of mechanical unfolding of RD1,49 showing two different types of force peaks. The force peaks with at the beginning are from the unfolding of residues before the FeS_4 center (in blue) leading to a ΔLc~ 4.5 nm. Then, rupture of the FeS_4 center and the extension of the previously buried residues (in red) give rise to force peaks with ΔLc ~13 nm. D) The histogram of the ruptured force for the FeS_4 center with an average value: 227±79 pN (n=1063). E) Schematic shows the FeS_4 center is stretched from direction of Fe-Cys5 and Fe-Cys41 bonds in RD1,49 (highlighted by black arrow and the cysteines are circled) and possible force distribution in the ferric thiolate bonds (in red arrow).

As shown in Figure 5.2B, the result agrees perfectly with a two-step unfolding scenario of RD1,49. The residues 1-4 and 41-49 outside the metal center are unfolded first and their extension leads to force peaks with a ΔLc of 4.4±1.0 nm (13aa*0.36 colored in blue). The FeS_4 center is then ruptured leading to the other force peaks with a ΔLc of 13.0±0.8 nm (residues 5-41, 37*0.36, colored in red) in the metal center. Thus, the force value 227±79 pN (n=1063) from force peak with ΔLc ~13 nm can be

unambiguously assigned to the rupture force of the FeS_4 center stretching initially from Fe-Cys5 and Fe-41bond.

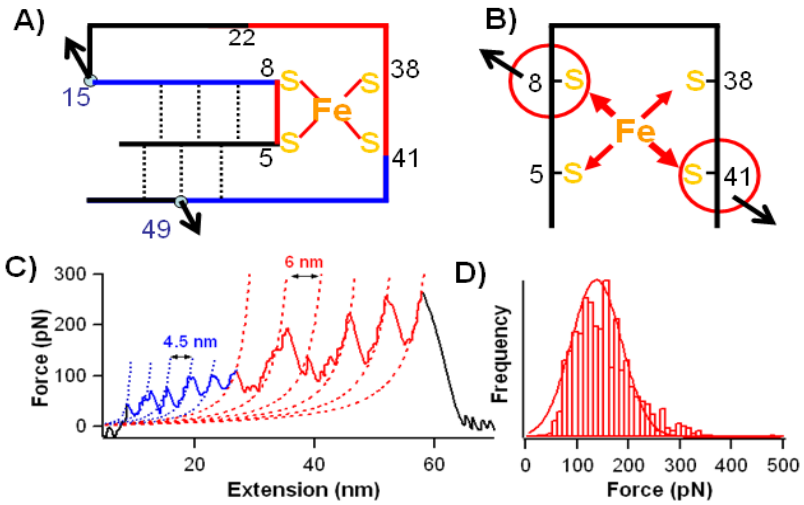

Figure 5.3 Mechanical stability of FeS_4 stretching from Fe-Cys8 and Fe-Cys41.A) Schematic shows the stretching scenario of RD15,49. The residues colored in red are further extended after the rupture of metal center. B) Simple structure of FeS_4 shows Fe-Cys8 and Fe-Cys41 bonds are stretched initially in RD15,49. C) Typical force extension curve of mechanical unfolding of RD15,49 shows a two-step unfolding scenario with force peak of ΔLc1~ 4.5 nm and ΔLc2 ~6.5 nm. D) Histogram of the rupture force with an average value of 152±60 pN (n = 708).

We then study the rupture of FeS_4 center with force applied initially through Fe-Cys8 and Fe-Cys41 bonds using polyprotein $(RD15,49)_n$ (Fig. 3A&B). Similarly as the two-step unfolding scenario of RD1,49, the AFM results show a first step with a force peak ΔLc of 4.5±0.8 nm ((8+9)*0.36-1.6 = 4.5 nm). Then, the FeS_4 center ruptures through Fe-Cys8 and Fe-Cys41 showing a force peak with ΔLc of 6.4±0.7 nm (Fig. 5D inset). Because only ~19 residues (residues 22-41, 19*0.36=6.8 nm) is further extended after the rupture of the metal center, the ΔLc is much less compared with that of RD1,49. It clearly demonstrates that the FeS_4 center in RD15,49 is stretched under a different direction. The corresponding rupture force is 152±60 pN

(n = 708), which is smaller than the value of RD1,49. It indicates that the stability of the FeS_4 varied considerable when stretched from different directions and proofs the mechanical anisotropy of FeS_4 center in rubredoxin. Compared with stretching the metal center in RD1,49, the only difference for RD15,49 is that the Fe-Cys8 bond is stretched as one direction instead of the Fe-Cys5 bond. Thus, the resultant mechanical anisotropy should be attributed to the different single ferric-thiolate bond.

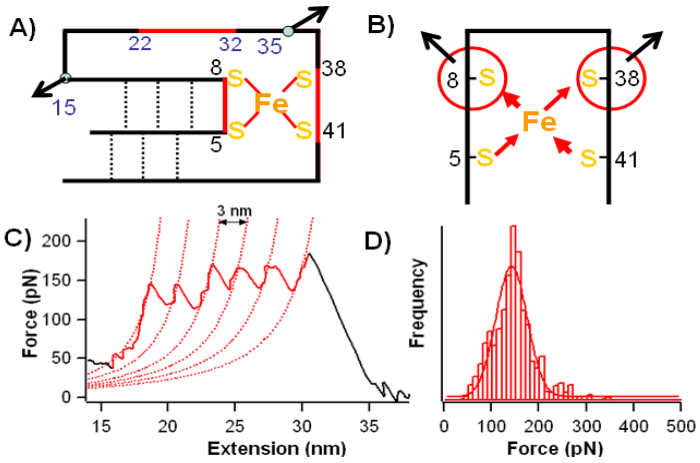

Figure 5.4 Mechanical stability of FeS_4 stretching from Fe-Cys8 and Fe-Cys38. A) Schematic shows the stretching scenario of RD15,35. B) Simple structure of FeS_4 shows Fe-Cys8 and Fe-Cys38 bonds are stretched initially in RD15,35. C) Typical force extension curve of mechanical unfolding of RD15,35 shows force peak with ΔLc of ~3 nm from the rupture of FeS_4 center (fitted by red line). D) Histogram of the force shows the average rupture force of the metal center is 146±49 pN (n = 985). The inset is the histogram of the ΔLc with average value of 2.7±0.4 nm.

We then carried out similar experiment on $(RD15,35)_n$ where the force applied first through Fe-Cys38 and Fe-Cys8 bonds. As a result, a different unfolding scenario is observed accordingly where only single force peak from metal center rupture event is present (Fig.4 A&B).The ΔLc of the force peak is only 2.7±0.4 nm which is from extension of residues from 22 to 32 after the rupture of metal center. The stability of FeS_4 measured at this force direction is 146±49 (n = 985).

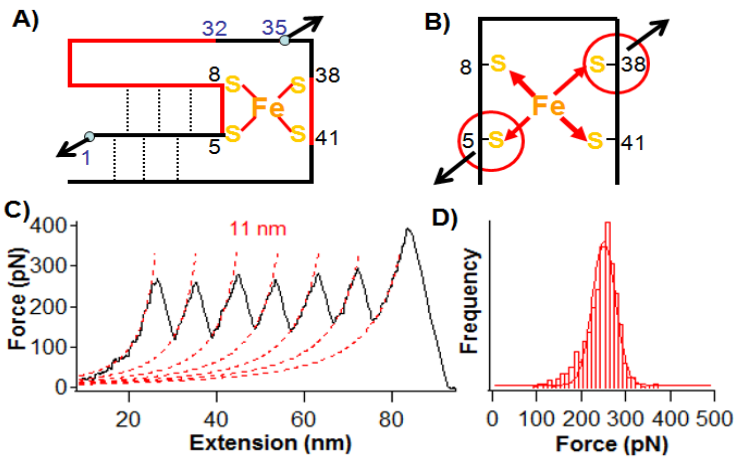

Figure 5.5 Mechanical stability of FeS_4 stretching from Fe-Cys5 and Fe-Cys41. A) Schematic shows the stretching scenario of RD1,35. B) Simple structure of FeS_4 shows Fe-Cys5 and Fe-Cys38 bonds are stretched initially in RD1,35. C) Typical force extension curve of mechanical unfolding of RD1,35 shows force peak with ΔLc ~11 nm from the rupture of FeS_4 center. D) The histogram of rupture force of metal center with an average value of 242±40 pN (n=1340).

Finally, stretching the metal center from Fe-Cys5 and Fe-Cys38 bonds is achieved using RD1,35 (Fig. 5.5A&B). Similar as RD15,35, there is only one type of force peak in the force extension curve from the rupture of metal center (Fig. 5.5C). The force peak shows a contour length increment of 11.1±1.2 nm (residues 5-32 and 38-41, 32*0.36). The force measured is 242±40 pN (n=1340) which is the highest among all four directions. Consequently, the measured stability of FeS_4 center under mechanical force applied from four different directions are all varied. It demonstrates the mechanical anisotropy of such a small structure.

5.3.3 The stochastic rupture mechanism of FeS_4 center stretched from Cys5 and Cys38 bond

The rupture of FeS_4 metal center from different directions shows a single force peak as a one step process. It indicates the multiple ferric-thiolate bonds are

cooperative in which all the four bonds appear to rupture simultaneously.

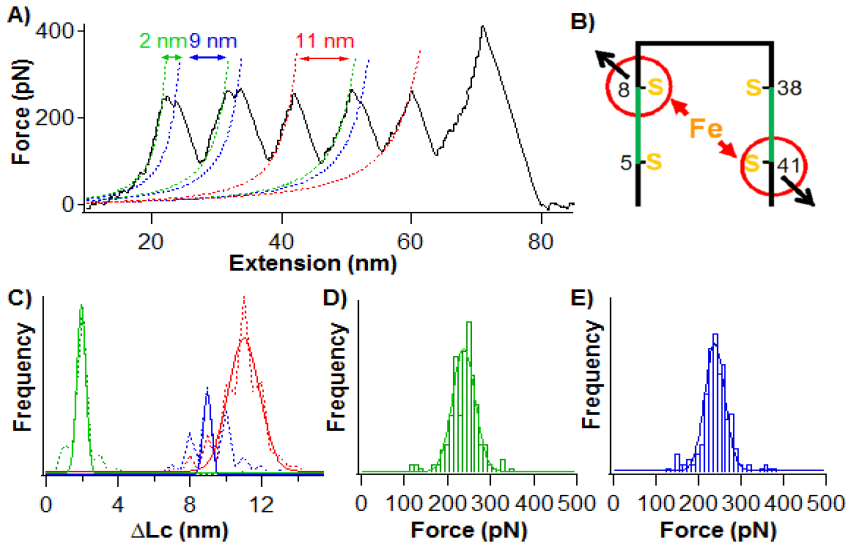

Figure 5.6 Stochastic rupture scenario found when stretched from Cys5 and Cys38. A) Force-extension curve reveals a two-step rupture event of metal center B) Simple structure of FeS_4 shows Fe-Cys8 and Fe-Cys38 bonds are stretched after the break of Fe-Cys5 and Fe-Cys38 which leads to the extension of six residues (colored in green). C) Histogram of the ΔLc reveals three different types of peaks: ΔLc1=11.1±1.2 nm (red), ΔLc2=1.9±0.3 nm (green) and ΔLc3=9.0±0.3 nm (blue). D-E) Histogram of the rupture forces from peaks with ΔLc2 (D) and ΔLc3 (E). The average force is 235 ±35 pN (n=177) and 239±38pN (n=173), respectively.

Interestingly, a close analysis of the force peaks in RD1,35 reveals small population of additional two-step rupture events (~10%). An intermediate force peak with ΔLc of ~2 nm is shown which can be assigned to an extension from ~6 amino acids. It is followed by another force peak with a ΔLc of ~9 nm. It suggests that the two Fe-Cys5 and Fe-Cys38 bonds break simultaneously while the other two Fe-Cys8 and Fe-C41 bonds are still present (Fig. 5.5B). Consequently, six residues (C5K6I7 and C38P39I40) from two CXXC binding motif can be extended at the same time which matches exactly with the experimental results. In addition, the further rupture of the metal center will lead to an extension of ~ 9 nm as observed. Moreover, the

sum of the ΔLc of the two peak is ~11nm which agrees with the result of the dominate one-step rupture scenario. The histogram of ΔLc shows clearly three different types of force peaks: ΔLc1=11.1±1.2 nm, ΔLc2=1.9±0.3 nm and ΔLc3=9.0±0.3 nm. The corresponding force value for the last two peaks are 235±35 pN (n=177) and 239±38pN (n=173), respectively. Thus, it indicates that a small population of metal center rupture in a two-step scenario. Consequently, a stochastic rupture mechanism in which the ferric-thiolate bonds can break both in cooperative and stepwise fashion is identified. It demonstrates that the direction of the applied force can also mediate the rupture mechanism of the metal center.

5.4 Discussion

5.4.1 Mechanical anisotropy is identified for the FeS_4 center.

By stretching the FeS_4 center from four different directions using single molecule AFM, we find that the mechanical stability of the metal center is depended on the direction of applied force. It demonstrates the mechanical anisotropy property of the simple FeS_4 center at single molecule level. Furthermore, one possible explanation is the difference among single ferric-thiolate bond which serves as the stretching direction. The measured rupture force of the metal center stretched from different directions is in the range from ~150 pN to ~250 pN. Compared with macromolecule such as proteins, the very simple FeS_4 center studied here still shows considerable mechanical anisotropy effect. Another possibility may arise from the asymmetric structure of the metal center. Consequently, the effective loading force on

the metal center may be different under different stretching sites. .

Moreover, from the aspect of metalloprotein, metal center is a critical part of metalloprotein. Besides its unique functional role, it can plays critical structural role as in the metalloprotein folding and stability, such as the iron in the iron-sulfur protein and zinc in zinc-finger protein. Thus, knowledge of the stability of metal center is of great importance. In addition, most metal center with multiple metal-ligand bonds is a three dimension structure inside metalloprotein. As a result, the investigation of the response of metal center to mechanical force from different directions is necessary to fully understand the structural role of metal center in metalloproteins. The AFM results here demonstrate a complete measurement of the mechanical stability of FeS_4 center in rubredoxin for the first time.

5.4.2 Applying well-controlled force to FeS_4 center from defined directions

In this experiment, it is critical to apply the mechanical force to the tiny structure of FeS_4 center under a defined direction. There are two important issues here: selecting the direction of applied force and indeed applying the force along this direction. Thanks to the simple structure of FeS_4 center, there are only four ferric-thiolate bonds in the structure which ultimately determine its stability. Thus, the four bonds naturally serve as the stretching direction. On comparison, it is a difficult task for a protein molecule. In principle, every residue in the protein molecule can be chosen as the stretching site. As a result, it is almost impossible to fully explore the mechanical anisotropy of this system. To simplify the selection, several models have

been proposed to define only a few arbitrary axes which need to be tested based on the topology of the protein. Here, the simple structure of FeS_4 center makes the choice of stretching axis much easier, and a complete investigation is achieved by stretching every possible combination of two ferric-thiolate bonds.

The most challenging part of this study is to ensure the force applied on the metal center is in a defined direction. It is achieved by unique protein unfolding result which is directly related to from which direction the FeS_4 center is ruptured. The force peaks from protein unfolding and rupture of FeS_4 center show characteristic contour length increment serving as a fingerprint to identify the force direction. In the AFM experiments of each four rubredoxin variants, the unfolding patterns upon the rupture of FeS_4 center are different. RD1,49 and RD15,49 show two force peaks upon unfolding while RD15,35 and RD1,35 show only one force peak. Additionally, the measured corresponding contour length increment is 13 nm, 6.5 nm, 3 nm and 11 nm, respectively. These values match perfectly with calculation from rupture of FeS_4 center from desired directions. Consequently, it is unambiguous that the force is applied to the metal center under defined directions.

5.4.3 A stochastic multiple ferric-thiolate bond rupture mechanism is found

Stretching the FeS_4 center in rubredoxin from different directions not only show different stabilities but also change the rupture mechanism in the terms of how these multiple metal-ligand bonds break. In most conditions, the rupture of the FeS_4 center occurs in a single step as a cooperative process. However, when stretched from

Fe-Cys5 bond and Fe-Cys38 bonds in RD1,35, two different force peak patterns from the rupture of the FeS_4 center are observed. Besides the dominant one step rupture event which is same as other conditions, there is a two-step rupture scenario. The intermediate force peak with 2 nm indicates a cooperative rupture mechanism in which the Fe-Cys5 and Fe-Cys38 bonds break concurrently. As a result, two different rupture mechanisms of the metal center in the unfolding of RD1,35 are observed. This stochastic rupture mechanism on the original FeS_4 center in rubredoxin agrees well with our previous finding on the center in a loop inserted rubredoxin mutant. Thus, the bond rupture mechanism is also dependent on the direction of applied force.

5.5 Method and materials

5.5.1 Protein engineering

The genes encoding RD1,49, RD15,49, RD15,35 and RD1,35 are engineered using a standard site-directed mutagenesis methods based on a WT-rubredoxin gene. Two selected original sites is substituted by cysteine condon in two sequential steps. The DNA sequences were confirmed directly by DNA sequencing. All the proteins were over-expressed in DH5α strain of *E. coli* and purified by Co^{2+}-affinity chromatography using TALON His-Tag purification resins (Clontech). Using a 3K MWCO Amicon ultra centrifugal filter, the protein solution was exchanged into Tris buffer (pH 8.5, 10 mM) (Millipore). Then the Fe-form rubredoxin variants were separated using an ion exchange chromatography. Finally, the polyproteins were obtained by a thiol-maleimide coupling reaction between the cysteines of rubredoxin and $BM(PEO)_3$[192].

5.5.2 Single molecule AFM experiments

A custom-built AFM is used for performing all the single-molecule AFM experiment[46]. The spring constant (typically around 40 pN/nm) of each Si_3N_4 cantilever (Veeco Corp.) was obtained in solution using the equipartition theorem before each experiment.

In an experiment, ~2 μL polyprotein solution at a concentration of ~2 mg/mL was added onto a clean glass coverslip. Covered by ~50 μL Tris buffer at pH ~7.5, the protein was allowed to absorb for ~5 minutes before starting the measurement. During the experiment, the cantilever was brought into contact with the substrate at a contact force of ~1 nN to pick up proteins with a pulling speed of 400 nm/s.

Chapter 6: The mechanical stability of FeS_4 center can be modulated by the protein structure[3]

6.1 Synopsis

Compared with previous AFM studies of chemical bonds, one of the most important features of metal-ligand bonds in metalloproteins is that these bonds are embedded within the protein structure, making the interplay between protein structure and the mechanical strength of metal-ligand bonds a novel and interesting research question. In this chapter, rubredoxin is utilized as an example to study how the mechanical stability of the FeS_4 center in rubredoxin is regulated by the surrounding protein structure.

