

Structural Determinants of Protein Dynamics, Cooperativity and Kinetic Stability
in Alpha-lytic Protease

by

F. Pinar Erciyas Bailey

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION

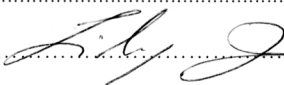
of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:


.....


.....


.....

Committee in Charge

UMI Number: 3398876

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3398876

Copyright 2010 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Copyright 2010

by

F. Pinar Erciyas Bailey

I dedicate this thesis

To my parents,

Tülin Bengüer and İsmail Erciyaş,

Who instilled in me an endless intellectual curiosity,

And taught me

To always do my best.

Acknowledgements

I would like to thank my thesis advisor David Agard, who has been instrumental in my intellectual development as a scientist. His sharp intellect and endless curiosity made working with him very enjoyable in a challenging way. Convincing him about anything takes utmost orientation to detail and he requires a thorough understanding of all the immediate and distant consequences. These expectations made me a better scientist than I could have been and I would like to thank him for that. In addition, Dave is an outstanding communicator and debater and expects his students to excel in these areas, as well. I learned a lot from him about how to communicate science in a clear and interesting way. Last, Dave puts great importance in establishing a friendly and collaborative environment in the lab, which makes our workplace very enjoyable. I appreciated this very much throughout my time in the lab.

Ken Dill, Lily Jan, Susan Marqusee, Jonathan Weissman, and Tack Kuntz deserve many thanks for not only having graciously served in my qualifying and thesis committees, but also for inspiring me with their outstanding science. I have learned a lot from them. Ken Dill was a deciding factor for me to choose UCSF for graduate school and while I didn't end up working in his lab, I did learn a lot from him scientifically. He is also an excellent communicator, who has been an inspiration to me.

Agard Lab has been a fantastic place to work with some of the smartest and most fun people I know. There are too many who helped me and shaped my scientific

development to list, so I would like to send all of them a big thank you. Strangely, it also consistently attracts outstanding cooks and what you bring in for lunch becomes a big deal. As a result, I became a decent cook myself and I would like to thank them for this unexpected benefit of graduate school, as well.

Brian Kelch, Sheila Jaswal, and Luke Rice were both great friends and mentors and all deserve many thanks. Brian started the projects constituting a substantial part of my thesis and has tremendous insight into questions regarding kinetic stability. Sheila has been always interested in my research, even though she and I actually never overlapped in the Agard Lab. She is a fine scientist and an inspiring person, whom I respect very much. Luke Rice is not only one of the smartest people I know, he has an outstanding ability and interest to teach developing scientists. He has an extremely deep knowledge in multiple vast areas of science. Not unlike a doctor making morning rounds, he would frequently walk around the lab and ask people what they were doing and what challenges they were facing. He supported me and many others by providing great guidance and motivating us with his interest in our research. He also has excellent culinary skills and will make an outstanding advisor.

I would like to thank Neema Salimi, Timo Street and Bosco Ho. I have exchanged many ideas with them in the lab about science. They are fine scientists and their input helped me improve as one, as well.

Alexandra Schnoes, Cynthia Fuhrmann Kelch, Mary Jane Budny, and Toral Surti are friends that all wore multiple hats for me. Alex and I were classmates, as well as roommates and her manners saved me from almost killing our landlady in many occasions. She handles tough situations extremely well and later has been influential in my job search, as well as being a supportive friend. Cynthia and I were roommates, as well as labmates. When it comes to α LP crystallography, she is the best mentor you can have. She has an infectious smile and an attitude to always make the best out of any situation. MJ and I, in the lab and as a roommate, had the best conversations. She is filled with a feeling of responsibility to help everybody around her, while having a great sense of all things fun and stylish. Our friendship stayed strong even though she left town many years ago and I can't wait to reunite one day. Toral Surti and I have been classmates and very close friends. She is the kindest person I know. She has a great aptitude to have a person feel listened. She is very giving and understanding, has a great sense of humor and was the best person who could have performed our marriage ceremony.

Chris Waddling and Michelle Moritz have been in the lab the entire period of my graduate school. Michelle and I shared a bay together for most of the time and she has been a great person to talk to, vent and laugh together. A roll of eyes accompanied by a sarcastic "Brother!" will forever be etched in my memory. Chris and I, not only got to be collaborating colleagues and are sharing a publication, we share many interests. He is a very positive and fun guy, so he is often the go-to person when I need a lift of spirits. Life in the Agard Lab wouldn't have been the same without these two.