It has long been recognized that hydrogen bonds formed by protein backbone amides with cysteinyl S^{γ} atoms play important roles in modulating the functional and structural properties of iron-sulfur centers in proteins. Using single molecule AFM, cyclic voltammetry and protein engineering techniques, the influence of N-H $\cdots S^{\gamma}$

[3] A version of this chapter has been published as " [Zheng P.], Takayama SJ. Mauk AG and Li H. (2012) Hydrogen bond strength modulates the mechanical strength of ferric-thiolate bonds in rubredoxin, Journal of the American Chemistry Society, 134(9), 4124-4131."

hydrogen bonds in the secondary coordination sphere on mechanical stability of Fe(III)-thiolate bonds of rubredoxin was previously investigated. Our results show that the mechanical stability of Fe(III)-thiolate bonds in rubredoxin correlate with the strength of N-H $\cdots S^{\gamma}$ hydrogen bonds as reflected by the midpoint reduction potential. This provides direct evidence that N-H $\cdots S^{\gamma}$ hydrogen bonds play important roles in modulating the mechanical and kinetic properties of the Fe(III)-thiolate bonds in iron-sulfur proteins, corroborating the important roles of the protein environment in tuning the properties of metal-thiolate bonds.

6.2 Introduction

Iron-sulfur proteins are ubiquitous in nature and play critical roles in a wide range of biological processes[154,156,159,190]. Facilitated by the highly covalent Fe-S bonds and the unique chemical properties of FeS clusters, iron-sulfur proteins are among the most important electron transfer proteins in nature and exhibit a broad range of reduction potential (-700 mV to +400 mV). Apart from the intrinsic characteristics of FeS clusters, it has long been recognized that the protein environment modulates the properties of Fe-S bonds to achieve the desired function and stability[159,191]. In particular, protein backbone amides form hydrogen bonds with cysteinyl S^{γ} atoms. These N-H $\cdots S^{\gamma}$ hydrogen bonds in the secondary coordination sphere are proposed to play important roles in modulating the functional and structural properties of the iron-sulfur centers[224-226]. For example, both crystallography and NMR studies on point mutants of the simplest iron-sulfur protein

rubredoxin revealed that the reduction potential of rubredoxin is correlated with the strength of N-H $\cdots S^{\gamma}$ hydrogen bonds[227-229]. These N-H $\cdots S^{\gamma}$ hydrogen bonds were also proposed to be responsible for the decreased covalency of Fe(III)-thiolate bonds of rubredoxins relative to their inorganic analogues, as the formation of hydrogen bonds influences electron delocalization between sulfur and iron[169,225,230]. However, direct experimental evidence concerning the quantitative contributions of these backbone hydrogen bonds to the stability of rubredoxins and the strength of Fe(III)-thiolate bonds remains limited.

Over the last two decades, the development of single molecule atomic force microscopy (AFM) has enabled measurement of the mechanical and kinetic properties of chemical bonds (both covalent and non-covalent) at the single molecule level along a well-defined reaction coordinate set by the vector of the applied stretching force[1,14 33,106,231]. In Chapter 3, we reported that the stability of a metal-thiolate bond in a protein could be measured for the ferric-thiolate bond at the active site of rubredoxin by single molecule AFM[201]. We found that, despite their highly covalent nature, Fe(III)-thiolate bonds exhibit surprisingly low mechanical stability (~200 pN at a pulling speed of 400 nm/s). Furthermore, the rupture force of ferric-thiolate bonds is much greater than that of Fe(II)-thiolate bonds, and the unfolding force of Fe(III)-thiolate bonds in *C. pasteurianum* rubredoxin (*cp*Rd) is higher than that in *P. furiosus* rubredoxin (*pf*Rd). The order of the mechanical stability of Fe-S bonds correlates with bond covalency.

To evaluate the contributions of hydrogen bonds to the properties of Fe-S bonds

and the correlation between covalency and mechanical stability in greater depth, we have now combined protein engineering, cyclic voltammetry (CV) and single molecule AFM technique to test directly whether the strength of hydrogen bonds formed by backbone amides and cysteinyl S^{γ} atoms can modulate the mechanical/kinetic stability of metal-thiolate bonds in the model protein *pf*Rd directly.

6.3 Results

6.3.1 Design of proline and glycine *pf*RD variants

In the second coordination sphere of *pf*Rd, multiple residues (Ile7, Cys8, Tyr10, Ile40, Cys41 and Ala43) have been identified to form N-H $\cdots S^{\gamma}$ hydrogen bonds (Figure 6.1) that involve backbone amides and the Fe-S center[163,232,233]. To investigate the effect of hydrogen bond strength on the mechanical properties of Fe(III)-thiolate bonds, we chose residue Ile7 and Ala43 as the sites to introduce point mutations for specific perturbation of the N…H-S^{γ} hydrogen bond strength (Figure 6.1B-C).

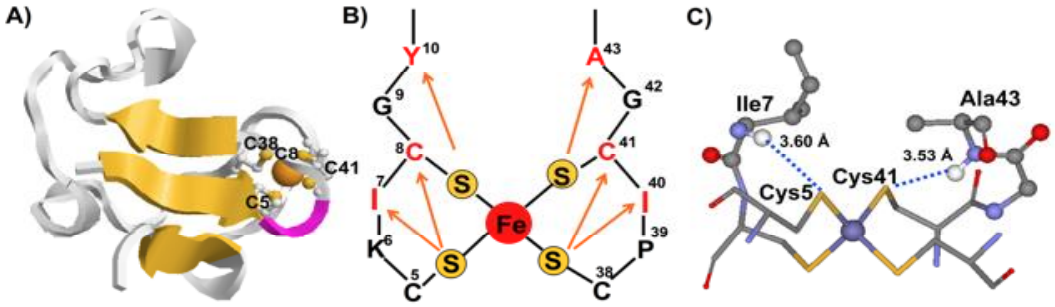

Figure 6.1 Schematics of the hydrogen bond network formed by backbone amides and cysteinyl S^{γ} atoms in *pf*Rd. A) 3D structure of *pf*Rd. The three-stranded antiparallel β-sheet structure and the β-turn including Lys6 and Ile7 are highlighted in yellow and pink, respectively. Cys5 and Cys38 are interior residues and Cys8 and Cy41 are exterior residues. B) Cartoon representation of residues in the secondary coordination sphere that are involved in the formation of hydrogen bond with FeS_4 center as identified

from X-ray diffraction studies[232,234]. The four cysteinyl S^{γ} atoms are denoted by S, and the other letters are single-letter amino acid designation. C) Close-up view of the structure of *pf*Rd in the vicinity of the N-H $\cdots S^{\gamma}$ bond between Ile7 and Cys5 and between Ala43 and Cys41. N-H $\cdots S^{\gamma}$ bonds are indicated by dashed lines.

To achieve the greatest possible range of hydrogen bond strength by mutation, proline and glycine substitutions were introduced at both sites (Ile7Pro and Ile7Gly, Ala43Pro and Ala43Gly). Proline is an imino acid and lacks a backbone amide hydrogen to serve as a hydrogen bond donor in the formation of an N…H-S^{γ} hydrogen bond. Introduction of a proline residue has been used previously to assess the role of backbone hydrogen bonds in protein structure and function, including the blue-copper protein cupredoxin[235-237]. Replacing Ile7 with proline prevents residue 7 from forming an N…H-S^{γ} hydrogen bond with residue Cys5, while replacing Ala43 with proline eliminates the N…H-S^{γ} hydrogen bond between Ala43 and Cys41. On the other hand, glycine was shown to form the strongest hydrogen bond N-H $\cdots S^{\gamma}$ in *cp*Rd, as indicated by NMR and electrochemical analysis[1,191]. Consequently, these two pairs of variants allow assessment of the extremes in strength of the N-H $\cdots S^{\gamma}$ hydrogen bond that can be formed by residue 7 with Cys5 and by residue 43 with Cys41.

6.3.2 The FeS_4 centers and the secondary structures of rubredoxin variants remain intact

To confirm that the FeS_4 center is intact in the engineered rubredoxin variants, we measured the electronic absorption spectra of the ferric form of all four variants (Ile7Pro, Ile7Gly, Ala43Pro and Ala43Gly). As shown in Figure 6.2A, the electronic

spectra of the variants are indistinguishable from that of wt-*pf*Rd: all three spectra superimpose with each other and show identical characteristic absorption signals with maxima centered at 390 nm, 495 nm and 570 nm, which are attributed to the ligand to metal charge-transfer transitions of oxidized rubredoxins. In addition, the A_{495}/A_{280} ratio, which indicates the purity of the Fe-form rubredoxin, is ~0.36 for all three proteins. These results suggest that neither substitution caused significant structural changes to the FeS_4 center.

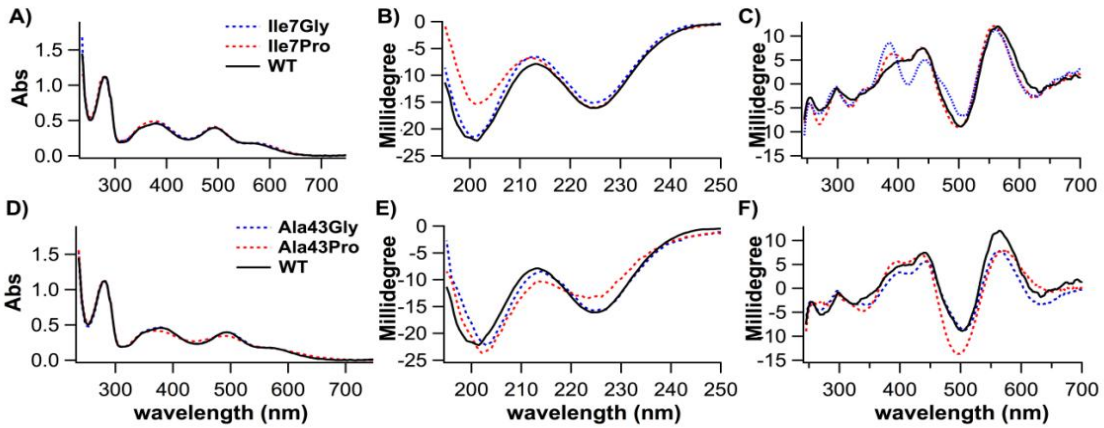

Figure 6.2 Absorption spectra of rubredoxin variants at positions 7 and 43 compared with wt-*pf*Rd. A) UV/Vis absorption spectra of Ile7Pro and Ile7Gly variants show indistinguishable feature as wt-*pf*Rd. The blue dash line is for Ile7Gly, the red is for Ile7Pro and the black solid line is for wt-*pf*Rd. All three spectra superimpose with each other. B) Far-UV CD spectra of Ile7Pro and Ile7Gly. C) UV/Vis CD spectrum of Ile7Pro and Ile7Gly. Color coding of the curves is the same for A-C). D) UV/Vis absorption spectra of Ala43Pro and Ala43Gly variants. The blue dash line is for Ala43Gly, the red is for Ala43Pro and the black solid line is for wt-*pf*Rd. All three spectra superimpose with each other. E) far-UV CD spectra of Ala43Pro and Ala43Gly F) UV/Vis CD spectra of Ala43Pro and Ala43Gly. Color coding of the curves is the same for D-F).

Moreover, we also obtained far UV and UV/Vis CD spectra to evaluate further the secondary structures of the variants as well as the environments of their aromatic amino acids and Fe(III) site. The far-UV CD spectra of the four variants are similar to

that of wt-*pf*Rd, suggesting that the three-stranded antiparallel β-sheet structure remains largely intact (Fig. 6.2A). The change of the band at 200 nm for Ile7Pro is consistent with the anticipated influence of the proline substitution on the β-turn structure. The CD spectra of the four variants in the visible range are also similar to that of wt-*pf*Rd and exhibit highly similar patterns with maxima at 320 nm (—), 350 nm (—), 400 nm (+), 440 nm (+), 505 nm (—), 560 nm (+) and 630 nm (—). This feature is also very similar to that observed for rubredoxins from other species[238]. The intensities of several transitions differ notably among the variants (wt-*pf*Rd, Ile7Pro, Ile7Gly, Ala43Pro and Ala43Gly), an observation that may also result from fine-tuning of the FeS_4 center by mutation-induced alteration of hydrogen bond strength.

6.3.3 CV confirms the relative order of the hydrogen bond strength

To confirm the relative order of hydrogen bond strengths of the pairs of rubredoxin variants at residues 7 and 43 (proline variant versus glycine variant) relative to wt-*pf*Rd, we undertook CV measurements to determine the reduction potentials of all variants as well as the wild-type protein because the reduction potential is correlated with the strength of the N-H $\cdots S^{\gamma}$ hydrogen bond[1,191]. A weaker H bond to the sulfur atom leads to a lower reduction potential. As shown in Figure 6.3A, the CVs of Ile7Pro shifted to a significantly lower reduction potential than observed for wt-*pf*Rd (-91 mV vs SHE (standard hydrogen electrode) for Ile7Pro vs -4 mV for wt), while the Ile7Gly variant exhibited a much higher reduction potential (42

mV vs SHE). Substitutions at residue 43 resulted in similar shifts in reduction potential: Ala43Pro(-40 mV) < wt(-4 mV) < Ala43Gly(13 mV) (Figure 6.3B).

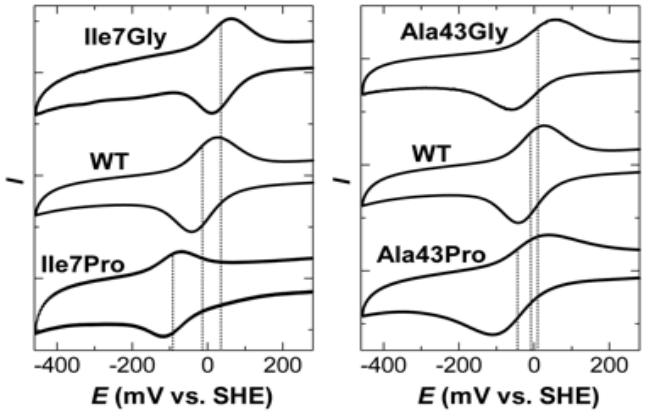

Figure 6.3 Reduction potential of *pf*Rd variants measured by CV. A) Compared with wt-*pf*Rd, the Ile7Pro shows a significantly lower reduction potential while Ile7Gly shows a higher reduction potential. This result suggests that the strength of the N…H-S^{γ} hydrogen bond increases following the order Ile7Pro<wt<Ile7Gly. B) CV measurement on the Ala43Pro and Alal43Gly shows a similar trend in reduction potential shift.

These results provide strong support for the conclusion that the glycine substitution strengthens the N-H $\cdots S^{\gamma}$ hydrogen bond significantly while the hydrogen bond is significantly weakened by proline substitution. This conclusion is consistent with the anticipated effect of a proline substitution to prevent the formation of an N-H $\cdots S^{\gamma}$ hydrogen bond.

6.3.4 The effect of the hydrogen bond on mechanical stability of the ferric-thiolate bonds

Having established the hydrogen bond strengths of these *pf*Rd variants, we used single molecule AFM to measure the mechanical stability of the ferric-thiolate bond in both *pf*Rd variants. To identify single molecule stretching events unambiguously, we

employed the well-established fingerprint polyprotein chimera approach and constructed polyprotein chimeras ($Rd_{variant}$-GB1)$_n$, in which the well-characterized GB1 domain serves as a fingerprint for identifying single molecule stretching events as well as an internal force caliper for the rupture force measurements of the ferric-thiolate bonds of rubredoxin[73,84,201,239]. We obtained electronic absorption and CD spectra (both far UV and UV/Vis) of the Rdmutant-GB1 protein chimeras and confirmed that as indicated by these criteria, the presence of GB1 in RD-GB1 chimera does not perturb the structure or electronic properties of the FeS_4 center relative to those of native rubredoxin.

We then used a maleimide-thiol coupling reaction to construct polyprotein chimeras ($Rd_{variant}$-GB1)$_n$ by reacting each $Rd_{variant}$-GB1 chimera which carries a cysteine residue at its N- and C-termini, with $BM(PEO)_3$[192,201]. The degree of polymerization n ranges from 2 to 5, as determined from SDS-PAGE. This result is similar to previous results for polyproteins constructed using the disulfide approach[98,99,240] or thiol-maleimide chemistry[192]. Although the dominant forms are dimers and trimers, polyproteins of higher degree of polymerization are clearly observed by SDS-PAGE. Because the heterogeneity of the length of the polyprotein has little effect on the measured unfolding force of proteins[42,99], the resulting polyproteins were used directly in the AFM pulling experiments without further purification. Moreover, the thiol-maleimide coupling reaction does not affect the properties of the FeS_4 centers of the RD domains coupled within the polyproteins.

As shown in Fig. 6.4A, stretching the polyprotein chimera ((Ile7Pro-GB1)$_n$

resulted in characteristic sawtooth-like force-extension curves in which each force peak corresponds to the unfolding of each domain. Fitting the force-extension curves to the Worm-Like Chain model of polymer elasticity[31] revealed that the unfolding force peaks exhibit two distinct types of contour length increment (ΔLc): 18 nm and 13 nm (Fig. 6.4A). The mechanical unfolding of the fingerprint GB1 domains, which has been studied in detail[192,201], is characterized by ΔLc of ~18 nm[187,196]. Thus, unfolding events with ΔLc of 18 nm can be readily attributed to the unfolding of GB1 domains[73,84,201]. Because Ile7Pro alternates with GB1 domains in the polyprotein chimera, unfolding events of ΔLc of 13 nm can accordingly be assigned to the unfolding of Ile7Pro without any ambiguity. The contour length increment of Ile7Pro (13 nm) is identical to that of wt-*pf*Rd and is consistent with the anticipated length increase resulting from rupture of the FeS_4 center. In Chapter 3[201,241], we showed that the unfolding of apo-rubredoxin itself does not contribute to the measured unfolding force of holo-Rd as the apo-Rd unfolds at forces <20 pN[196,236]. Therefore, the major event during the mechanical unfolding of the Ile7Pro variant can be readily assigned to the mechanical rupture of the FeS_4 center in this protein domain, and the unfolding force of Ile7Pro variant can be attributed to the rupture force of the FeS_4 center in the Ile7Pro domain. It is of note that two mechanisms of rupturing FeS_4 center in rubredoxin are possible. However, due to the limited resolution of the AFM, current data are insufficient to select between the two possible mechanisms.

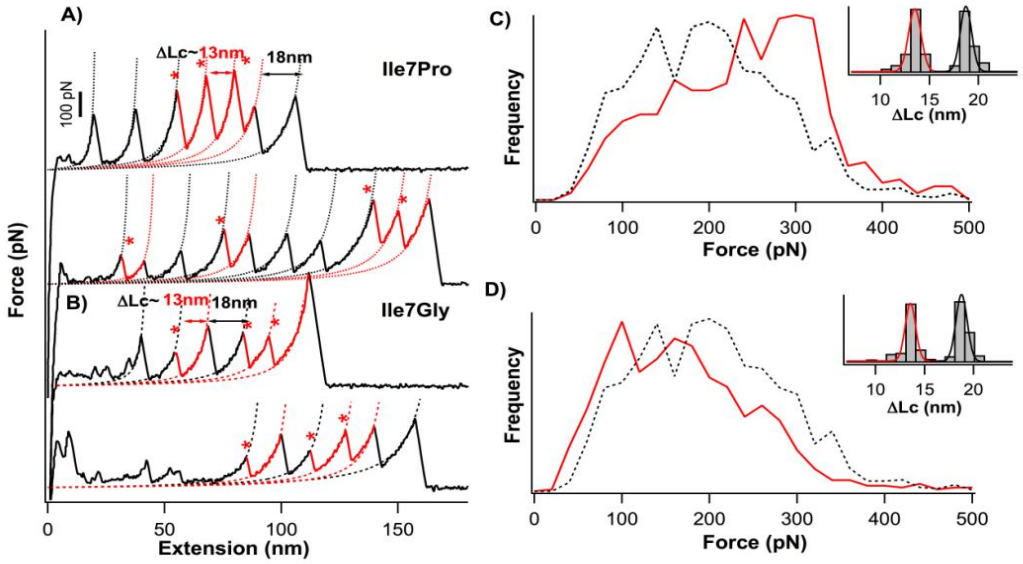

Figure 6.4 Measurement of mechanical stability of metal center in Ile7Pro and $Ile7Gly_n$. A-B) Typical force-extension curves of $(Ile7Pro\text{-}GB1)_n$ (A) and $(Ile7Gly\text{-}GB1)_n$ (B). The unfolding events of Ile7Pro and Ile7Gly are of ΔLc of ~13 nm. (C-D) Histograms of the mechanical rupture forces of Fe(III)-thiolate bonds in Ile7Pro (C) and Ile7Gly (D). The average rupture force of Fe-thiolate bonds is 255 ±4 pN for Ile7Pro (n = 564) and 182 ± 3 pN for Ile7Gly (n = 829), respectively. For comparison, the rupture force histogram of Fe(III)-thiolate bonds in wt-*pf*Rd is also shown (black dashed line). The inset is the ΔLc histogram of the mechanical unfolding of the corresponding polyprotein chimera.

As shown in Fig. 6.4C, the rupture force of FeS_4 in Ile7Pro exhibited a broad distribution that reflects the intrinsically short distance between the bound state and rupture transition state[33,201]. However, the average rupture force of the Ile7Pro variant is shifted towards a higher value (255 ± 4 pN for Ile7Pro (n = 564) versus 214 ± 3 pN for wt (n = 1534), the data are represented as mean ± standard error of the mean, where n represents the number of independent observations), suggesting that eliminating the $N\text{-}H\cdots S^{\gamma}$ hydrogen bond increases the mechanical stability of Fe(III)-thiolate bonds.

Similar analysis of the Ile7Gly variant revealed that the Fe(III)-thiolate bond of this variant ruptures at a lower force of 182 ± 3 pN (n = 829), suggesting that

increased N…H-S^{γ} bond strength leads to a decrease in the mechanical stability of Fe(III)-thiolate bonds. These results are consistent with the conclusion that the strength of the N-H $\cdots S^{\gamma}$ hydrogen bond indeed modulates the mechanical strength of the ferric-thiolate bonds of *pf*Rd.

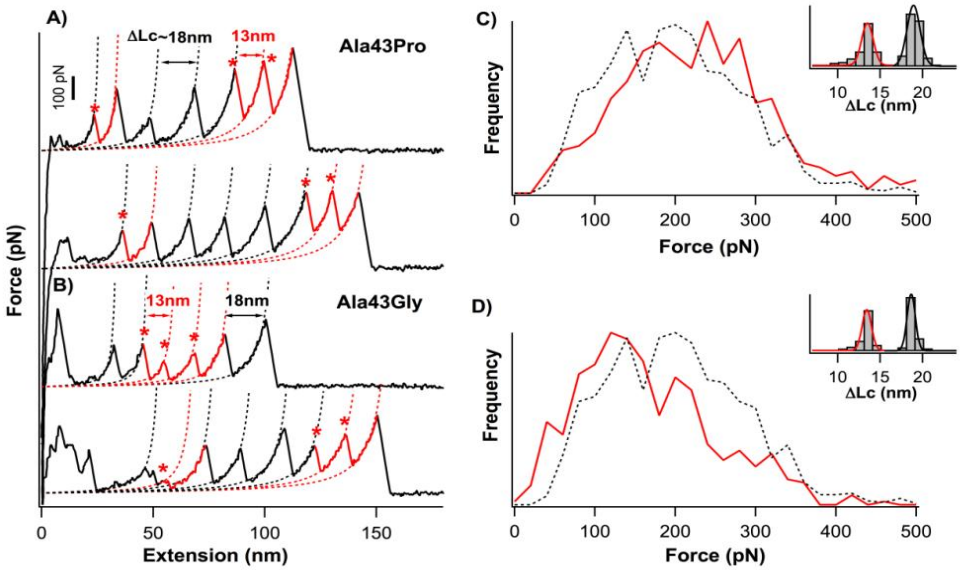

Figure 6.5 Mechanical unfolding of polyprotein chimeras $(Ala43Pro\text{-}GB1)_n$ and $(Ala43Gly\text{-}GB1)_n$. Typical force-extension curves of (A) $(Ala43Pro\text{-}GB1)_n$ and (B) $(Ala43Gly\text{-}GB1)_n$.(C-D) Histogram of the mechanical rupture forces of Fe(III)-thiolate bonds in (C) Ile43Pro and(D) Ile43Gly. The average rupture force of Fe-thiolate bonds is 248 ± 6 pN for Ala43Pro (n = 425) and 182 ± 4 pN for Ala43Gly (n = 482), respectively. It is clear that a much greater rupture force is required for the Ala43Pro variant than for the Ala43Gly variant. The inset is the histogram of the ΔLc of the mechanical unfolding of the corresponding polyprotein chimera.

To evaluate the generality of this effect, we undertook similar single molecule AFM analyses of *pf*Rd variants with substitutions for residue 43. Single molecule AFM experiments of the polyprotein chimeras $(Ala43Pro\text{-}GB1)_n$ and $(Ala43Gly\text{-}GB1)_n$ exhibited the expected increase in the mechanical unfolding force for Ala43Pro (248 ± 6 pN, n = 425) as well as the expected decrease for Ala43Gly (182 ± 4 pN, n = 482) (Fig. 6.5). These results for both pairs of variants reveal the

correlation between the reduction potential and the mechanical rupture force for ferric-thiolate bonds in *pf*Rd (Fig. 6.6), suggesting that this modulating effect is a general feature for rubredoxins.

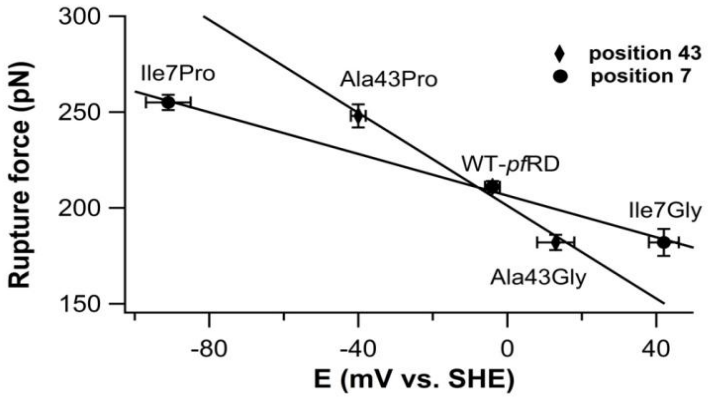

Figure 6.6 Dependence of the mechanical rupture force of the ferric-thiolate bond in *pf*Rd and its variants on reduction potential: (♦) position 43 variants; (●) position 7 variants. The solid lines are linear fits to the data.

We also carried out the mechanical rupture experiments at various pulling speeds (Figure 6.7). The dependence of the rupture force on the pulling speed exhibited by all variants remains, within experimental error, similar to that of wt-*pf*Rd, suggesting that the distance between the bound state and the mechanical rupture transition state Δx_u remains unchanged (0.14 nm) in both variants and that the difference in mechanical stabilities of ferric-thiolate bonds of the *pf*Rd variants studied here is largely attributable to the change of the lifetime of ferric-thiolate bonds at zero force. Monte Carlo simulations revealed that the average spontaneous dissociation rate constant of the Fe(III)-thiolate bond at zero force is 0.06 s^{-1} for Ile7Pro and Ala43Pro, 0.15 s^{-1} for wt and 0.25 s^{-1} for Ile7Gly and Ala43Gly, respectively.

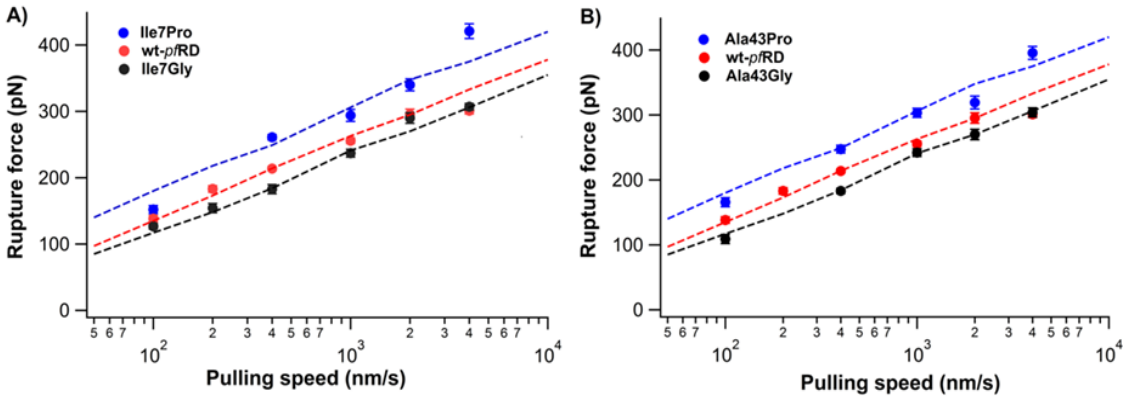

Figure 6.7 Different pulling speeds result on rubredoxin variants. The rupture forces of the Ile7Pro and Ile7Gly variants exhibited similar speed dependence to that of wt-*pf*Rd, suggesting that the distance between the bound state and mechanical rupture transition state Δx_u remains largely unchanged by the substitutions.

6.4 Discussion

6.4.1 Hydrogen bonds modulate the stability of ferric-thiolate bonds

Combining protein engineering, cyclic voltammetry and single molecule AFM technique, we have demonstrated that the mechanical stability of the ferric-thiolate bond is correlated with the reduction potentials of *pf*Rd variants. Because the reduction potential correlates with the strength of the hydrogen bonds, our results provide direct experimental evidence that the strength of the N-H $\cdots S^{\gamma}$ hydrogen bonds involving the protein backbone amide and cysteinyl S^{γ} atoms can play important roles in modulating the mechanical and kinetic properties of the ferric-thiolate bonds of iron-sulfur proteins. These results demonstrate another important role that the protein environment plays in tuning the properties of metal-thiolate bonds. To the best of our knowledge, the correlation of hydrogen bond strength with mechanical properties of ferric-thiolate bonds in metalloproteins has not been reported previously. It is now

clear that the electron delocalization between sulfur and iron atoms, which can be modulated by N-H $\cdots S^{\gamma}$ hydrogen bonds, directly determines the mechanical stability of Fe(III)-thiolate bonds. This conclusion is consistent with our observation that mechanical stability of Fe(III)-thiolate bonds is correlated with bond covalency[201].

It is well known that hydrogen bonds from the secondary coordination sphere that are widely observed in many different metalloproteins, play important roles in regulating the biological functions of metalloproteins[156,224,230]. Our current results demonstrate the utility of single molecule AFM in investigating the effect of such hydrogen bonds on active site stabilities of rubredoxin. We anticipate that this method can be applied to other metalloproteins to quantify the roles of these hydrogen bonds in modulating the stability of metal centers.

Furthermore, our experiments also provide potentially interesting model systems for spectroscopic studies of metal-thiolate bonds. It has long been proposed that the relatively low covalency of Fe(III)-thiolate bonds in rubredoxin is attributable to hydrogen bonding interactions formed with the sulfur atoms[169,170], The variants we report here exhibit a broad range of reduction potential, mechanical stability, and N-H $\cdots S^{\gamma}$ hydrogen bond strength that should serve as useful models for further use in K-edge X-ray absorption spectroscopy experiments to test this hypothesis in greater depth.

6.4.2 The influence of hydrogen bonds is site dependent

Although mutations at two different positions, Ile7 and Ala43, affect the

mechanical stability of the ferric-thiolate bond in rubredoxin, we note that the modulation of mechanical strength is more sensitive to the change of reduction potential at residue 43 than that at residue 7 (Fig. 6.6). The experimental slopes observed for this dependence were 1.24 pN/mV for the Ala43 series and 0.55 pN/mV for the Ile7 series. This observation suggests a site-dependent relationship between the hydrogen bond strength and mechanical stability of ferric-thiolate bond in rubredoxin. It is interesting to note that NMR studies of *cp*Rd also revealed a similar site-dependent relationship between hydrogen bond strength and reduction potential for variants at residue 44 (equivalent to the residue 43 series in *pf*Rd (*cp*Rd possesses an additional Met residue at the N-terminus)) and variants at residue 8 (equivalent to residue 7 in *pf*Rd).

Residues Ile7 and Ala43 in *pf*Rd (or residues 8 and 44 in *cp*Rd) exhibit considerably different structural characteristics[227,229]. Ile7 is located within the bi-cysteine coordination loop C5XXC8, which forms a rigid turn structure around the iron atom. In contrast, Ala43 is outside the C38XXC41 coordination loop and is subject to lesser steric strain upon mutation. Thus, it is possible that the site-specific effects on the mechanical stability of ferric-thiolate bond observed here arise from these structural differences.

Another possible basis for the functional inequivalence of these two sites is distinct difference in location of the two cysteine residues that are hydrogen bonded to residues 7 and 43 (Figure 6.1) in that Cys5 is located at an interior position and Cys41 has a more exterior location. These two cysteines may play different roles during the

mechanical rupture of FeS_4 center upon rubredoxin unfolding. Our previous results for variants of *pf*Rd in which Cys residues were replaced with His showed that the shorter Fe-thiolate bonds (Fe-Cys8 and Fe-Cys41) exhibit greater mechanical stability than do the longer Fe-thiolate bonds (Fe-Cys5 and Fe-Cys38)[201]. In addition, the hydrogen bond distance between Ile7 and Cys5 (3.60 Å) is greater than the hydrogen bond distance between Ala43 and Cys41 (3.53 Å). Thus, it is likely that the mechanical stability of the FeS_4 center is more sensitive to changes at residue 43, as residue 43 is hydrogen bonded to Cys41.

It is important to note that the effect of active site substitutions on the N…H-S^{γ} hydrogen bond strength is complex. Detailed NMR studies of *cp*Rd showed that substitutions of this type can have localized or aggregate effect on hydrogen bond strength[229]. For example, replacements for Val8 in *cp*Rd resulted in an aggregate effect on the strength of multiple hydrogen bonds that are not only localized to Cys6. In contrast, the effect of substitutions for Val44 can be largely attributed to Cys42 alone. Thus, it is also plausible that multiple factors are responsible for the observed site-dependent relationship between the hydrogen bond strength and mechanical stability of ferric-thiolate bond.

6.5 Materials and methods

6.5.1 Protein engineering

The genes encoding the Ile7Pro, Ile7Gly, Ala43Pro and Ala43Gly *pf*-rubredoxin variants were generated by standard site-directed mutagenesis methods using the

wild-type *pf*-rubredoxin gene as the template. Similarly, the genes encoding chimera Cys-Ile7Pro-GB1-Cys, Cys-Ile7Gly-GB1-Cys, Cys-Ala43Pro-GB1-Cys and Cys-Ala43Gly-GB1-Cys were constructed using Cys-wt-*pf*Rd-GB1-Cys gene as the template. All these constructs contain an N and C-terminal cysteine residue to facilitate the construction of polyproteins using maleimide-thiol chemistry. The gene was then cloned in the expression vector pQE80L between the *BamH*I and *Kpn*I sites, and the sequence was confirmed by direct DNA sequencing. Proteins were over-expressed and purified using the same method as in the previous chapter.

The degree of polymerization (n) was determined using SDS-PAGE. The linkages between consecutive RD-GB1 heterodimers (head-to-tail, head-to-head and tail-to-tail) in the resulting polyproteins are random. Because the stretching force is a vector, protein domains in the polyprotein will be subject to the same stretching force regardless of their linkages, and thus the AFM measurements are not affected by the orientation of the linkage between consecutive Rd-GB1[98,99,192].

6.5.2 Cyclic voltammetry

CV experiments were carried out with an Autolab PGSTAT12 potentiostat-galvanostat (Eco Chemie, The Netherlands) with an edge-plane pyrolytic carbon working electrode (PGE). The PGE was polished with alumina slurry and then sonicated in deionized water for 30 second before use. Typically, protein solution (2 μL, 2 mM) was spread onto the surface of the PGE with a micro-syringe and then covered with a semi-permeable membrane. A saturated calomel (SCE) electrode

(Radiometer, France) and platinum wire were used as the reference and counter electrodes, respectively. All experiments were carried out in sodium phosphate buffer (pH 7.0, 200 mM).

6.5.3 UV/Vis absorption spectroscopy experiments

The electronic absorption spectra of wild-type and variant RDs and of the Rd-GB1 chimeric proteins were recorded in Tris buffer (10 mM, pH 8.5) with a NanoDrop Model ND-1000 spectrophotometer at room temperature. The protein concentration was ~0.5 mM as determined from the absorbance of the solution at 495 nm (ε = 9.22 mM^{-1} cm^{-1})[162].

6.5.4 Circular dichroism spectroscopy experiments

CD spectra were recorded with a Jasco Model J810 spectropolarimeter using a quartz cuvette with a path length of 0.2 cm. For these measurements, protein samples with concentration of ~1 mM were used. For the far-UV CD measurements, the same protein samples were diluted with distilled water to ~10 μM for measurements.

6.5.5 Single molecule AFM experiments

Single molecule AFM experiments were performed on a custom-built AFM as reported[46]. The spring constant of each Si_3N_4 cantilever (Bruker Corp.) was calibrated in solution using the equipartition theorem prior to each experiment (typically ~40 pN/nm). All experiments were performed in Tris buffer (pH 7.4) at room temperature.

In a typical experiment, the polyprotein sample (2 μL, 2mg/mL) was added onto

a clean glass coverslip covered by Tris buffer (~50 μL, pH 7.4, 100 mM). The protein was allowed to absorb onto the coverslip for ~5 minutes before the AFM experiment. The polyprotein was picked up randomly along the contour of the polyprotein by means of non-specific adhesion with the AFM tip, leading to different number of unfolding force peaks in each force-extension curve. The use of polyproteins and GB1 fingerprint domains in this manner afforded unambiguous identification of single molecule stretching events.