Audrey Kanemoto, and Heidi and Gerhard Matzen are some of the many friends I made while swing dancing during graduate school. We have danced and performed together, but more importantly became great friends. They supported me with their joy, wisdom and often food throughout graduate school. My life wouldn't have been nearly as enjoyable without having known these people. Thank you!

My parents Tülin Bengüer and İsmail Erciyaş are the reason I am successful at anything, at all. From an early age, they encouraged me to try new things and follow my curiosity. My mother would try to answer my never-ending questions about anything and everything from science and math to legal matters and politics and would never get bored. When she didn't know, instead of making up an answer, she would show me how to find it out. My father's best advice to me was to tell people what I appreciate about them. He also showed me that if people say no, there may be better ways of asking. He was also very interested in my intellectual development and went as far as making an appointment with a well respected history professor when I was in elementary school, so I could have not only answers to my questions, but also witness the scholarly method. Together they taught me that it is important to ask questions instead of taking things at face value. They have instilled in me a self-confidence that graduate school couldn't even break. They were my biggest cheerleaders in everything I tried and taught me to always try my best. Their conviction in me let me grow into being who I am and their love and devotion can never be paid back. Thank you, mom and dad.

Last, but not least, my husband Greg Bailey is my best friend. The best aspect of being in San Francisco for graduate school was meeting him. He has been a pillar of my life, especially when things went wrong, and things went wrong a lot in graduate school. He has not been an unquestioning supporter, but went the extra mile to try to provide perspective. He has provided me with joy, which helped a lot when I couldn't see the light. Since June 2008, he has been a most loving and kind husband and I am so lucky and grateful to have him as my partner in life. I am looking forward to growing our family and spending the rest of our lives together.

Abstract

Structural information on nonnative states of proteins, including folding intermediates and folding and unfolding transition states is crucial for understanding folding and unfolding mechanisms. Kinetically stable proteins such as α -Lytic protease (α LP) combine a very high barrier to unfolding with extraordinary unfolding cooperativity to uncouple their native state from the unfolded states. This unusual energetic landscape results in a remarkable resistance to proteolytic destruction and thus is crucial for α LP's biological function. The uncoupling of native and unfolded states makes the differences between the native state and the transition state most relevant for protein function and thus, the height of the unfolding barrier rather than the thermodynamic stability is the relevant metric to investigate.

In order to characterize the structural determinants of protein unfolding in α LP, I used salt bridges to probe for structural rearrangements and cooperative contributions in the unfolding transition. From these studies, I have identified protein regions that are frequently unfolded in the transition state, as well as a mechanism to couple cooperative contributions from distant regions in the protein. This analysis led to an energetic dissection of cooperativity allowing for a quantitative assessment of this property for the first time. Also, two other findings, support the previous findings of an extremely rigid α LP native state: The extremely low native state pK_a values for the salt bridge carboxylates point to a significantly dampened native state dynamics strengthening the salt bridges. In addition, experiments investigating the role of the disulfide bridges in α LP unfolding found an extreme insensitivity of the unfolding barrier to reducing conditions,

suggesting that these disulfide bridges are also protected by rigid native state dynamics and that protein unfolding is all or nothing.

α LP folding transition involves an even higher barrier than that for the unfolding transition. Uncatalyzed, a molten-globule like intermediate converts extremely slowly to mature protease. Understanding this transition with structural detail has been the motivation to develop a structure determination method for nonnative states of proteins. Nonnative proteins provide extreme challenges for structure determination; a multiplicity of structural states and aggregation at high enough concentrations to name a couple. The structure determination method I developed involves the use of cross-linkers to identify distance constraints in the protein structure via mass spectrometry. While the structure determination method was not sensitive enough to provide the level of structural detail I aimed to obtain, findings facilitated by precursor ion scanning during the development of this method are crucial to improve the detection limit for future studies.

Last, I studied the substrate length dependent k_{cat} effect for α LP substrate catalysis. Crystal structures of α LP with varying lengths of boronic acid inhibitors that are thought to mimic one of the tetrahedral intermediates in catalysis were compared in terms of active site protein dynamics. Through the analysis of anisotropic B-factors, a correlation was found with increased thermal motion in catalytic atoms and decreased thermal motion in the substrate binding pocket, and increasing inhibitor length. The latter with the higher overall disorder may provide an explanation for the k_{cat} effect with an increased order in catalytic transition states in the presence of longer substrates.

Table of Contents

Chapter 1. Introduction	1
Cooperativity and Kinetic Stability	1
Cooperativity and kinetic stability in disease context	3
Alpha-lytic Protease (α LP) and its folding landscape	4
The native-like TS and current models of TS	8
Cooperative unfolding of α LP	10
Characterization of the Intermediate state	12
α LP catalytic mechanisms	15
Specific aims	17
Chapter 2. Quantifying protein unfolding cooperativity with acid sensitive probes: inter-domain salt bridge contributions to unfolding cooperativity are combined efficiently in α-Lytic Protease	18
Preface	19
Abstract	20
Introduction	22
Results	27
pH dependent unfolding profiles of single salt bridge mutants	27
pH dependent unfolding profiles of double salt bridge mutants	31
Crystal structures of SB2+3 and SB1+2	33
Cooperativity through non-local interactions between salt bridges	36
Number of salt bridges	38
Discussion	40
A strategy for broad range pH stability	41
Cooperative contributions of salt bridges	43
Structure in the TS ensemble and the cooperative unfolding transition	48
Extraordinarily low pK_a 's linked to high protection factors	50
Materials and Methods	53
Chapter 3. Investigating the role of substrate wall dynamics in αLP k_{cat} effect	62
Preface	63
Introduction	64
Materials and Methods	70
Results and Discussion	72
Conclusions	82
Chapter 4. A method to characterize nonnative states of proteins like αLP Int using cross-linkers	84
Preface	85
Introduction	86

Results	95
Cross-linker activation	97
Labeling primary amines: lysines and N-terminus	99
Eliminating excess cross-linker	101
Concentration and buffer exchange	101
Dilution into photoactivation buffer	101
Photoactivation	102
Purification of monomers	103
Proteolytic Digest	104
Purification cross-linker labeled peptides	105
Reducing disulfides and carboxymethylate	106
LC-MS with precursor ion scanning	107
Conclusions	112
Materials and Methods	113
<i>Chapter 5. The role of disulfide bridges in αLP folding and unfolding</i>	<i>121</i>
Introduction	122
Results	125
α LP unfolding by loss of activity with GSH	125
α LP unfolding by Trp fluorescence	126
Uncatalyzed α LP refolding with GSH	128
Discussion	129
Materials and Methods	132
<i>Conclusions</i>	<i>134</i>
Role of the domain-domain interface in TS	135
Extremely low native state pK _a 's	136
Further characterization unfolding barrier using salt bridge mutants	137
Recapturing stability at the salt bridge 3 site	139
Further structural elements are thought to contribute to cooperativity	140
A method for structural characterization of nonnative protein species motivated by α LP Int	141
Uncatalyzed folding of α LP under reducing conditions	143
α LP unfolding rates are unaffected by reducing conditions	144
References	145

List of Tables

<i>Table 2.1 Data collection and refinement statistics.</i>	35
<i>Table 3.1 Kinetic parameters for substrates of varying chain length (αLP)</i>	67
<i>Table 4.1 Recovery of insulin peptides in mass spectra. CM: Number of carboxymethyls after reducing and alkylating cysteins.</i>	111
<i>Supplemental Table 2.1 Overall unfolding rate and pK_{TS} values.</i>	60