The mechanical rupture process of the ferric-thiolate bonds can be modeled as a two-state dissociation process with force-dependent rate constants:

$$\alpha_0(F) = \alpha_0 \times \exp(F * \Delta x_u / k_B T)$$

where $\alpha_0(F)$ is the rate constant for dissociation at a stretching force F, α_0 is the spontaneous dissociation rate constant at zero force, Δx_u is the distance between the bound and transition states, k_B is Boltzmann constant and T is the absolute temperature. We estimated the dissociation rate constant α_0 at zero force and Δx_u by means of Monte Carlo simulations as reported previously[43,79].

Chapter 7: Thiocyanate substitution reaction on the ferric-thiolate bond in rubredoxin

7.1 Synopsis

In this chapter, the study of rubredoxin focuses on the chemical reactivity of metal-ligand bonds in metal center of metalloproteins. Metal-ligand bonds are chemically active among different types of chemical bonds which are subject to various chemical reactions. For example, ligand substitution reaction is a typical reaction in which new ligands compete with old ligands for metal ions. It is well known that ferric-thiolate bonds in synthetic analog of Fe-S cluster of iron-sulfur proteins can be substituted by a series of RS^- ligands.

However, it is difficult to directly probe the chemical reactivity between exogenous ligand and the metal-ligand bonds in metalloproteins. Because these bonds are sometimes buried deeply inside the protein structure, they are physically separated from the ligand in solution. Consequently, direct observation of chemical reaction between them is rare.

Here, we use single molecule AFM technique to solve this challenge. Previously, it was demonstrated that rubredoxin unfolds in a two-step scenario: the protein structure outside the FeS_4 center is unfolded firstly and these residues are extended, leading to the exposure of the previously buried FeS_4 center to the solution. If ligand

for ferric ion is present in the solution under this condition, a reaction between the ligand and the ferric-thiolate bonds in FeS_4 center can occur. Finally, the breaking of old ferric-thiolate bonds can be accelerated by applied force and monitored by the characteristic contour length increment upon rubredoxin unfolding.

7.2 Introduction

Iron-sulfur protein is ubiquitous in nature where Fe-S cluster is present as a critical metal site[156,159]. The incorporation of inorganic ferric/ferrous ion greatly expands the functionality and stability of proteins[154,186]. Thanks to the highly covalent nature of chemical bond between iron and sulfur atoms, these proteins play critical roles in a wide range of life processes, such as electron transfer, catalysis and gene regulation[159,167]. Thus, the investigation of chemical reactivity of iron sulfur centers/clusters in proteins is of great interest toward understanding of their unique stability and functions. Although the protein matrix protects the chemically active iron ion from aqueous environment, it impedes direct characterization of the metal center reactivity. Pioneering synthesis of inorganic analogs of the Fe-S cluster have revealed valuable information concerning their structure and reactivity[242]. For example, Dr. Holm and colleagues reported the first kinetic study of a ligand substitution reaction on a synthetic Fe_4S_4 cluster which is an analog of the active site from ferredoxin[243,244]. However, investigating chemical reactivity of metal centers in natural proteins is still a challenge under physiological conditions.

One example is the simple iron sulfur protein rubredoxin. It is a small protein of

53 residues which has a FeS_4 center as the active site[161,162]. The only ferric ion within the protein structure is coordinated by four sulfur atoms of cysteines forming four ferric-thiolate bonds. As shown in Fig. 7.1A, the ferric ion is buried inside protein structure and shielded from exogenous reactants in the solution. For example, addition of thiocyanate (SCN^-) which is a classic ferric ligand into the protein solution produces no reaction.

Mechanochemistry is an emerging research field which studies the influence of mechanical force on chemical reaction[1,47]. Similar to the classic thermal or photo induced reaction, the external mechanical force can also activate the chemical bond and facilitate the chemical reactions. The development of single molecule force spectroscopy techniques enable the exploration of bond rupture process and chemical reaction kinetics in a single protein molecule[14,106,200,202]. Furthermore, it is also widely used to study protein unfolding/folding process in which protein structure can be mechanically unfolded[42,76,86,245]. Consequently, it is an ideal tool to fulfill the requirement of studying chemical reactivity of metal center in metalloproteins. By applying external force to stretch metalloproteins, the protein structure which encloses the metal center will be unfolded, the metal center is exposed, and chemical reactions occurring between the exposed metal center and reactant in the solution can be monitored.

In previous chapters, we demonstrated the proof of concept for this type of study using rubredoxin.[201,202]. The AFM results show that the residues (1-5 and 41-53) which are outside the metal center can be unfolded and then extended as a first step

during protein unfolding. As a result, the FeS_4 center will be exposed and ruptured mechanically through the breaking of ferric-thiolate bonds. During the bond rupture process, it is highly possible that the exposed metal center is available for chemical reactions with reactant in the solution Thus it serves as an excellent model system to probe the chemical reactivity of metal center in metalloproteins.

This chapter details the mechanical rupture experiment on the ferric-thiolate bonds from rubredoxin in the present of ligand competing agent thiocyanate (SCN^-). We found that the SCN^- indeed accelerates the bond rupture process through a ligand substitution reaction with the ferric-thiolate bonds in the partially unfolded rubredoxin. Furthermore, the kinetics of the substitution reaction is revealed to be first order with respect to the SCN^-. This study opens up a new avenue towards studying the chemical reactivity of metal centers in metalloproteins using single molecule force spectroscopy techniques.

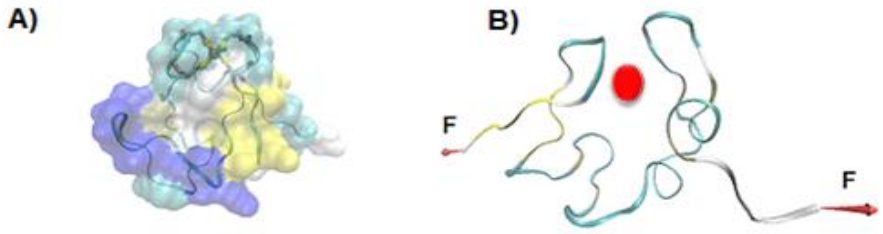

Figure 7.1 Schematic of folded and unfolded rubredoxin structure. A) Structure of the rubredoxin depicted in both ribbon and surface model reveals the FeS_4 center is buried by the protein structure. B) The cartoon shows that the FeS_4 center can be attacked by the SCN^- when the metal center is exposed by force.

7.3 Results

To investigate chemical reactivity of the FeS_4 center in rubredoxin, we construct a $(RD\text{-}GB1)_n$ polyprotein chimera for single molecule AFM experiment. It incorporates the well-studied GB1 domain as a fingerprint to help identify single molecule stretching event[73,84]. In addition, as a non-metalloprotein whose unfolding process (force) is not affected by the SCN^-, the GB1 domain serves as a control protein and force caliper in the experiment.

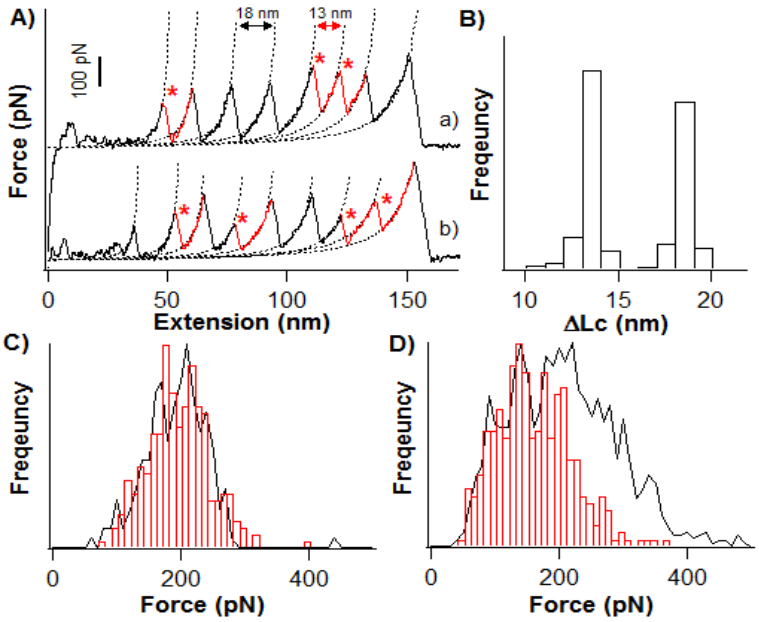

Figure 7.2 Mechanical unfolding of rubredoxin exposes the FeS_4 center to aqueous environment and allows the attacking of SCN^- to the ferric-thiolate bonds. A) Typical force-extension curves of stretching polyprotein $(RD\text{-}GB1)_n$ in Tris buffer (a) or 500 mM KSCN solution (b) show that the rupture of FeS_4 center leads to force peak with ΔLc of ~13 nm (highlight in red star) and unfolding of fingerprint GB1 domain lead to force peaks with ΔLc of ~18 nm. B) The histogram of contour length increment of $(RD\text{-}GB1)_n$ in 500 mM SCN shows two different distributions. C) The unfolding force histogram of non-metalloprotein GB1 domain shows similar value with/without SCN^-. D) The histogram of unfolding force in Tris buffer (black line) and 500 mM KSCN solution (red bar). It clearly shows the rupture force under the second condition is shifted toward a lower value (158 ± 58 pN (n = 342)).

Stretching $(RD\text{-}GB1)_n$ polyprotein in the presence of 500 mM KSCN solution results in typical saw-tooth like force-extension peaks (Fig. 7.2A curve b) with two types of contour length increments (ΔLc) which is similar as the result of previous experiments without KSCN (Fig. 7.2A curve a)[31,201]. The ΔLc1 of ~ 18 nm is from the unfolding of GB1 domain and the ΔLc2 of ~ 13 nm is from the breaking of FeS_4 center during the unfolding of rubredoxin[84,201]. The ΔLc histogram also shows these two distributions (Fig. 7.2B). In addition, the average unfolding force of GB1 domain is similar between two conditions (194 ± 48 pN (n = 284, with SCN^-) and 198 ± 54 pN (n =202, without SCN^-)). The AFM result indeed shows the addition of SCN^- does not affect the unfolding process of GB1.

On contrast, the rupture process of FeS_4 center is affected significantly by the addition of SCN^-. The measured rupture force of ferric-thiolate bonds under KSCN solution is considerable lower (average value: 158 ± 58 pN (n = 342)) than the one without SCN^- (211 ± 86 pN). Rupture forces with lower value are clearly more frequent on the histogram (Fig. 7.2C). Furthermore, the distribution of the rupture force which describes about the reaction transition state is quite different (58 pN vs 86 pN). It indicates that the bond rupture mechanism changes when SCN^- is added during the ferric-thiolate bond rupture process[2]. Thus, the AFM result exhibits that the thiocyanate indeed participate the rupture process of the ferric-thiolate bond when the FeS_4 center is exposed to solution under mechanical force. In addition, it also facilitates the dissociation of the original thiolate ligand.

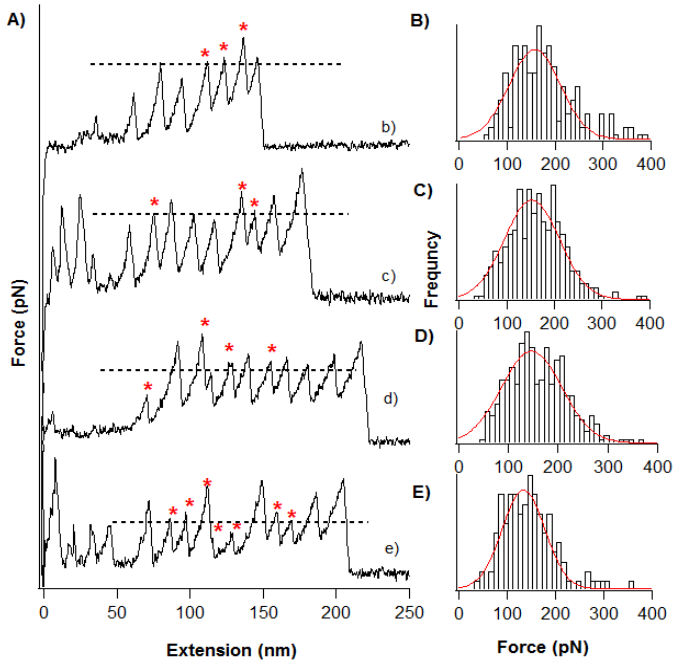

Figure 7.3 Mechanical unfolding of rubredoxin at various KSCN concentrations. A) Typical force-extension curve of polyprotein $(RD\text{-}GB1)_n$ under b) 5 mM; c)100 mM d) 500 mM e) 800 mM KSCN solution. B-E) the histogram of the rupture force of the FeS_4 center are shown and the force is 175 ± 68 pN, n =191; 162 ± 59 pN, n = 345; 158 ± 58 pN, n = 342 and 144 ± 55 pN, n = 130; respectively.

Because many ligand substitution reactions on the synthetic analogs of Fe-S cluster show that the reaction rate depends linearly on the concentration of the competing ligand[242,243]. To identify the effect of thiocyanate is from a SCN^- substitution reaction on the original ferric-thiolate bonds, we perform the AFM experiment as a function of KSCN concentration. The force extension curves of stretching the polyprotein $(RD\text{-}GB1)_n$ at various concentrations of KSCN solution are shown (Fig. 7.3A, 5 mM, 100 mM, 500 mM and 800 mM from top to bottom). As the concentration of SCN^- increases, the rupture force of ferric-thiolate bonds deceases accordingly. The average rupture forces are: 175 ± 68 pN (n = 191), 162 ± 59 pN (n = 345), 158 ± 58 pN (n=342) and 144 ± 55pN (n=130), respectively. Consequently, the

SCN^- effect is dependent on concentration and the result strongly supports a SCN^- substitution reaction on ferric-thiolate bonds during the mechanical unfolding of rubredoxin.

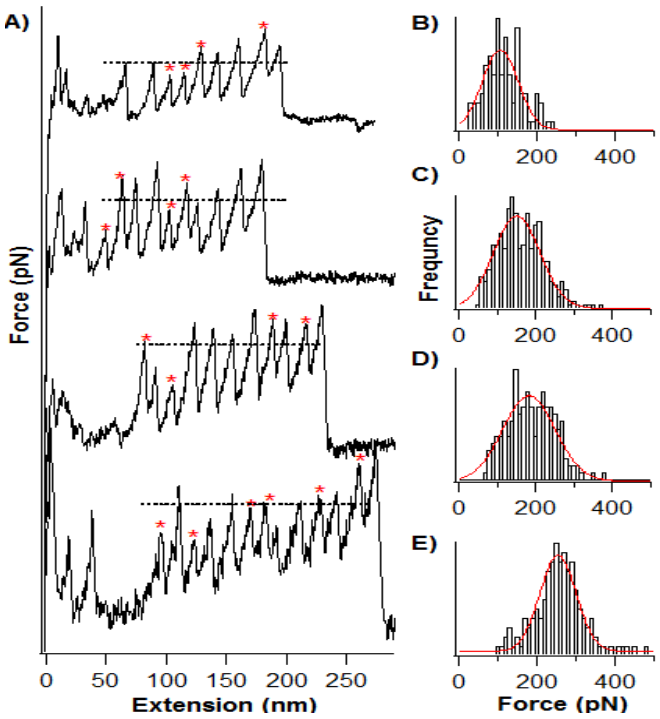

Figure 7.4 Mechanical unfolding of rubredoxin at different pulling velocity in the presence of 500 mM KSCN. A) typical force-extension curve of $(RD\text{-}GB1)_n$ under various pulling velocity from up to bottom: 100 nm/s, 400 nm/s, 1000 nm/s and 4000 nm/s B) The corresponding histogram of rupture force showing the average force is: 114 ± 46 pN n = 98, 158 ± 55 pN n =342, 185 ± 60 pN n = 184 and 251 ± 77 pN n=146, respectively.

To study kinetics of the SCN^- substitution reaction on the ferric-thiolate bonds, we carried out AFM experiment as a function of pulling velocity. Chemical reaction is accelerated exponentially under mechanical force whose degree is dependent on the exact force value and pulling velocity[33]. For example, when stretched under 500 mM KSCN solution (Fig. 7.4A), the measured rupture force is higher when the pulling velocity is increased. Moreover, we can extract two important kinetic parameters at

this condition using Monte Carlo simulation[43]: the spontaneous ferric-thiolate bond dissociation rate (k_0 = 0.37 s^{-1}) and the distance (Δx = 0.15 nm) between native bonded state to the mechanical transition state.

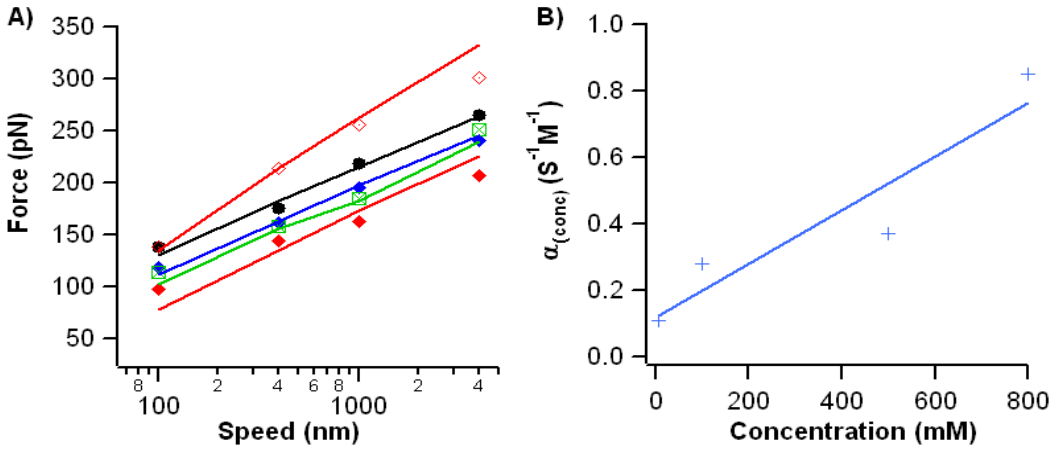

Figure 7.5 Kinetics of ferric-thiolate bond dissociation in the presence of thiocyanate. A) The plot of pulling velocity dependent experiment at different concentrations of KSCN: 800 mM (red), 500 mM (green), 100 mM (blue) and 5 mM (black). The solid line is the Monte Carlo fitting with a bond dissociation rate of 0.87, 0.37, 0.28 and 0.11 s^{-1}, respectively. Their distance between the bound state to the mechanical transition state are similar ~0.15 nm. The result in Tris buffer (red) is shown for comparison: k = 0.15 s and r = 0.11 nm. B) The relationship between the ferric-thiolate bond dissociation rate and the concentration of SCN^- is shown.

Furthermore, we repeated these different pulling velocity experiments at various SCN^- concentrations and obtained the relation between the ferric-thiolate bond dissociation rate and KSCN concentration (Fig. 7.5). The Δx for all conditions are similar ~ 0.15 nm, but the k_0 are largely different, with the value of 0.87, 0.37, 0.28 and 0.11 s^{-1} under 800 mM, 500 mM, 100 mM, and 5 mM KSCN solution, respectively. The relationship between them is plotted in Fig. 7.5B which shows a linear dependency and can be empirically fitted using equation: $r = k*[SCN^-]+0.11$

Where r is the ferric-thiolate bond dissociation rate in the presence of SCN^-, k is

a rate constant equals to 0.83 $M^{-1}s^{-1}$.

7.4 Discussion

7.4.1 SCN^- substitutes the ferric-thiolate bonds during mechanical unfolding of rubredoxin

Here, we demonstrate a direct thiocyanate substitution reaction on the ferric-thiolate bonds in rubredoxin by single molecule AFM technique. The AFM result shows that the presence of SCN^- reduces the mechanical stability of ferric-thiolate bonds and increases the bond dissociation rate. For example, the rupture force decreases significantly from 214 pN to 144 pN when 800 mM KSCN is added into the solution. Moreover, the bond dissociation rate increases to 0.87 s^{-1} from 0.15 s^{-1}, which is about 6 times higher. The presence of thiocyanate dramatically increased the leaving rate of the original thiolate ligand. Thus it clearly indicates the participation of SCN^- during the dissociation of ferric-thiolate bonds.

Furthermore, we found that this effect was linearly dependent on the SCN^- concentration which agreed with a typical second order ligand substitution reaction on ferric ion. Many ligand substitution reactions on the synthetic analogs of Fe-S cluster in iron-sulfur protein show a second order reaction which is both first order on the metal cluster and the competing ligand. For example, several ligand substitution reactions on the alkylthiolate tertramer dianion $Fe_4S_4(SR)_4$ which is the analog of active site of ferredoxin show that the reaction rate is linearly dependent on the concentration of the competing agent R'SH[243,244,246,247]. In our experiment, we

obtained the kinetic of the SCN^- substitution reaction on the ferric-thiolate bonds in a naturally occurring iron-sulfur protein. The substitution reaction is first order on the thiocyanate when pH is kept at 7.4 which agrees with the result on the inorganic analog of FeS cluster. For example, our substitution rate constant (0.87 s^{-1}) is similar as (2.10 s^{-1}) the rate of reaction: $[Fe_4S_4(S\text{-}t\text{-}Bu)_4]^{2-} + R'SH = [Fe_4S_4(S\text{-}t\text{-}Bu)_3(SR')]^{2-} +$ t-BuSH. R = *p*- $C_6H_4NH_2$. Thus, we provide the first evidence that a chemical reaction can be monitored directly in the wild-type iron sulfur protein under physiological condition at single molecule level.

7.4.2 AFM is a promising tool to study reactivity of metal center in metalloproteins

AFM is a simple and powerful tool to manipulate single molecule and measure tiny force in many research fields[1,20,60,76,141]. For example, by applying an external force to a single protein molecule, it can mechanically unfold a protein. In addition, it can also explore chemical bond rupture process. Here, we partially unfold rubredoxin which exposes the FeS_4 center to the surrounding solution by AFM. Consequently, it allows the attacking of thiocyanate in the solution to the ferric-thiolate bonds and the measurement of a SCN^- substitution reaction is achieved.

In this fashion, it solves two critical challenges of measurement of chemical reaction for metal center in metalloproteins. The first is to decouple the protein structure which separates the metal center from the competing agents. Some experiments on metal center which is not completely buried in metalloprotein

structure show that the addition of denaturant can significant increase the reaction rate[248]. It is believed that the denaturant distorts the protein structure and exposes more parts of the metal center to the solution. The second is to conduct the measurement in a limited time. Because many metal centers is kinetically labile in aqueous solution. Using single molecule AFM, the applied force increase the bond dissociation rate exponentially and allow the substitution reaction completed less than a second. Consequently, AFM can be a suitable tool to probe the reactivity of metal center in metalloproteins.

7.4.3 The nature of mechanical rupture of ferric-thiolate bonds

In this thesis, we demonstrate the mechanical rupture of ferric-thiolate bonds in rubredoxin. However, the mechanism of this mechanical bond activation (dissociation) process is largely unknown. Depending on how the electron split between the two atom after bond breaking, it can be heterolytic or hemolytic rupture process with different product (iron ion or iron radical)[60,197]. The AFM result here shows that the ferric-thiolate bond dissociation process can be accelerated by the addition of thiocyanate. As a second order ligand substitution reaction between the SCN^- and the ferric-thiolate bonds, a heterolytic rupture process is more likely to happen. It results in ferric ion which is ready to be attacked by thiocyanate. Similar heterolytic bond rupture result is obtained by computer simulation of PEG in solution[197].

In addition, we also found that the distance (Δx) between bound state to the mechanical transition state during the rupture of ferric-thiolate bond was different

with/without thiocyanate. Δx is parameter describing the mechanical transition state and is characteristic of particular reaction[2,33]. For example, the thiol-initiated and phosphine-initiated disulfide bond reduction which involves different reduction mechanism show two different Δx values[21]. Our result indicates that there is indeed a different bond rupture mechanism when the SCN^- is present.

7.4.4 The interplay between force and chemical reaction

The interplay between force and chemical reaction has been found for centuries. Similar as thermal energy, mechanical force can also induce chemical reaction. It is predicted that applying force on a chemical reaction leads to an exponential increase of the reaction rate based on Bell-Evans model[2,105]. However, different from the thermal energy, mechanical force is a vector with specific direction. As a result, the alignment of the direction of mechanical force and the reaction coordinate is a challenge in the bulk studies. The development of single molecule AFM enables the measurement of chemical reaction by mechanical force. Recently, a few chemical reactions have been probed by AFM. Here, our result on a substitution reaction on a metal-ligand bond strongly supports this idea.

7.5 Materials and methods

7.5.1 Protein engineering

The gene encoding protein chimera RD-GB1 was constructed as previously reported[192,201]. The protein was expressed using pQE80L vector in *Escherichia coli*

strain DH5α and purified by Co^{2+}-affinity chromatography using TALON His-Tag purification resins (Clontech) followed by an ion exchange chromatography using Mono Q 5/50 anion exchange column (GE healthcare). The resultant ferric-form rubredoxin chimera was reacted with a $BM(PEO)_3$ (1, 8-bis (maleimido)triethylene glycol, Molecular Biosciences) through a thiol-maleimide coupling reaction forming the polyprotein $(RD\text{-}GB1)_n^{192}$.

7.5.2 Single molecule AFM experiments

Single molecule AFM experiments were carried out on a custom-built AFM as reported[20]. Prior to each experiment, each Si_3N_4 cantilever (Bruker Corp.) was calibrated in solution using the equipartition theorem (with typical value ~40 pN/nm). All experiments were performed in Tris buffer (pH 7.4) at room temperature.

In a typical AFM stretching experiment, the polyprotein sample (2 μL, 2 mg/mL) was added onto a clean glass coverslip covered by Tris buffer (~50 μL, pH 7.4, 100 mM). The protein was allowed to absorb onto the coverslip for ~5 minutes before the experiment. 5 mM, 100 mM, 500 mM and 800 mM of KSCN were obtained by adding appropriate quantity of the 1M KSCN stock solution into the Tris buffer.

Chapter 8: Conclusion

8.1 Thesis summary

Mechanical force is ubiquitous in nature and plays critical roles for many biologically critical processes[34]. The importance of mechanical force has been increasingly recognized over the past two decades, and involves multiple fields of studies, such as biology, physics and chemistry[22,25,44,115,200]. This thesis demonstrates mechanical rupture processes of metal-ligand bonds in a single metalloprotein as elucidated using single molecule AFM[192,201,202,249]. It is found that ferric-thiolate bonds in a FeS_4 center of rubredoxin can be ruptured under mechanical force. This is the first unambiguous example demonstrating that a metal center in a biological system can be mechanically ruptured.

The investigation of metal-ligand bonds in metalloproteins using single molecule force spectroscopy demonstrates several unique features when compared with previous AFM studies of simple inorganic chemical bonds.

Firstly, the rupture of metal-ligand bonds in metalloprotein is typically a step during protein unfolding which shows a clear contour length increment. Consequently, the single or multiple bond rupture event can be identified without ambiguity. Using rubredoxin as a model system for single molecule AFM studies, we show that this protein unfolds in two steps: the protein structure outside the metal center in

rubredoxin is unfolded at first, and the FeS_4 center is then mechanically ruptured, leading to the extension of previously enclosed residues with a ΔLc of ~13 nm. This contour length increment is characteristic of the FeS_4 center rupture event, matching perfectly with the extension of the 36 residues enclosed by the metal center in the protein.

Secondly, multiple metal-ligand bonds in the metal sites are a common feature of metalloproteins. Thus, the investigation of a metal center involving multiple bonds raises many new questions regarding the rupture sequence/pathway. In rubredoxin, the FeS_4 center contains four ferric-thiolate bonds where the rupture of at least two bonds results in a break of the metal center. Using a loop insertion rubredoxin variant to increase the distance between the two ferric-thiolate bonds, we clearly distinguish between the two different rupture pathways. A complex stochastic rupture mechanism is identified in which both cooperative and sequential rupture scenarios happen. From the view of protein unfolding, these AFM results provide valuable information about the release process of iron from metal center, which is of important biological consequences for metalloproteins.

Thirdly, mechanical anisotropy of the FeS_4 is revealed. Similar as single protein molecule, we found the rupture force is dependent on the direction of applied force. Compared with protein, the FeS_4 is an extremely small object consists of only five atoms. The stretching direction can be naturally chosen as the four ferric-thiolate bonds. In addition, the challenge of applying force on this small object with well-controlled direction is overcome by stretching the four designed sites.

Consequently, the property of mechanical anisotropy is found valid for such as small system.

In addition, the influence of protein structure on embedded metal-ligand bonds is a unique feature of metalloproteins. The effect of amide hydrogen bond to the ferric-thiolate bond strength in rubredoxin is chosen as a model system. Using proline and glycine mutation to rationally decrease and increase the strength of hydrogen bond to the sulfur atom, the corresponding mechanical stability of the FeS_4 center in rubredoxin is varied significantly.

Finally, the metal-ligand bonds which are chemically active serve as a good system to study their chemical reactivity. Using a classic ferric ligand thiocyanate, we demonstrate a ligand substitution reaction on the ferric-thiolate bonds during the mechanical unfolding of rubredoxin. The reaction is found to be first order on thiocyanate which agrees with the result of typical second order ligand substitution reaction on iron cluster.

In summary, using rubredoxin as a model system, the mechanical rupture of metal-ligand bonds in metalloprotein is reported for the first time. Many aspects of ferric-thiolate bonds in rubredoxin are systematically studied, such as the direct measurement of bond strength, the rupture mechanism of the multiple bonds, the interplay between the bond strength and protein structure and the chemical reactivity of metal center. These findings increase our knowledge of metal-ligand bonds in metalloproteins, as well as providing a general methodology towards investigating metal-ligand bonds in different metalloproteins.

8.2 Future directions

In fact, metalloproteins are ubiquitous in nature which constitute more than one third of all proteins. Rubredoxin is the simplest iron sulfur protein, and there are hundreds of iron sulfur proteins with a more complex iron center/cluster. Further investigating this class of metalloproteins could provide more evidences for the result discovered here, and provide further information regarding the nature and characteristics of metal-ligand bonds.

Many other transition metals are also present in metalloproteins besides iron, such as copper and zinc. Moreover, many residues in proteins can bind to metal. As a result, metalloproteins provide a huge amount of subjects to probe different types of metal-ligand bonds for AFM studies. One future direction is to study other types of metalloproteins besides those involving iron. Particularly, many metal-ligand bonds in the metal center of metalloproteins play critical structural roles in protein stability and folding. For example, the zinc-thiolate bonds present in the zinc-finger proteins are required for its folding. Consequently, direct measurements of the mechanical stability of these bonds are of great biological interest.

In chapter 7, a simple SCN^- ligand substitution reaction on the ferric-thiolate bonds in rubredoxin is discussed. This experiment demonstrates the potential of using metalloproteins as a system to study the chemical reactivity of metal centers by AFM. As described above, the huge numbers of metalloproteins, and the wide range of metal-ligand bonds they represent, offer a great opportunity to study many different

types of chemical reactions that could occur within biological systems. For example, protonation chemistry of Fe-S cluster is critical for function of several enzymes. It is proposed the conversion of N_2 to NH_3 in the nitrogen fixation process by certain enzymes is operated in a protic environment. Consequently, the investigation of mechanical rupture process of metal centers in acidic environment is important to understand chemical reactions as they occur in biologically relevant reactions.

In summary, we believe metalloproteins serve as excellent subjects for AFM studies and we hope more interesting discoveries would be revealed in the near future.

Reference

(1) Beyer, M. K.; Clausen-Schaumann, H. *Chem. Rev.* **2005**, *105*, 2921.

(2) Evans, E.; Ritchie, K. *Biophys. J.* **1997**, *72*, 1541.

(3) Lea, M. C. *Philosophical Magazine Series 5* **1892**, *34*, 46.

(4) Caruso, M. M.; Davis, D. A.; Shen, Q.; Odom, S. A.; Sottos, N. R.; White, S. R.; Moore, J. S. *Chem. Rev.* **2009**, *109*, 5755.

(5) Nguyen, T. Q.; Kausch, H. H. *Macromolecules* **1990**, *23*, 5137.

(6) Smith, S. B.; Finzi, L.; Bustamante, C. *Science* **1992**, *258*, 1122.

(7) Tskhovrebova, L.; Trinick, J.; Sleep, J. A.; Simmons, R. M. *Nature* **1997**, *387*, 308.

(8) Florin, E.; Moy, V.; Gaub, H. *Science* **1994**, *264*, 415.

(9) Alon, R.; Hammer, D. A.; Springer, T. A. *Nature* **1995**, *374*, 539.

(10) Cluzel, P.; Lebrun, A.; Heller, C.; Lavery, R.; Viovy, J. L.; Chatenay, D.; Caron, F. *Science* **1996**, *271*, 792.

(11) Evans, E.; Ritchie, K.; Merkel, R. *Biophys. J.* **1995**, *68*, 2580.

(12) Greenleaf, W. J.; Woodside, M. T.; Block, S. M. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 171.

(13) Neuman, K. C.; Nagy, A. *Nat. Methods* **2008**, *5*, 491.

(14) Grandbois, M.; Beyer, M.; Rief, M.; Clausen-Schaumann, H.; Gaub, H. E. *Science* **1999**, *283*, 1727.

(15) Binnig, G.; Quate, C. F.; Gerber, C. *Phys. Rev. Lett.* **1986**, *56*, 930.

(16) Binnig, G.; Rohrer, H. *Ibm Journal of Research and Development* **1986**, *30*, 355.

(17) Hansma, P. K.; Elings, V. B.; Marti, O.; Bracker, C. E. *Science* **1988**, *242*, 209.

(18) Florin, E. L.; Rief, M.; Lehmann, H.; Ludwig, M.; Dornmair, C.; Moy, V. T.; Gaub, H. E. *Biosensors & Bioelectronics* **1995**, *10*, 895.

(19) Hutter, J. L.; Bechhoefer, J. *Review of Scientific Instruments* **1993**, *64*, 1868.

(20) Carrion-Vazquez, M.; Oberhauser, A. F.; Fisher, T. E.; Marszalek, P. E.; Li, H.; Fernandez, J. M. *Prog. Biophys. Mol. Biol.* **2000**, *74*, 63.

(21) Ainavarapu, S. R. K.; Wiita, A. P.; Dougan, L.; Uggerud, E.; Fernandez, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 6479.

(22) Borgia, A.; Williams, P. M.; Clarke, J. *Annu. Rev. Biochem.* **2008**, *77*, 101.

(23) Cao, Y.; Yoo, T.; Li, H. B. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 11152.

(24) Guzman, D. L.; Randall, A.; Baldi, P.; Guan, Z. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1989.
(25) Li, H. B. *Adv. Funct. Mater.* **2008**, *18*, 2643.
(26) Peng, Q.; Li, H. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 1885.
(27) Schmidt, S. W.; Beyer, M. K.; Clausen-Schaumann, H. *J. Am. Chem. Soc.* **2008**, *130*, 3664.
(28) Schwaderer, P.; Funk, E.; Achenbach, F.; Weis, J.; Brauchle, C.; Michaelis, J. *Langmuir* **2007**, *24*, 1343.
(29) Oberhauser, A. F.; Hansma, P. K.; Carrion-Vazquez, M.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 468.
(30) Oesterhelt, F.; Rief, M.; Gaub, H. E. *New J. Phys.* **1999**, *1*, 6.
(31) Marko, J. F.; Siggia, E. D. *Macromolecules* **1995**, *28*, 8759.
(32) Bell, G. I. *Science* **1978**, *200*, 618.
(33) Evans, E. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 105.
(34) Bustamante, C.; Chemla, Y. R.; Forde, N. R.; Izhaky, D. *Annu. Rev. Biochem.* **2004**, *73*, 705.
(35) Lee, C.-K.; Wang, Y.-M.; Huang, L.-S.; Lin, S. *Micron* **2007**, *38*, 446.
(36) De Paris, R.; Strunz, T.; Oroszlan, K.; Güntherodt, H.-J.; Hegner, M. *Single Molecules* **2000**, *1*, 285.
(37) Merkel, R.; Nassoy, P.; Leung, A.; Ritchie, K.; Evans, E. *Nature* **1999**, *397*, 50.
(38) Lo, Y.-S.; Zhu, Y.-J.; Beebe, T. P. *Langmuir* **2001**, *17*, 3741.
(39) Piramowicz, M. D.; Czuba, P.; Targosz, M.; Burda, K.; Szymonski, M. *Acta Biochi. Pol.* **2006**, *53*, 93.
(40) Moy, V. T.; Florin, E. L.; Gaub, H. E. *Science* **1994**, *266*, 257.
(41) Teulon, J.-M.; Delcuze, Y.; Odorico, M.; Chen, S.-w. W.; Parot, P.; Pellequer, J.-L. *J. Mol. Recognit.* **2011**, *24*, 490.
(42) Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3694.
(43) Rief, M.; Fernandez, J. M.; Gaub, H. E. *Phys. Rev. Lett.* **1998**, *81*, 4764.
(44) Wielert-Badt, S.; Hinterdorfer, P.; Gruber, H. J.; Lin, J.-T.; Badt, D.; Wimmer, B.; Schindler, H.; Kinne, R. K. H. *Biophys. J.* **2002**, *82*, 2767.
(45) Brant, D. A. *Curr. Opin. Struct. Biol.* **1999**, *9*, 556.
(46) Fernandez, J. M.; Li, H. B. *Science* **2004**, *303*, 1674.
(47) Puchner, E. M.; Gaub, H. E. *Annu.Rev. Biophys.* **2012**, *41*, 497.
(48) Pethica, J. B.; Oliver, W. C. *Phys. Scr.* **1987**, *T19A*, 61.
(49) Lantz, M. A.; Hug, H. J.; Hoffmann, R.; van Schendel, P. J. A.; Kappenberger, P.; Martin, S.; Baratoff, A.; Güntherodt, H.-J. *Science* **2001**, *291*, 2580.
(50) Perez, R.; Payne, M. C.; stich, I.; Terakura, K. *Phys. Rev. Lett.* **1997**, *78*, 678.
(51) Hoffmann, R.; Lantz, M. A.; Hug, H. J.; van Schendel, P. J. A.;