List of Figures

<i>Figure 1.1 Folding landscape of αLP.</i> 1. Rate of unfolding by fluorimetry or CD with GdnHCl and extrapolation to 0 M. 2. Rate of folding with a sensitive thio-benzyl substrate. 3. Pro catalyzed folding, k_{cat} and K_M . 4. Inhibition of α LP by Pro, K_i	6
<i>Figure 1.2 Proposed Unfolding TS Model for αLP.</i> The N- and C-terminal β -barrel domains retain much of the native structure, but separate from each other allowing for solvent exposure of the domain-domain interface. Reprinted with permission from (Kelch, Eagen et al. 2007).	9
<i>Figure 1.3 Model for the unfolding transition state for the Pro-mediated proteases.</i> In the TS, the domain bridge breaks many key contacts holding the N- and C- terminal domains, which remain intact, but separate from each other and allow for water to flow in between. Reprinted with permission from (Kelch and Agard 2007).	10
<i>Figure 2.1 Salt bridges across αLP's two domains correlate with the TS model.</i> (a) Proposed model for α LP TS: Water intercalates in the domain-domain interface (shaded pink), while the N- and C-terminal domains maintain a high degree of structural integrity. Modified and reprinted with permission from (Kelch, Eagen et al. 2007) (b) The changes made in α LP to generate the salt bridge mutants. Salt bridges are eliminated by mutating them to their NAPase counterparts: Arg105→Ser and Glu8→Ala for salt bridge 1, Arg78→Leu for salt bridge 2, and Arg64→Ala and Glu182→Gln for salt bridge 3. The double salt bridge mutants SB2+3, SB1+3 and SB1+2 have the salt bridge 1, 2, or 3 eliminated, respectively. SB1, SB2, and SB3 include the salt bridges 1, 2, and 3, respectively and have the other two eliminated.	24
<i>Figure 2.2 pH-dependent unfolding of single salt bridge mutants.</i> (a) Natural logs of SB1 (purple, solid circles), SB2 (green, solid diamonds), SB3 (black, solid triangles), and wt α LP (red, solid squares) unfolding rates are plotted as a function of pH at 60 °C. (b) Fold increase of unfolding rates from neutral pH baseline. SB1 (purple), SB2 (green), SB3 (black) and wt α LP (red).	28
<i>Figure 2.3 Salt bridge energetics in the single salt bridge mutants.</i> The TS and N state pK_a 's associated with salt bridges 1, 2, and 3 in single salt bridge mutants SB1, SB2, and SB3, respectively. The shift in pK_a 's from TS to N state is translated to a free energy contribution of each salt bridge to the unfolding transition.	29
<i>Figure 2.4 pH-dependent unfolding of double salt bridge mutants.</i> (a) Natural logs of SB2+3 (purple, solid circles), SB1+3 (green, solid diamonds), SB1+2 (black, solid triangles), and wt α LP (red, solid squares) unfolding rates are plotted as a function of pH at 60 °C. (b) Fold increase of unfolding rates from neutral pH baseline. SB2+3 (purple), SB1+3 (green), SB1+2 (black) and wt α LP (red).	32
<i>Figure 2.5 Energetic coupling of salt bridge pairs result in pK_a shifts.</i> (a) TS pK_a 's of each salt bridge in the context of the salt bridge mutants, SB1, SB2 and SB3, SB2+3 (purple), SB1+3 (green), SB1+2 (black) and wt α LP (red). (b) Coupling energies corresponding to pK_a shifts in double salt bridge mutants and wt α LP associated with the reciprocal interactions of the salt bridges in the context of each protein.	33
<i>Figure 2.6 Crystal structures of salt bridge mutants.</i> (a) Overlay of the wt α LP (pink), SB1+2 (teal), and SB2+3 (green) crystal structures. (b) Salt bridge 1 in wt α LP (pink), SB1+2 (teal) and corresponding NAPase residues in SB2+3 (green). (c) Salt bridge 2 in wt α LP (pink), SB1+2 (teal), and in SB2+3 (green). (d) Salt bridge 3 in wt α LP (pink), SB2+3 (green) and corresponding NAPase residues in SB1+2 (teal).	34
<i>Figure 2.7 Cooperativity measured by reduction in number of states.</i> In red, the reduction in number of states in the TS ensemble calculated for each mutant and wt α LP. In green, contribution of each salt bridge to the reduction in number of states is calculated using linear combinations of individual contributions describing the coupling energies in the double salt bridge mutants. In dashed arrows, fold difference in number of states in TS between double salt bridge mutants and wt α LP. Green line indicates total attainable reduction in number of states.	45
<i>Figure 3.1 The general mechanism for catalysis by chymotrypsin-like serine proteases.</i> In the first stage (acylation) active site serine – deprotonated by His- attacks the carbonyl for the substrate's scissile bond forming a metastable tetrahedral intermediate (TI_1). His donates the proton to the amide of the scissile bond allowing the release of the C-terminal substrate. During the second stage, a water molecule becomes the active nucleophile. Another tetrahedral intermediate (TI_2) forms prior to the release of the product. Figure modified from ((Fersht 1999)).	65

<i>Figure 3.2 Structure of the αLP+boronic acid inhibitor complex. a) Boronic acid inhibitors are mimics of the tetrahedral intermediate TI_2. b) Structure of the boronic acid inhibitor Ala-Ala-Pro-Ala-BO_2 in αLP active site. Reprinted with permission from (Bone, Silen et al. 1989).</i>	66
<i>Figure 3.3 Anisotropy distributions in crystal structures with boronic acid inhibitors of varying length.</i>	73
<i>Figure 3.4 Anisotropy distributions of various structures with anisotropic B-factors in the Protein Data Bank. Figure reproduced from (Merritt 1999).</i>	73
<i>Figure 3.5 Inhibitor anisotropy plotted as inverse anisotropy.</i>	75
<i>Figure 3.6 Anisotropic B-factor ellipsoid volumes representing the magnitude of thermal motion for boronic acid inhibitors of varying length.</i>	76
<i>Figure 3.7 Thermal ellipsoids around the boronic acid do not show significant changes in size and directionality with inhibitors of varying length.</i>	77
<i>Figure 3.8 Anisotropy along the substrate binding pocket, first wall.</i>	78
<i>Figure 3.9 Anisotropy in the catalytic triad.</i>	79
<i>Figure 3.10 Thermal ellipsoids for the catalytic triad residues show reduced motion compared to the rest of the structure. Thermal ellipsoids of the Leu16 (41 in chymotrypsin numbering) side chain extending to the catalytic triad are anomalously large.</i>	80
<i>Figure 4.1 Folding landscape of αLP. 1. Rate of unfolding by fluorimetry or CD with GdnHCl and extrapolation to 0 M. 2. Rate of folding with a sensitive substrate. 3. Pro catalyzed folding, k_{cat} and K_M. 4. Inhibition of αLP by Pro, K_i.</i>	87
<i>Figure 4.2 Revised model of Pro-catalyzed αLP folding. (a) The pro domain (green) of the Pro-αLP precursor folds separately from the N- and C-terminal domains of αLP (blue and red, respectively), which associate with one another to form substantial local secondary structure. (b) The three-stranded β-sheet from the Pro C-terminal domain pairs with the β-hairpin of $C\alpha$LP to form a continuous five-stranded β-sheet. (c) Both N- and C-terminal domains of Pro bind to $C\alpha$LP to help arrange key structural elements, enabling (d) $N\alpha$LP and $C\alpha$LP to simultaneously fold. Interdependent folding of the αLP domains completes the protease active site, which can then process the Pro-αLP junction. The new N-terminus of αLP repositions to its native conformation, while the Pro C-terminal tail remains bound to the αLP active site, inhibiting protease activity. (e) Intermolecular cleavage of Pro by αLP, or other exogenous proteases, leads to the degradation of Pro and the release of mature, active αLP (Cunningham and Agard 2003).</i>	90
<i>Figure 4.3 Cross-linker design. Cross-linkers that are best suited are specific on one end, non-specific on the other end, soluble, rigid, have good cross-linking yield, can be enriched for and are easily detectable. Desthiobiotin has lower than biotin affinity for streptavidin for ease of purification. Isotopic labeling allows better mass spectrometric detection. The linker length can be adjusted to span different distances. Sulfonated NHS_ester and hydrophilic groups in the linker increase solubility.</i>	96
<i>Figure 4.4 Labeling of bovine insulin with NHS-esters. a) Labeling with NHS-PEO4-Biotin in 10 mM Na Carbonate buffer, pH 9/150mM NaCl at RT was stopped with 100 mM Tris, pH 8.5 after 1-2 hrs. b) 3 mg/ml Insulin (0.5mM) was mixed with 21 mM (estimated) cross-linker and 3.5 mM triethylamine in DMF and stirred 1-2 hrs at RT in dark followed by addition of acetic acid to 1M to stop the reaction.</i>	98
<i>Figure 4.5 EDC activation of Compound I. Time course followed by Insulin labeling with 35 molar equivalents of estimated cross-linker in DMF. a) MALDI mass spectra show significant mass shifts indicating multiple cross-linkers per Insulin molecule. b) Westerns with streptavidin-HRP showed significant amounts of cross-linker labeling activity after 1hr of EDC activation. Biotinylated insulin control was prepared with NHS-PEO4-Biotin.</i>	99
<i>Figure 4.6 Cross-linker to protein ratio. Varying molar equivalents of cross-linker was mixed with insulin and the labeling efficiency was followed with MALDI spectra.</i>	100
<i>Figure 4.7 Comparison of photoactivation methods. Intermolecular cross-linking as an indicator of photoactivation with a handheld UV lamp (long (L) and short (S) UV), an HP spectrometer deuterium lamp and a 1kW laser with a 300-400 nm pass filter. Samples were run in SDS-PAGE gels with reducing (R) and non-reducing (NR) sample buffers.</i>	103
<i>Figure 4.8 V8 digest of insulin. V8 is able to generate reasonably sized peptides from insulin. a) The list of expected peptides with V8 specificity. b) Insulin disulfide bridges keep peptides linked after digest in MALDI spectra. c) Disulfide linked peptides are separated upon reduction and carboxymethylation.</i>	105
<i>Figure 4.9 Precursor ion scan. In a QTRAP mass spectrometer precursor ion scan Q1 sweeps a given mass range and Q3 is fixed on mass of diagnostic fragments. Only the ions that pass through Q1 and produce a diagnostic fragment will produce signal.</i>	107