Kappenberger, P.; Martin, S.; Baratoff, A.; Guntherodt, H. J. *Phys. Rev. B* **2003**, *67*.
(52) Rubio-Bollinger, G.; Bahn, S. R.; Agrait, N.; Jacobsen, K. W.; Vieira, S. *Phys. Rev. Lett.* **2001**, *87*.
(53) Florin, E. L.; Moy, V. T.; Gaub, H. E. *Science* **1994**, *264*, 415.
(54) Hinterdorfer, P.; Baumgartner, W.; Gruber, H. J.; Schilcher, K.; Schindler, H. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 3477.
(55) Kamruzzahan, A. S. M.; Ebner, A.; Wildling, L.; Kienberger, F.; Riener, C. K.; Hahn, C. D.; Pollheimer, P. D.; Winklehner, P.; Hölzl, M.; Lackner, B.; Schörkl, D. M.; Hinterdorfer, P.; Gruber, H. J. *Bioconjugate Chemistry* **2006**, *17*, 1473.
(56) Barattin, R.; Voyer, N. *Chem. Commun.* **2008**, *0*, 1513.
(57) Bowers, C. M.; Carlson, D. A.; Shestopalov, A. A.; Clark, R. L.; Toone, E. J. *Biopolymers* **2012**, *97*, 761.
(58) Alegre-Cebollada, J.; Kosuri, P.; Rivas-Pardo, J. A.; Fernandez, J. M. *Nat. Chem.* **2011**, *3*, 882.
(59) Garnier, L.; Gauthier-Manuel, B.; van der Vegte, E. W.; Snijders, J.; Hadziioannou, G. *J. Chem. Phys.* **2000**, *113*, 2497.
(60) Ribas-Arino, J.; Marx, D. *Chem. Rev.* **2012**, *112*, 5412.
(61) Conti, M.; Falini, G.; Samori, B. *Angew. Chem. Int. Ed.* **2000**, *39*, 215.
(62) Schmitt, L.; Ludwig, M.; Gaub, H. E.; Tamper, R. *Biophys. J.* **2000**, *78*, 3275.
(63) Kienberger, F.; Kada, G.; Gruber, H. J.; Pastushenko, V. P.; Riener, C.; Trieb, M.; Knaus, H.-G.; Schindler, H.; Hinterdorfer, P. *Single Mol.* **2000**, *1*, 59.
(64) Tang, J.; Ebner, A.; Ilk, N.; Lichtblau, H.; Huber, C.; Zhu, R.; Pum, D.; Leitner, M.; Pastushenko, V.; Gruber, H. J.; Sleytr, U. B.; Hinterdorfer, P. *Langmuir* **2007**, *24*, 1324.
(65) Kudera, M.; Eschbaumer, C.; Gaub, H. E.; Schubert, U. S. *Adv. Funct. Mater.* **2003**, *13*, 615.
(66) Zapotoczny, S.; Auletta, T.; de Jong, M. R.; Schonherr, H.; Huskens, J.; van Veggel, F. C. J. M.; Reinhoudt, D. N.; Vancso, G. J. *Langmuir* **2002**, *18*, 6988.
(67) Xu, B.; Xiao, X.; Tao, N. J. *J. Am. Chem. Soc.* **2003**, *125*, 16164.
(68) Reddy, P.; Jang, S.-Y.; Segalman, R. A.; Majumdar, A. *Science* **2007**, *315*, 1568.
(69) VazquezH; SkoutaR; SchneebeliS; KamenetskaM; BreslowR; VenkataramanL; HybertsenM.S *Nat. Nano* **2012**, *7*, 663.
(70) Kim, C. M.; Bechhoefer, J. *J. Chem. Phys.* **2013**, *138*, 014707.
(71) Frei, M.; Aradhya, S. V.; Koentopp, M.; Hybertsen, M. S.; Venkataraman, L. *Nano Lett.* **2011**, *11*, 1518.
(72) Cao, Y.; Er, K. S.; Parhar, R.; Li, H. B. *Chemphyschem* **2009**, *10*, 1450.
(73) Cao, Y.; Li, H. *Nat. Mater.* **2007**, *6*, 109.

(74) Galera-Prat, A.; Gómez-Sicilia, A.; Oberhauser, A. F.; Cieplak, M.; Carrión-Vázquez, M. *Curr. Opin. Struct. Biol.* **2010**, *20*, 63.
(75) Labeit, S.; Kolmerer, B. *Science* **1995**, *270*, 293.
(76) Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E. *Science* **1997**, *276*, 1109.
(77) Kellermayer, M. S. Z.; Smith, S. B.; Granzier, H. L.; Bustamante, C. *Science* **1997**, *276*, 1112.
(78) Lee, G.; Abdi, K.; Jiang, Y.; Michaely, P.; Bennett, V.; Marszalek, P. E. *Nature* **2006**, *440*, 246.
(79) Oberhauser, A. F.; Marszalek, P. E.; Erickson, H. P.; Fernandez, J. M. *Nature* **1998**, *393*, 181.
(80) Oberhauser, A. F.; Badilla-Fernandez, C.; Carrion-Vazquez, M.; Fernandez, J. M. *J. Mol. Biol.* **2002**, *319*, 433.
(81) Borgia, A.; Williams, P. M.; Clarke, J. In *Annu. Rev. Biochem.*; Annual Reviews: Palo Alto, 2008; Vol. 77, p 101.
(82) Ng, S. P.; Billings, K. S.; Ohashi, T.; Allen, M. D.; Best, R. B.; Randles, L. G.; Erickson, H. P.; Clarke, J. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 9633.
(83) Jollymore, A.; Lethias, C.; Peng, Q.; Cao, Y.; Li, H. *J. Mol. Biol.* **2009**, *385*, 1277.
(84) Cao, Y.; Lam, C.; Wang, M.; Li, H. *Angew. Chem. Int. Ed.* **2006**, *45*, 642.
(85) Carrion-Vazquez, M.; Li, H.; Lu, H.; Marszalek, P. E.; Oberhauser, A. F.; Fernandez, J. M. *Nat. Struct. Biol.* **2003**, *10*, 738.
(86) Randles, L. G.; Rounsevell, R. W.; Clarke, J. *Biophys. J.* **2007**, *92*, 571.
(87) Brockwell, D. J.; Beddard, G. S.; Paci, E.; West, D. K.; Olmsted, P. D.; Smith, D. A.; Radford, S. E. *Biophys. J.* **2005**, *89*, 506.
(88) Fowler, S. B.; Best, R. B.; Toca Herrera, J. L.; Rutherford, T. J.; Steward, A.; Paci, E.; Karplus, M.; Clarke, J. *J. Mol. Biol.* **2002**, *322*, 841.
(89) Chyan, C.-L.; Lin, F.-C.; Peng, H.; Yuan, J.-M.; Chang, C.-H.; Lin, S.-H.; Yang, G. *Biophys. J.* **2004**, *87*, 3995.
(90) Lenne, P. F.; Raae, A. J.; Altmann, S. M.; Saraste, M.; Hörber, J. K. H. *FEBS Letters* **2000**, *476*, 124.
(91) Hoffmann, T.; Dougan, L. *Chemical Society Reviews* **2012**, *41*, 4781.
(92) Gosline, J.; Lillie, M.; Carrington, E.; Guerette, P.; Ortlepp, C.; Savage, K. *Philos. T. Roy. Soc. B.* **2002**, *357*, 121.
(93) Valbuena, A.; Oroz, J.; Hervás, R.; Vera, A. M.; Rodríguez, D.; Menéndez, M.; Sulkowska, J. I.; Cieplak, M.; Carrión-Vázquez, M. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 13791.
(94) Aggarwal, V.; Kulothungan, S. R.; Balamurali, M. M.; Saranya, S. R.; Varadarajan, R.; Ainavarapu, S. R. K. *J. Biol. Chem.* **2011**, *286*, 28056.
(95) Fuson, K. L.; Ma, L.; Sutton, R. B.; Oberhauser, A. F. *Biophys. J.* **2009**, *96*, 1083.
(96) Lv, C.; Tan, C.; Qin, M.; Zou, D.; Cao, Y.; Wang, W. *Biophys. J.* **2012**,

102, 1944.
(97) Yang, G.; Cecconi, C.; Baase, W. A.; Vetter, I. R.; Breyer, W. A.; Haack, J. A.; Matthews, B. W.; Dahlquist, F. W.; Bustamante, C. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 139.
(98) Dietz, H.; Bertz, M.; Schlierf, M.; Berkemeier, F.; Bornschlogl, T.; Junker, J. P.; Rief, M. *Nat. Protoc.* **2006**, *1*, 80.
(99) Dietz, H.; Berkemeier, F.; Bertz, M.; Rief, M. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12724.
(100) Brockwell, D. J.; Paci, E.; Zinober, R. C.; Beddard, G. S.; Olmsted, P. D.; Smith, D. A.; Perham, R. N.; Radford, S. E. *Nat. Struct. Mol. Biol.* **2003**, *10*, 731.
(101) Lee, W.; Zeng, X. C.; Rotolo, K.; Yang, M.; Schofield, C. J.; Bennett, V.; Yang, W. T.; Marszalek, P. E. *Biophys. J.* **2012**, *102*, 1118.
(102) Jagannathan, B.; Elms, P. J.; Bustamante, C.; Marqusee, S. *Proc. Natl. Acad. Sci. U. S. A.* **2012**.
(103) Li, Yongnan D.; Lamour, G.; Gsponer, J.; Zheng, P.; Li, H. *Biophys. J.* **2012**, *103*, 2361.
(104) Wiita, A. P.; Perez-Jimenez, R.; Walther, K. A.; Grater, F.; Berne, B. J.; Holmgren, A.; Sanchez-Ruiz, J. M.; Fernandez, J. M. *Nature* **2007**, *450*, 124.
(105) Wiita, A. P.; Ainavarapu, S. R. K.; Huang, H. H.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 7222.
(106) Garcia-Manyes, S.; Liang, J.; Szoszkiewicz, R.; Kuo, T. L.; Fernandez, J. M. *Nat. Chem.* **2009**, *1*, 236.
(107) Liang, J.; Fernandez, J. M. *ACS Nano* **2009**, *3*, 1628.
(108) Ainavarapu, S. R. K.; Bruji, J.; Huang, H. H.; Wiita, A. P.; Lu, H.; Li, L.; Walther, K. A.; Carrion-Vazquez, M.; Li, H.; Fernandez, J. M. *Biophysical journal* **2007**, *92*, 225.
(109) Zhang, X.; Rico, F.; Xu, A.; Moy, V. In *Handbook of Single-Molecule Biophysics*; Hinterdorfer, P., Oijen, A., Eds.; Springer US: 2009, p 555.
(110) Cao, Y.; Balamurali, M. M.; Sharma, D.; Li, H. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 15677.
(111) Garcia-Manyes, S.; Dougan, L.; Fernández, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 10540.
(112) Zheng, P.; Cao, Y.; Bu, T. J.; Straus, S. K.; Li, H. B. *Biophys. J.* **2011**, *100*, 1534.
(113) Aioanei, D.; Lv, S.; Tessari, I.; Rampioni, A.; Bubacco, L.; Li, H.; Samorì, B.; Brucale, M. *Angew. Chem. Int. Ed.* **2011**, *50*, 4394.
(114) Aioanei, D.; Brucale, M.; Tessari, I.; Bubacco, L.; Samori, B. *Biophys. J.* **2012**, *102*, 342.
(115) Ma, L.; Xu, M.; Oberhauser, A. F. *J. Biol. Chem.* **2010**, *285*, 38438.
(116) Junker, J. P.; Ziegler, F.; Rief, M. *Science* **2009**, *323*, 633.
(117) Bertz, M.; Rief, M. *J. Mol. Biol.* **2009**, *393*, 1097.
(118) Han, X.; Qin, M.; Pan, H.; Cao, Y.; Wang, W. *Langmuir* **2012**, *28*, 10020.

(119) Li, H. *Organic & Biomolecular Chemistry* **2007**, *5*, 3399.
(120) Lu, W.; Negi, S. S.; Oberhauser, A. F.; Braun, W. *Proteins* **2012**, *80*, 1308.
(121) Forman, J. R.; Yew, Z. T.; Qamar, S.; Sandford, R. N.; Paci, E.; Clarke, J. *Structure (London, England : 1993)* **2009**, *17*, 1582.
(122) Zhuang, S.; Peng, Q.; Cao, Y.; Li, H. *J. Mol. Biol.* **2009**, *390*, 820.
(123) Borgia, A.; Steward, A.; Clarke, J. *Angew. Chem. Int. Ed.* **2008**, *47*, 6900.
(124) Crampton, N.; Brockwell, D. J. *Curr. Opin. Struct. Biol.* **2010**, *20*, 508.
(125) Sadler, D. P.; Petrik, E.; Taniguchi, Y.; Pullen, J. R.; Kawakami, M.; Radford, S. E.; Brockwell, D. J. *J. Mol. Biol.* **2009**, *393*, 237.
(126) Lu, H.; Schulten, K. *Proteins* **1999**, *35*, 453.
(127) Hummer, G.; Szabo, A. *Biophys. J.* **2003**, *85*, 5.
(128) Dudko, O. K.; Hummer, G.; Szabo, A. *Phys. Rev. Lett.* **2006**, *96*, 108101.
(129) Hsin, J.; Strümpfer, J.; Lee, E. H.; Schulten, K. *Annu.Rev. Biophys.* **2011**, *40*, 187.
(130) Marszalek, P. E.; Lu, H.; Li, H.; Carrion-Vazquez, M.; Oberhauser, A. F.; Schulten, K.; Fernandez, J. M. *Nature* **1999**, *402*, 100.
(131) Sotomayor, M.; Schulten, K. *Science* **2007**, *316*, 1144.
(132) Guzmán, D. L.; Randall, A.; Baldi, P.; Guan, Z. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1989.
(133) Makarov, D. E. *Biophys. J.* **2009**, *96*, 2160.
(134) Aioanei, D.; Samorì, B.; Brucale, M. *Phys. Rev. E* **2009**, *80*, 061916.
(135) He, C.; Genchev, G. Z.; Lu, H.; Li, H. *Journal of the American Chemical Society* **2012**, *134*, 10428.
(136) Lupton, E. M.; Achenbach, F.; Weis, J.; Bräuchle, C.; Frank, I. *Physical Review B* **2007**, *76*, 125420.
(137) Christian, A. B.; Daniel, J. M. *Reports on Progress in Physics* **2011**, *74*, 086601.
(138) Frederix, P. L. T. M.; Bosshart, P. D.; Engel, A. *Biophys. J.* **2009**, *96*, 329.
(139) Rogers, J. M.; Steward, A.; Clarke, J. *J. Am. Chem. Soc.* **2013**.
(140) Cheng, S.; Cetinkaya, M.; Gräter, F. *Biophys. J.* **2010**, *99*, 3863.
(141) Hervás, R.; Galera-Prat, A.; Gómez-Sicilia, À.; Losada-Urzáiz, F.; Carmen Fernández, M.; Fernández-Bravo, D.; Santana, E.; Barrio-García, C.; Melero, C.; Carrión-Vázquez, M. In *Single-molecule Studies of Proteins*; Oberhauser, A. F., Ed.; Springer New York: 2013; Vol. 2, p 1.
(142) Rico, F.; Su, C.; Scheuring, S. *Nano Letters* **2011**, *11*, 3983.
(143) Ho, D.; Zimmermann, J. L.; Dehmelt, F. A.; Steinbach, U.; Erdmann, M.; Severin, P.; Falter, K.; Gaub, H. E. *Biophys. J.* **2009**, *97*, 3158.
(144) Janke, M.; Rudzevich, Y.; Molokanova, O.; Metzroth, T.; Mey, I.; Diezemann, G.; Marszalek, P. E.; Gauss, J.; Bohmer, V.; Janshoff, A. *Nat. Nano* **2009**, *4*, 225.
(145) Loksztejn, A.; Scholl, Z.; Marszalek, P. E. *Chem. Commun.* **2012**, *48*,

11727.
(146) Marszalek, P. E.; Dufrene, Y. F. *Chem. Soc. Rev.* **2012**, *41*, 3523.
(147) Brantley, J. N.; Wiggins, K. M.; Bielawski, C. W. *Science* **2011**, *333*, 1606.
(148) Davis, D. A.; Hamilton, A.; Yang, J.; Cremar, L. D.; Van Gough, D.; Potisek, S. L.; Ong, M. T.; Braun, P. V.; Martinez, T. J.; White, S. R.; Moore, J. S.; Sottos, N. R. *Nature* **2009**, *459*, 68.
(149) Wiggins, K. M.; Bielawski, C. W. *Angew. Chem. Int. Ed.* **2012**, *51*, 1640.
(150) Kryger, M. J.; Munaretto, A. M.; Moore, J. S. *J. Am. Chem. Soc.* **2011**, *133*, 18992.
(151) Wildling, L.; Rankl, C.; Haselgrübler, T.; Gruber, H. J.; Holy, M.; Newman, A. H.; Zou, M.-F.; Zhu, R.; Freissmuth, M.; Sitte, H. H.; Hinterdorfer, P. *J. Biol. Chem.* **2012**, *287*, 105.
(152) Franz, C. M.; Taubenberger, A.; Puech, P.-H.; Muller, D. J. *Sci. STKE* **2007**, *2007*, pl5.
(153) Puntheeranurak, T.; Wildling, L.; Gruber, H. J.; Kinne, R. K. H.; Hinterdorfer, P. *Journal of Cell Science* **2006**, *119*, 2960.
(154) Holm, R. H.; Kennepohl, P.; Solomon, E. I. *Chem. Rev.* **1996**, *96*, 2239.
(155) Bertini, I.; Gray, H. B.; Lippard, S. J.; Valentine, J. S. *Bioinorganic Chemistry*; University Science Books: Mill Valley, California, 1994.
(156) Lippard, S. J.; Berg, J. M. *Principles of bioinorganic chemistry*; University Science Books: Mill Valley, California, 1994.
(157) Shimazaki, Y.; Takani, M.; Yamauchi, O. *Dalton Trans.* **2009**, 7854.
(158) Xiao, Z. G.; Wedd, A. G. *Nat. Prod. Rep.* **2010**, *27*, 768.
(159) Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K. *Annu. Rev. Biochem.* **2005**, *74*, 247.
(160) Hagelueken, G.; Wiehlmann, L.; Adams, T. M.; Kolmar, H.; Heinz, D. W.; Tummler, B.; Schubert, W. D. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 12276.
(161) Lovenber.W; Sobel, B. E. *Proc. Natl. Acad. Sci. U. S. A.* **1965**, *54*, 193.
(162) Blake, P. R.; Park, J. B.; Bryant, F. O.; Aono, S.; Magnuson, J. K.; Eccleston, E.; Howard, J. B.; Summers, M. F.; Adams, M. W. *Biochemistry* **1991**, *30*, 10885.
(163) Day, M. W.; Hsu, B. T.; Joshuator, L.; Park, J. B.; Zhou, Z. H.; Adams, M. W. W.; Rees, D. C. *Protein Sci.* **1992**, *1*, 1494.
(164) Eidsness, M. K.; Odell, S. E.; Kurtz, D. M.; Robson, R. L.; Scott, R. A. *Protein Eng.* **1992**, *5*, 367.
(165) Dauter, Z.; Wilson, K. S.; Sieker, L. C.; Moulis, J. M.; Meyer, J. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 8836.
(166) Eidsness, M. K.; Richie, K. A.; Burden, A. E.; Kurtz, D. M., Jr.; Scott, R. A. *Biochemistry* **1997**, *36*, 10406.
(167) Solomon, E. I.; Gorelsky, S. I.; Dey, A. *J. Comput. Chem.* **2006**, *27*, 1415.