<i>Figure 4.10 Cross-linker fragmentation pattern.</i> a) Fragmentation of the cross-linker gives ions corresponding to fragments from the desthiobiotin arm. b) The fragmentation susceptible bonds on the cross-linker.	108
<i>Figure 4.11 Precursor ion scans of the cross-linker.</i> a) Precursor ion scanning with the fragment ion $m/z=240$ returns the hydrolyzed cross-linker in the spectrum. b) Precursor ion scanning with the fragment ion $m/z=197$ returns the hydrolyzed cross-linker in the spectrum.	109
<i>Figure 4.12 Precursor ion scan of biotinylated insulin.</i> Biotinylated insulin was prepared with NHS-PEO4-Biotin. a) Total ion count chromatogram using Enhanced Multiple Charge (EMC) as survey scan. b) Total ion count chromatogram using Precursor Ion Scan (270; 50) as survey scan.	110
<i>Figure 4.13 Synthesis of cross-linking reagent 1.</i> a) Ethyl bromoacetate (1.0 eq), K_2CO_3 (2.0 eq), rt 1.5 h. b) 4-azido-2,3,5,6-tetrafluorobenzoic acid (5 , 0.30 eq.), 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) (0.33 eq.), diisopropylethylamine (DIPEA) (0.9 eq.), N,N-dimethylformamide (DMF), rt, 1.0 h, 50% from 5 . c) 4N HCl in 1,4-dioxane, rt, 5 h. d) desthiobiotin (1.5 eq.), HBTU (1.5 eq.), DIPEA (6.0 eq.), DMF, rt, 1 h, 84% from 3 . e) MeOH, NaOH aq. (3.0eq.), rt, 15 h, HCl work-up. f) HOSu (1.3 eq.), PL-EDC (2.3 eq.), 4Å molecular sieves, DMF, $CHCl_3$, rt, 15 h, 78% from 4 .	115
<i>Figure 5.1 αLP unfolding in reducing conditions followed by loss of activity.</i> The unfolding rates with or without reducing agent are indistinguishable.	126
<i>Figure 5.2 αLP unfolding under reducing conditions followed by Trp fluorescence.</i> No significant differences were found between reducing and oxidizing buffers. Instead, the experimental variation in GdnHCl concentration was a stronger determinant of the small differences in unfolding rate.	127
<i>Supplemental Figure 2.1 Coupled pK_a fits of wt αLP pH-dependent unfolding rates.</i> wt α LP pH titration of unfolding rates fit by either coupled TS pK_a 's (red, solid), fixed shifted TS pK_a 's determined using calculated cooperativity contributions of each salt bridge from double salt bridge mutant fits (purple, dashed), or fixed TS pK_a 's from single salt bridge mutant fits (green, dotted).	57
<i>Supplemental Figure 2.2 Coupled pK_a fits of double titration mutants' pH-dependent unfolding rates.</i> Double salt bridge mutant pH titration of unfolding rates fit by either coupled TS pK_a 's (solid), or fixed TS pK_a 's from single salt bridge mutant fits (dotted). (a) SB2+3, (b) SB1+3, and (c) SB1+2.	58
<i>Supplemental Figure 2.3 Autolysis product in SB1+2 crystal structure.</i> Bound peptide in the SB1+2 crystal structure matches the protein sequence between residues 184-187.	59
<i>Supplemental Figure 2.4 pH-dependent unfolding rates of relocation mutants.</i>	61

Chapter 1. Introduction

Cooperativity and Kinetic Stability

Thermodynamic stability with a lower free energy native state compared to unfolded states has been the norm in the field of protein folding. However, the function of a protein can be maintained within a biologically relevant time-scale, without the native state being the thermodynamically stable state compared to the non-functional states of the protein. A sufficiently high unfolding barrier coupled with a cooperative unfolding transition can uncouple the native state from all other states enough that thermodynamic stability is no longer required. This strategy utilizing high unfolding barriers has the advantage of never sampling the partially- or fully-unfolded states, protecting the protein from irreversible alterations such as proteolysis, chemical alterations, aggregation and undesirable interactions with other macromolecules. Both in crowded intracellular environments and under the harsh conditions of proteolytic extracellular milieus, a high cooperative unfolding transition contributes distinct advantages.

In contrast to downhill barrierless folding (Cho, Weinkam et al. 2008) that goes through a glassy transition (Garcia-Mira, Sadqi et al. 2002), a two state cooperative folding landscape is an evolved property (Watters, Deka et al. 2007) for a lot of proteins, even though it is not a physical requirement. Concerted structural rearrangements in protein chains result in bottlenecks in folding and unfolding, thus a certain degree of

cooperativity is a hallmark of barriers associated in these transitions. In kinetically stable proteins, a great barrier height and extreme cooperativity sufficiently uncouple the native state from the other states (Sohl, Jaswal et al. 1998) (Jaswal, Sohl et al. 2002). Kinetic stability coevolves with extremely cooperative unfolding transitions suppressing breathing motions and making the global unfolding barrier extremely high at the same time. As a result, the evolution of cooperative behavior makes sampling of nonnative states rare and lengthens the functional life time of the protein (Cunningham, Jaswal et al. 1999).

In addition to the extensively studied kinetically stable proteins ((Dee and Yada), (Truhlar, Cunningham et al. 2004), (Sohl, Jaswal et al. 1998), a proteome wide protease screen identified about 40 proteins with high and cooperative unfolding barriers in *E. coli* (Park, Zhou et al. 2007). Similarly, an SDS screen identified about 50 such proteins (Xia, Manning et al. 2007). A closer look to the protease resistant group heavily favored periplasm binding proteins and stress response proteins. The former clearly need protection from exogenous proteases. The latter are expressed during the stationary phase, when continued protein synthesis is expensive for the bacteria and thus the extended lifetime of these proteins provides an important advantage.

The protease screen identified proteins with high and cooperative unfolding barriers of diverse folds ruling out fold specific explanations for this characteristic. Similarly, TIM barrels of different species have diverse behaviors regarding their unfolding transition (Costas, Rodriguez-Larrea et al. 2009). Thus the structural basis for

kinetic stability, although poorly understood, is most likely in sub-fold level detailed interactions that differ from protein to protein within a structural family.

Cooperativity and kinetic stability in disease context

Misfolding diseases generally fall into two groups: 1) A lack of function due to inability to fold or kinetic instability (p53 in cancer (Friedler, Veprintsev et al. 2003), (Butler and Loh 2006), α -fetoprotein in Down syndrome (Gillespie and Uversky 2000), CFTR in cystic fibrosis (Thibodeau, Brautigam et al. 2005)), and 2) amyloidosis/toxic misfolded state (crystallin in cataract (Meehan, Berry et al. 2004), (Ecroyd and Carver 2009), A β in Alzheimer's disease (Bharadwaj, Dubey et al. 2009), Prion protein in bovine spongiform encephalopathy (Mad Cow Disease) (Bocharova, Makarava et al. 2006) and amylin in Type II Diabetes (Gotz, Ittner et al. 2009)). Mutations that cause proteins to be highly prone to misfolding are often eliminated by cellular quality control mechanisms (Balch, Morimoto et al. 2008) (Kaganovich, Kopito et al. 2008). Misfolding diseases in some cases involve very conservative mutations that may escape these mechanisms, but perturb the cooperative unfolding of the protein too much, like in the case of triose-phosphate isomerase (Rodriguez-Almazan, Arreola et al. 2008).

The unfolding free energy barrier can be more susceptible to mutations than the thermodynamic stability of the protein. A few kJ/mol change in the unfolding barrier can significantly affect the unfolding time scale a lot, according to

$$\tau \propto e^{\frac{\Delta G^\ddagger}{RT}} \quad (1),$$

where $\tau = 1/k$, ΔG^\ddagger is the free energy of unfolding, R is gas constant and T is temperature (Godoy-Ruiz, Ariza et al. 2006). In contrast, a few kJ/mol in thermodynamic stability will only affect the T_m a few degrees, which is already out of the physiological range for most proteins. In addition, the rate limiting step for amyloid formation is often protein unfolding (Canet, Sunde et al. 1999) (Johnson, Wiseman et al. 2005) (de Prat Gay, Johnson et al. 1994; Thirumalai, Klimov et al. 2003). The rate of unfolding is also crucial in the case of p53 (Butler and Loh 2005) (Butler and Loh 2006).

Alpha-lytic Protease (α LP) and its folding landscape

α LP is 198 residue serine protease secreted by the gram negative soil bacterium *Lysobacter enzymogenes*. Structurally, it consist of two beta-barrel domains in strong similarity to mammalian digestive serine proteases trypsin in chymotrypsin and its sequence resembles moderate sequence homology to these.

Like virtually all extracellular bacterial proteases that are synthesized with precursor pro regions (Cunningham, Jaswal et al. 1999), α LP is synthesized with a 166 residue N-terminal Pro, which is required for folding (Silen, Frank et al. 1989).