(168) Solomon, E. I.; Hedman, B.; Hodgson, K. O.; Dey, A.; Szilagyi, R. K. *Coord. Chem. Rev.* **2005**, *249*, 97.
(169) Rose, K.; Shadle, S. E.; Eidsness, M. K.; Kurtz, D. M.; Scott, R. A.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **1998**, *120*, 10743.
(170) Lane, R. W.; Ibers, J. A.; Frankel, R. B.; Holm, R. H. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 2868.
(171) Cavagnero, S.; Zhou, Z. H.; Adams, M. W.; Chan, S. I. *Biochemistry* **1998**, *37*, 3377.
(172) Cavagnero, S.; Debe, D. A.; Zhou, Z. H.; Adams, M. W. W.; Chan, S. I. *Biochemistry* **1998**, *37*, 3369.
(173) Morleo, A.; Bonomi, F.; Iametti, S.; Huang, V. W.; Kurtz, D. M. *Biochemistry* **2010**, *49*, 6627.
(174) Bao, G.; Suresh, S. *Nat. Mater.* **2003**, *2*, 715.
(175) Tatham, A. S.; Shewry, P. R. *Trends Biochem. Sci.* **2000**, *25*, 567.
(176) Carrion-Vazquez, M.; Oberhauser, A.; Diez, H.; Hervas, R.; Oroz, J.; Fernandez, J.; Martinez-Martin, A. In *Advanced Techniques in Biophysics*; Arrondo, J., Alonso, A., Eds.; Springer: 2006, p 163.
(177) Puchner, E. M.; Gaub, H. E. *Curr. Opin. Struct. Biol.* **2009**, *19*, 605.
(178) Fisher, T. E.; Marszalek, P. E.; Fernandez, J. M. *Nat. Struct. Biol.* **2000**, *7*, 719.
(179) Steward, A.; Toca-Herrera, J. L.; Clarke, J. *Protein Sci.* **2002**, *11*, 2179.
(180) Brockwell, D. J.; Beddard, G. S.; Clarkson, J.; Zinober, R. C.; Blake, A. W.; Trinick, J.; Olmsted, P. D.; Smith, D. A.; Radford, S. E. *Biophys. J.* **2002**, *83*, 458.
(181) Bornschlogl, T.; Anstrom, D. M.; Mey, E.; Dzubiella, J.; Rief, M.; Forest, K. T. *Biophys. J.* **2009**, *96*, 1508.
(182) Zhang, N.; Palmer, A. F. *Biotechnol. Prog.* **2010**, *26*, 1481.
(183) Hermanson, G. T. *Bioconjugate Techniques, 2nd Edition*; Academic Press, 2008.
(184) Green, N. S.; Reisler, E.; Houk, K. N. *Protein Sci.* **2001**, *10*, 1293.
(185) Odian, G. *Principles of Polymerziation*; 4th ed.; Wiley-Interscience, 2004.
(186) Beinert, H.; Holm, R. H.; Munck, E. *Science* **1997**, *277*, 653.
(187) Sharma, D.; Perisic, O.; Peng, Q.; Cao, Y.; Lam, C.; Lu, H.; Li, H. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 9278.
(188) Kruger, D.; Rousseau, R.; Fuchs, H.; Marx, D. *Angew. Chem. Int. Ed.* **2003**, *42*, 2251.
(189) Beyer, M. K. *J. Chem. Phys.* **2000**, *112*, 7307.
(190) Beinert, H. *J. Biol. Inorg. Chem.* **2000**, *5*, 2.
(191) Kennepohl, P.; Solomon, E. I. *Inorg. Chem.* **2003**, *42*, 689.
(192) Zheng, P.; Cao, Y.; Li, H. *Langmuir* **2011**, *27*, 5713.
(193) Strop, P.; Mayo, S. L. *J. Am. Chem. Soc.* **1999**, *121*, 2341.

(194) Xiao, Z. G.; Lavery, M. J.; Ayhan, M.; Scrofani, S. D. B.; Wilce, M. C. J.; Guss, J. M.; Tregloan, P. A.; George, G. N.; Wedd, A. G. *J. Am. Chem. Soc.* **1998**, *120*, 4135.
(195) Valbuena, A.; Oroz, J.; Hervas, R.; Vera, A. M.; Rodriguez, D.; Menendez, M.; Sulkowska, J. I.; Cieplak, M.; Carrion-Vazquez, M. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 13791.
(196) Glaser, T.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *Acc. Chem. Res.* **2000**, *33*, 859.
(197) Aktah, D.; Frank, I. *J. Am. Chem. Soc.* **2002**, *124*, 3402.
(198) Reddi, A. R.; Gibney, B. R. *Biochemistry* **2007**, *46*, 3745.
(199) Hickenboth, C. R.; Moore, J. S.; White, S. R.; Sottos, N. R.; Baudry, J.; Wilson, S. R. *Nature* **2007**, *446*, 423.
(200) Schmidt, S. W.; Filippov, P.; Kersch, A.; Beyer, M. K.; Clausen-Schaumann, H. *ACS Nano* **2012**, *6*, 1314.
(201) Zheng, P.; Li, H. *J. Am. Chem. Soc.* **2011**, *133*, 6791.
(202) Zheng, P.; Takayama, S.-i. J.; Mauk, A. G.; Li, H. *J. Am. Chem. Soc.* **2012**, *134*, 4124.
(203) Winkler, J. R.; WittungStafshede, P.; Leckner, J.; Malmstrom, B. G.; Gray, H. B. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 4246.
(204) Wilson, C. J.; Apiyo, D.; Wittung-Stafshede, P. *Q. Rev. Biophys.* **2004**, *37*, 285.
(205) Dukes, G. R.; Holm, R. H. *J. Am. Chem. Soc.* **1975**, *97*, 528.
(206) Cavagnero, S.; Debe, D. A.; Zhou, Z. H.; Adams, M. W. W.; Chan, S. I. *Biochemistry* **1998**, *37*, 3369.
(207) Shimazaki, Y.; Takani, M.; Yamauchi, O. *Dalton. Trans.* **2009**, 7854.
(208) Bau, R.; Rees, D. C.; Kurtz Jr, D. M.; Scott, R. A.; Huang, H.; Adams, M. W. W.; Eidsness, M. K. *J. Biol. Inorg. Chem.* **1998**, *3*, 484.
(209) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657.
(210) Blanco, F. J.; Rivas, G.; Serrano, L. *Nat. Struct. Biol.* **1994**, *1*, 584.
(211) Day, M. W.; Hsu, B. T.; Joshuator, L.; Park, J. B.; Zhou, Z. H.; Adams, M. W. W.; Rees, D. C. *Protein Sci.* **1992**, *1*, 1494.
(212) Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3694.
(213) Peng, Q.; Li, H. B. *J. Am. Chem. Soc.* **2009**, *131*, 14050.
(214) Stahl, S. W.; Puchner, E. M.; Alexandrovich, A.; Gaute, M.; Gaub, H. E. *Biophy. J.* **2011**, *101*, 1978.
(215) Dudev, T.; Lim, C. *Annu. Rev. Biophys.* **2008**, *37*, 97.
(216) Wittung-Stafshede, P. *Acc. Chem. Res.* **2002**, *35*, 201.
(217) Arnesano, F.; Banci, L.; Piccioli, M. *Q. Rev. Biophys.* **2005**, *38*, 167.
(218) Bertini, I.; Cowan, J. A.; Luchinat, C.; Natarajan, K.; Piccioli, M. *Biochemistry* **1997**, *36*, 9332.

(219) Dietz, H.; Rief, M. *Physical Review Letters* **2008**, *100*, 098101.
(220) Eyal, E.; Dutta, A.; Bahar, I. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 426.
(221) Bensimon, D. *Structure* **1996**, *4*, 885.
(222) Owen, A. J.; Ward, I. M. *J. Macro. Sci. B* **1973**, *7*, 417.
(223) Gavara, N.; Chadwick, R. S. *PLoS ONE* **2009**, *4*, e4877.
(224) Dey, A.; Okamura, T.; Ueyama, N.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **2005**, *127*, 12046.
(225) Gámiz-Hernández, A. P.; Galstyan, A. S.; Knapp, E.-W. *J. Chem. Theory Comput.* **2009**, *5*, 2898.
(226) Lu, Y.; Yeung, N.; Sieracki, N.; Marshall, N. M. *Nature* **2009**, *460*, 855.
(227) Eidsness, M. K.; Burden, A. E.; Richie, K. A.; Kurtz, D. M., Jr.; Scott, R. A.; Smith, E. T.; Ichiye, T.; Beard, B.; Min, T.; Kang, C. *Biochemistry* **1999**, *38*, 14803.
(228) Lin, I. J.; Gebel, E. B.; Machonkin, T. E.; Westler, W. M.; Markley, J. L. *J. Am. Chem. Soc.* **2003**, *125*, 1464.
(229) Lin, I. J.; Gebel, E. B.; Machonkin, T. E.; Westler, W. M.; Markley, J. L. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 14581.
(230) Sun, N.; Dey, A.; Xiao, Z. G.; Wedd, A. G.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *J. Am. Chem. Soc.* **2010**, *132*, 12639.
(231) Kienberger, F.; Ebner, A.; Gruber, H. J.; Hinterdorfer, P. *Acc. Chem. Res.* **2005**, *39*, 29.
(232) Bau, R.; Rees, D. C.; Kurtz Jr, D. M.; Scott, R. A.; Huang, H.; Adams, M. W. W.; Eidsness, M. K. *J. Biol. Inorg. Chem.* **1998**, *3*, 484.
(233) Kurihara, K.; Tanaka, I.; Chatake, T.; Adams, M. W.; Jenney, F. E., Jr.; Moiseeva, N.; Bau, R.; Niimura, N. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 11215.
(234) Day, M. W.; Hsu, B. T.; Joshua-Tor, L.; Park, J. B.; Zhou, Z. H.; Adams, M. W.; Rees, D. C. *Protein Sci* **1992**, *1*, 1494.
(235) Yanagisawa, S.; Banfield, M. J.; Dennison, C. *Biochemistry* **2006**, *45*, 8812.
(236) Marshall, N. M.; Garner, D. K.; Wilson, T. D.; Gao, Y.-G.; Robinson, H.; Nilges, M. J.; Lu, Y. *Nature* **2009**, *462*, 113.
(237) Li, H. B.; Carrion-Vazquez, M.; Oberhauser, A. F.; Marszalek, P. E.; Fernandez, J. M. *Nature Structural Biology* **2000**, *7*, 1117.
(238) Christensen, H. E. M.; Hammerstad-Pedersen, J. M.; Holm, A.; Iversen, G.; Jensen, M. H.; Ulstrup, J. *European Journal of Biochemistry* **1994**, *224*, 97.
(239) Cao, Y.; Lam, C.; Wang, M. J.; Li, H. B. *Angew. Chem. Int. Edit.* **2006**, *45*, 642.
(240) Dietz, H.; Bertz, M.; Schlierf, M.; Berkemeier, F.; Bornschlogl, T.; Junker, J. P.; Rief, M. *Nat Protoc* **2006**, *1*, 80.
(241) Zheng, P.; Cao, Y.; Li, H. *Langmuir* **2011**, *10*, 5713.

(242) Venkateswara Rao, P.; Holm, R. H. *Chem. Rev.* **2004**, *104*, 527.
(243) Que, L., Jr.; Bobrik, M. A.; Ibers, J. A.; Holm, R. H. *J. Am. Chem. Soc.* **1974**, *96*, 4168.
(244) Bobrik, M. A.; Que, L., Jr.; Holm, R. H. *J. Am. Chem. Soc.* **1974**, *96*, 285.
(245) Brockwell, D. J. *Curr. Nanosci.* **2007**, *3*, 3.
(246) Job, R. C.; Bruice, T. C. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 2478.
(247) Lo, W.; Scott, T. A.; Zhang, P.; Ling, C. C.; Holm, R. H. *J. Inorg. Biochem.* **2011**, *105*, 497.
(248) Foster, M. W.; Bian, S. M.; Surerus, K. K.; Cowan, J. A. *J. Biol. Inorg. Chem.* **2001**, *6*, 266.
(249) Zheng, P.; Li, H. *Biophys. J.* **2011**, *101*, 1467.

Appendix A: Protein engineering

A1. Protein sequence and corresponding cDNA

A1.1 Wild-type *pf*RD (without endogenous *KpnI* site)

Protein: A K W V C K I C G Y I Y D E D A G D P D N G I S P G T K F E E L P D D W V C P I C G A P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCAAAATCTGCGGTTACATCTACGACGAAGACGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACTGCCGGACGACTGGGTTTGTCCGATCTGCGGTGCTCCGAAATCCGAATTCGAAAAACTGGAAGAC

A1.2 Cys-*pf*RD-GB1-Cys

Protein: C A K W V C K I C G Y I Y D E D A G D P D N G I S P G T K F E E L P D D W V C P I C G A P K S E F E K L E D R S M D T Y K L I L N G K T L K G E T T T E A V D A A T A E K V F K Q Y A N D N G V D G E W T Y D D A T K T F T V T E C

cDNA:TGCGCTAAATGGGTTTGCAAAATCTGCGGTTACATCTACGACGAAGACGCTGGTGACCCGGACAACGGTATCT**CCCCGG**GTACCAAATTCGAAGAACTGCCGGACGACTGGGTTTGTCCGATCTGCGGTGCTCCGAAATCCGAAT

TCGAAAAACTGGAAGACAGATCCGACACCTACAAACTGATCCTGAACGGT

AAAACCCTGAAAGGTGAAACCACCACCGAAGCTGTAGACGCTGCTACTGC

AGAAAAAGTTTTCAAACAGTACGCTAACGACAACGGTGTCGACGGTGAAT

GGACCTACGACGACGCTACCAAAACCTTCACGGTTACCGAATGC

A1.3 Apo-*pf*RD

Protein: A K W V K K I T G Y I Y D E D A G D P D N G I S P G T K F E E L P

D D W V A P I T G A P K S E F E K L E D

cDNA:GCTAAATGGGTTAAGAAAATCACCGGTTACATCTACGACGAAGA

CGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACT

GCCGGACGACTGGGTTGCGCCGATCACCGGTGCTCCGAAATCCGAATTCG

AAAAACTGGAAGAC

A1.4 C5,8H *pf*RD

Protein: A K W V H K I H G Y I Y D E D A G D P D N G I S P G T K F E E L P

D D W V C P I C G A P K S E F E K L E D

cDNA:GCTAAATGGGTTCACAAAATCCACGGTTACATCTACGACGAAGA

CGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACT

GCCGGACGACTGGGTTTGTCCGATCTGCGGTGCTCCGAAATCCGAATTCGA

AAAACTGGAAGAC

A1.5 C38H *pf*RD

Protein: A K W V C K I C G Y I Y D E D A G D P D N G I S P G T K F E E L P D D W V H P I C G A P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCAAAATCTGCGGTTACATCTACGACGAAGACGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACTGCCGGACGACTGGGTTCACCCGATCTGCGGTGCTCCGAAATCCGAATTCGAAAAACTGGAAGAC

A1.6 C41H *pf*RD

Protein:A K W V C K I C G Y I Y D E D A G D P D N G I S P G T K F E E L P D D W V C P I H G A P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCAAAATCTGCGGTTACATCTACGACGAAGACGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACTGCCGGACGACTGGGTTTGTCCGATCCACGGTGCTCCGAAATCCGAATTCGAAAAACTGGAAGAC

A1.7 C38,41H *pf*RD

Protein:A K W V C K I C G Y I Y D E D A G D P D N G I S P G T K F E E L P D D W V H P I H G A P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCAAAATCTGCGGTTACATCTACGACGAAGA
CGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACT
GCCGGACGACTGGGTTCACCCGATCCACGGTGCTCCGAAATCCGAATTCGA
AAAACTGGAAGAC

A1.8 *cp*RD

Protein:M K K Y T C T V C G Y I Y N P E D G D P D N G V N P G T D F K D
I P D D W V C P L C G V G K D Q F E E V E E

cDNA:ATGAAGAAATACACCTGCACCGTTTGCGGTTACATCTACAACCC
GGAAGACGGTGATCCGGACAACGGTGTTAACCCGGGCACCGACTTTAAAG
ATATCCCGGACGACCTGGGTTTGCCCGCTGTGCGGTGTTGGTAAAGACCAA
TTCGAAGAAGTTGAAGAA

A1.9 RD1β

Protein: A K W V C L G E W T Y D D A T K T F I V T E L G C G Y I Y D E D
A G D P D N G I S P G T K F E E L P D D W V C P I C G A P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCCTCGGGGAATGGACCTACGACGACGCTAC
CAAAACCTTCACGGTTACCGAACTCGGGTGCGGTTACATCTACGACGAAGA
CGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACT
GCCGGACGACTGGGTTTGTCCGATCTGCGGTGCTCCGAAATCCGAATTCGA

AAAACTGGAAGAC

A1.10 RD2β

Protein: A K W V C K I C G Y I Y D E D A G D P D N G I S P G T K F E E L P D D W V C L G E W T Y D D A T K T F I V T E L G C G A P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCAAAATGTGCGGTTACATCTACGACGAAGACGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACTGCCGGACGACTGGGTTTGTCTCGGGGAATGGACCTACGACGACGCTACCAAAACCTTCACGGTTACCGAACTCGGGTGCGGTGCTCCGAAATCCGAATTCGAAAAACTGGAAGAC

A1.11 RD1,49

Protein: C A K W V C K I C G Y I Y D E D A G D P D N G I S P G T K F E E L P D D W V C P I C G A P K S E F C K L E D

cDNA:TGCGCTAAATGGGTTTGCAAAATCTGCGGTTACATCTACGACGAAGACGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACTGCCGGACGACTGGGTTTGTCCGATCTGCGGTGCTCCGAAATCCGAATTCTGCAAACTGGAAGAC

A1.12 RD1,35

Protein: C A K W V C K I C G Y I Y D E D A G D P D N G I S P G T K F E E L P D C W V C P I C G A P K S E F E K L E D

cDNA:TGCGCTAAATGGGTTTGCAAAATCTGCGGTTACATCTACGACGAAGACGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACTGCCGGACTGCTGGGTTTGTCCGATCTGCGGTGCTCCGAAATCCGAATTCGAAAAACTGGAAGAC

A1.13 RD15,35

Protein: A K W V C K I C G Y I Y D E C A G D P D N G I S P G T K F E E L P D C W V C P I C G A P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCAAAATCTGCGGTTACATCTACGACGAATGCGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACTGCCGGACTGCTGGGTTTGTCCGATCTGCGGTGCTCCGAAATCCGAATTCGAAAAACTGGAAGAC

A1.14 RD15,49

Protein: A K W V C K I C G Y I Y D E C A G D P D N G I S P G T K F E E L P D D W V C P I C G A P K S E F C K L E D

cDNA:GCTAAATGGGTTTGCAAAATCTGCGGTTACATCTACGACGAATG

CGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACT
GCCGGACGACTGGGTTTGTCCGATCTGCGGTGCTCCGAAATCCGAATTCTG
CAAACTGGAAGAC

A1.15 RD6,40

Protein: A K W V C C I C G Y I Y D E D A G D P D N G I S P G T K F E E L P
D D W V C P C C G A P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCTGCATCTGCGGTTACATCTACGACGAAGA
CGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACT
GCCGGACGACTGGGTTTGTCCGTGCTGCGGTGCTCCGAAATCCGAATTCGA
AAAACTGGAAGAC

A1.16 I7P *pf*RD

Protein: A K W V C K P C G Y I Y D E D A G D P D N G I S P G T K F E E L P
D D W V C P I C G A P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCAAACCGTGCGGTTACATCTACGACGAAGA
CGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACT
GCCGGACGACTGGGTTTGTCCGATCTGCGGTGCTCCGAAATCCGAATTCGA
AAAACTGGAAGAC

A1.17 I7G *pf*RD

Protein: A K W V C K G C G Y I Y D E D A G D P D N G I S P G T K F E E L P D D W V C P I C G A P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCAAAGGTTGCGGTTACATCTACGACGAAGACGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACTGCCGGACGACTGGGTTTGTCCGATCTGCGGTGCTCCGAAATCCGAATTCGAAAAACTGGAAGAC

A1.18 A43P *pf*RD

Protein: A K W V C K I C G Y I Y D E D A G D P D N G I S P G T K F E E L P D D W V C P I C G P P K S E F E K L E D

cDNA:GCTAAATGGGTTTGCAAAATGTGCGGTTACATCTACGACGAAGACGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACTGCCGGACGACTGGGTTTGTCCGATCTGCGGTCCGCCGAAATCCGAATTCGAAAAACTGGAAGAC

A1.19 A43G *pf*RD

Protein: A K W V C K I C G Y I Y D E D A G D P D N G I S P G T K F E E L P D D W V C P I C G P G K S E F E K L E D

cDNA:GCTAAATGGGTTTGCAAAATGTGCGGTTACATCTACGACGAAGA

CGCTGGTGACCCGGACAACGGTATCTCCCCGGGTACCAAATTCGAAGAACT

GCCGGACGACTGGGTTTGTCCGATCTGCGGTGGTCCGAAATCCGAATTCGA

AAAACTGGAAGAC

A2 Engineering of chimera protein

A2.1 Construction of gene encoding chimera protein

To provide an unambiguous fingerprint for identification of single molecule stretching event, a well-characterized protein domain is usually fused into the target protein as a protein chimera for AFM studies. Here, we use multiple cloning techniques to construct the plasmid encoding protein chimera.

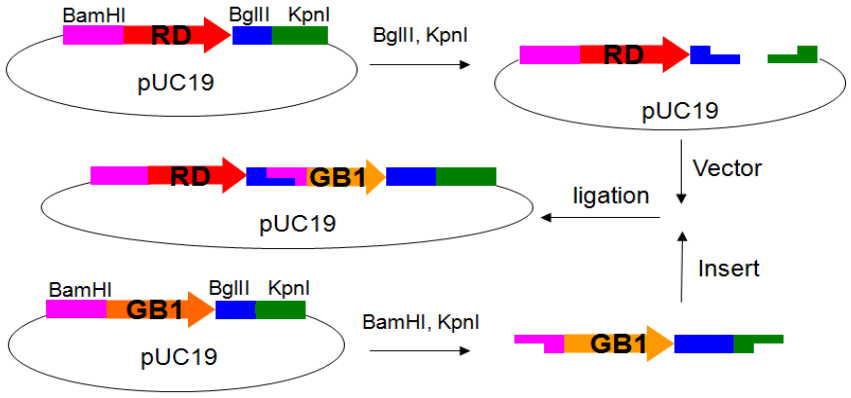

Figure A1 Construction of rubredoxin-GB1 protein chimera. The vector containing rubredoxin gene and insert containing the GB1 gene are obtained digestion. The two fragments are ligated together resulting in the plasmid pUC19-RD-GB1 including the desired genes.

Using PCR techniques, selected restriction sites are engineered into the gene of interest. The *BamH*I (G'GATCC), *Kpn*I (G'GTACC) and *Blg*II (A'GATCT) are the

three mostly used sites. It is noted that the digested *BamH*I and *Blg*II site share the same overhang (GATC) and thus can be ligated together. Furthermore, the new sequence (GGATCT) is not a restriction site and cannot be digested by any of the three enzymes.

The typical construction of a RD-GB1 protein chimera is shown in Figure A1. The pUC19 plasmid (2686 bps, New England Biolabs, Ipswich, MA) containing the gene of rubredoxin is digested with restriction enzyme *Kpn*I and *Blg*II and severs as a vector. The plasmid containing the gene of GB1 is then digested with enzyme *BamH*I and *Kpn*I as the insert. Consequently, the insert is ligated into the vector forming the gene encoding protein chimera RD-GB1. Finally, the gene is transferred into the expression vector pQE80L (4751 bps, Qiagen, Valencia, CA) which contains a His-tag for protein purification. All these digested fragments (the insert ~200 bps, the vector ~3000 bps and the ligated gene encoding protein chimera ~ 400 bps) are identified and purified by an agarose gel electrophoresis.

A2.2 Protein expression

E. coli expression system is used for protein expression. The plasmid containing the gene of interest is transformed into *E. coli* DH5α strain. It is then grown in 10 mL of 2.5% LB with 10 mg/mL ampicillin overnight at 37 ˚C and 225 rpm. Then the 10 mL culture is transferred into a larger quantity of ~500 mL of LB and grown at the same condition. When the OD of the culture is ~ 0.8 (usually takes ~3 hours), IPTG is added into the solution with a final concentration of ~ 1 mM to induce protein

expression. Incubated for ~4 hours, the cells are harvested by centrifuge at 5,000 g for 15 min and kept in storage at -80 ˚C.

The cell lysis is achieved by one freeze-thaw cycle at first. Then, a lysozyme digestion protocol is followed by adding lysozyme solution at a final concentration of ~ 1mg/mL. The solution is incubated for ~0.5 hour on ice and gradually becomes viscous due to the release of genomic DNA and RNA. The DNase I and RNase A are added to the solution at a final concentration ~ 0.005 mg/mL. It is incubated on ice and shaken until the solution become watery again. After centrifuging at 12,000 g for 1 hour, the supernatant is collected and purified by Co^{2+}-affinity chromatography using TALON resins. The protein is kept in Tris buffer in pH 7.4 at a concentration of ~ 2 mg/mL.

A2.3 Purification of pure Fe-form rubredoxin

Because Zn-RD and Fe-RD are co-expressed in *E. coli*, it is necessary to separate these two forms before the construction of poly-rubredoxin. The purification process of a pure Fe-form RD-GB1 is described as follows as an example: First, the protein chimera RD-GB1 is buffer exchanged into a 10 mM Tris at pH 8.5 and concentrated by centrifugal ultrafiltration (Amicon 3K MWCO filter, Millipore). The Fe and Zn form RDs are separated using Mono-Q 5/50GL anion exchange column. After loaded on the column, the protein mixture is eluted using a linear gradient elution (0-300 mM NaCl in 10 mM at pH 8.5) in AKTA FPLC system (GE Healthcare) at a flow rate of 2 mL/min. The Fe form RD-GB1 is eluted first at around 100 mM NaCl followed by Zn

form RD-GB1 around 150 mM (Fig. A2.A).

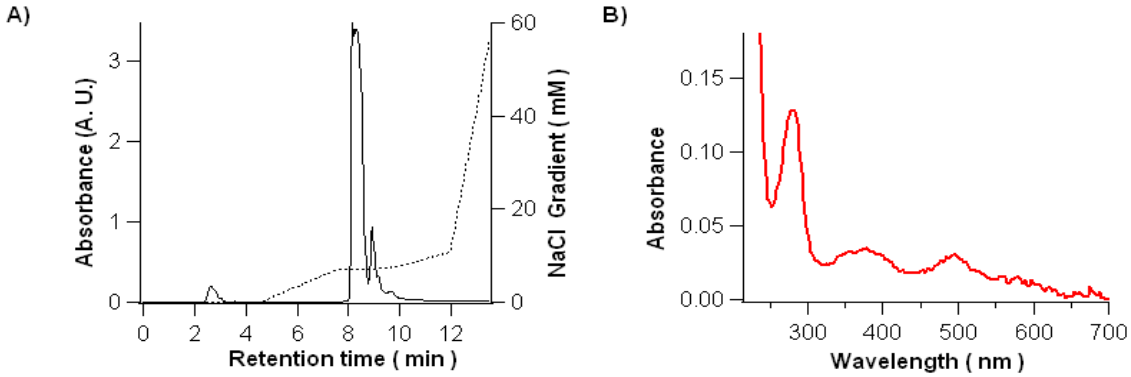

Figure A2 Purification of Fe-form rubredoxin. A) The elution profile of GB1-pfRD from anion-exchange chromatography using a Mono Q 5/50 column. Co2+-NTA purified protein was subjected to the column and eluted with a gradient of NaCl (dash line: 20% corresponds to 200 mM NaCl) in 50 mM Tris buffer (pH 8.5) at a flow rate of 2 mL/min. The solid line indicates the relative absorbance at 280 nm. B) UV-visible absorption spectra of purified protein *pf*RD-GB1.

The purity of Fe-form RD is confirmed by UV-Vis absorption spectroscopy (Fig. A2.B). Only Fe form rubredoxin shows characteristic absorption at visible light range with maxima at wavelength 494 nm as a red color solution. On comparison, the Zn form rubredoxin is a colorless solution. Thus, using the extinction coefficient of 9.22 mM^{-1} cm^{-1} at 494 nm unique for Fe form rubredoxin, the concentration of Fe-RD is determined. In addition, the sum of protein concentration can be obtained by measuring the absorption at 280 nm which is same from two types of rubredoxins. Consequently, purify of Fe-RD is estimated to be >90% after 2 to 3 times ion-exchange chromatography purification.

A3 UV-Vis spectrum of rubredoxin variants

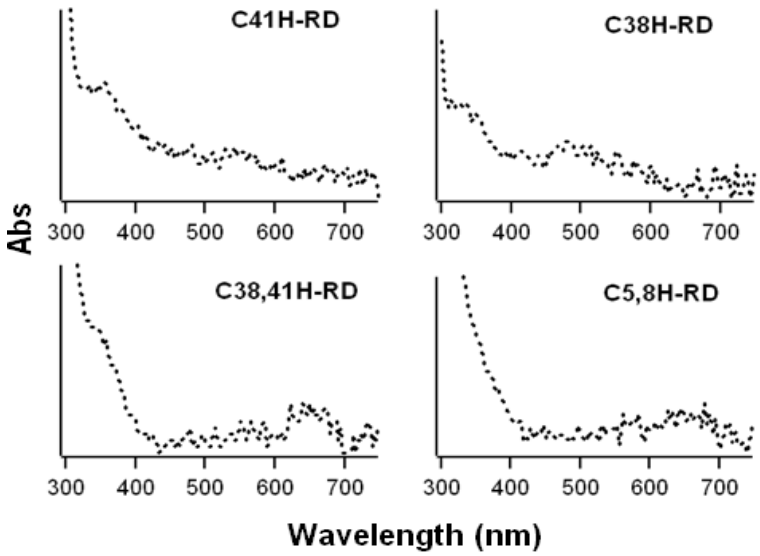

Figure A3 UV-Vis spectrum of cysteine to histidine rubredoxin variants.

The purified rubredoxin and variants are characterized by UV-Vis adsorption spectroscopy. The wild-type rubredoxin shows a characteristic UV-Vis spectrum with adsorption maxima at 280, 390 and 494 nm with a reddish color. The visible light absorption of rubredoxin arises from the ligand to metal charge transfer due to the ferric-thiolate bonds. Thus, measurement of the UV-Vis spectrum of rubredoxin is a simple approach to determine whether the mutation significantly distorts the structure of the FeS_4 center in rubredoxin variants. For example, the cysteine to histidine mutation in the FeS_4 center dramatically changes the property of the metal center (Fig. A4). If single cysteine residue is mutated to histidine in rubredoxin, the color of protein changes from red to orange. When two cysteine residues are mutated to histidine, it shows a blue color (Chapter 2).

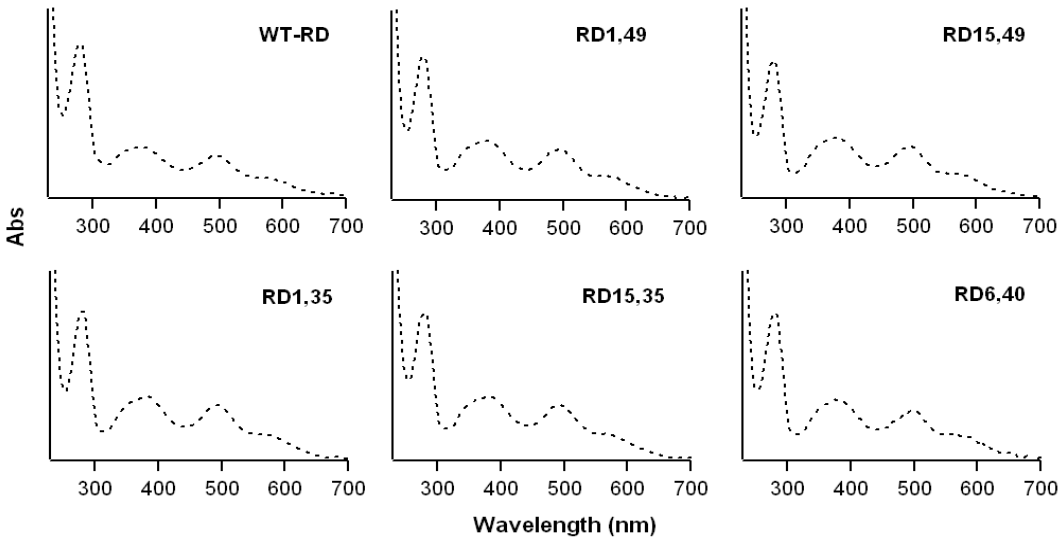

Figure A4 UV-Vis spectrum of rubredoxin cysteine variants for stretching the metal center from different directions. The name is labeled on the corresponding graphs.

On comparison, the cysteine-substituted rubredoxin variants engineered for mechanical anisotropy studies of rubredoxin show similar UV-Vis spectrum as that of WT-RD (Fig. A4). It demonstrates that the introduced cysteine residues does not affect the original FeS_4 center and function simply as handles in AFM studies (Chapter 6).

A4 Chemical coupling method for poly-rubredoxin

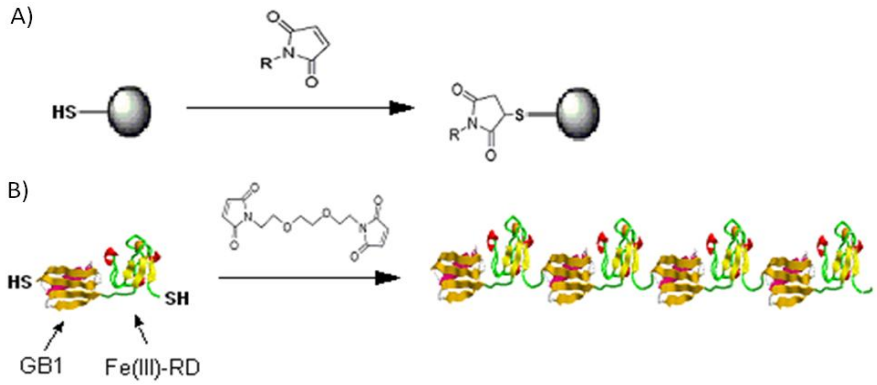

Figure A5 The construction of pure Fe-form RD-containing polyprotein based on

maleimide-thiol coupling chemistry. A) The sulfhydryl group of cysteine can readily react with maleimido group of $BM(PEO)_3$ to form a thioether bond. B) Thus, bi-functional cys-RD-GB1-cys can react with bi-functional maleimide compound $BM(PEO)_3$ to produce heteropolyprotein $(Fe\text{-}RD\text{-}GB1)_n$, where n denotes the degree of polymerization.

To obtain poly-rubredoxin $(RD\text{-}GB1)_n$ molecule, a chemical coupling reaction between maleimide and thiol groups is utilized (Fig A3). The two functional group forms a covalent thioether bond. As a result, the protein which carries two thiol groups at each ends is linked together by a bis-maleimide compound forming the desired polyprotein. Typically ~ 1 mL freshly purified Cys-RD-GB1-Cys protein solution with a concentration of 2 mg/mL is reacted with 40 μL of 10 mM concentrated stock solution of $BM(PEO)_3$ (1, 8-bis (maleimido)triethylene glycol, Molecular Biosciences) in Tris buffer at pH 7.4. After incubation at room temperature for ~2 hours, the resultant polyprotein solution can be used in AFM experiments directly.

Compared with classic polyprotein construction method at the DNA-level, this method does not have a precise control of the molecular weight of the polyprotein. Determined by SDS-PAGE, the polymerization degree of the resultant polyproteins is around 2-5. In addition, the orientations of each individual domain in the polyprotein chain here have three different arrangements: head-to-tail, head-to-head and tail-to-tail. However, these drawbacks do not affect the measurement of the mechanical properties of the protein of interest. Because applied force is vector and the measured mechanical stability should be the same regard least the arrangement in the linkage[99,192].