In order to investigate the folding landscape of α LP, folding and unfolding is studied separately. When α LP is denatured and the denaturant is diluted out an

intermediate (Int) forms under the same conditions as native protease (see below). Very small fractions of Int mature into the active native state (N) and this folding rate can be measured by following the amount of N protease using a very sensitive substrate assay. According to the initial rate obtained with this assay, the α LP folding rate is determined to be $1.18 \times 10^{-11} \text{ s}^{-1}$ at 4°C ($t_{1/2} = 1,800$ years), which corresponds to 30 kcal/mol according to transition state theory (Sohl, Jaswal et al. 1998) (Figure 1.1).

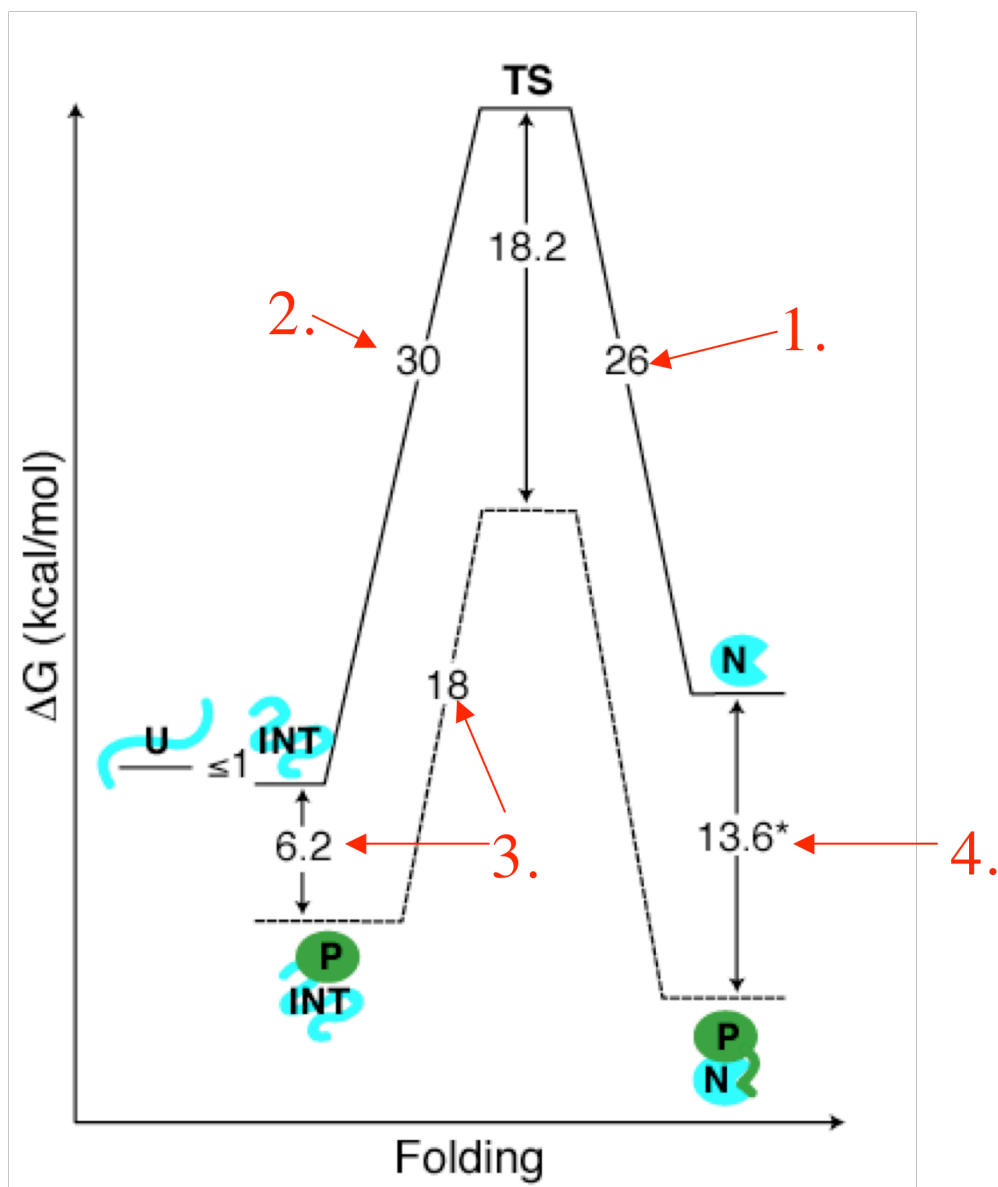


Figure 1.1 Folding landscape of α LP. 1. Rate of unfolding by fluorimetry or CD with GdnHCl and extrapolation to 0 M. 2. Rate of folding with a sensitive thio-benzyl substrate. 3. Pro catalyzed folding, k_{cat} and K_M . 4. Inhibition of α LP by Pro, K_i

Using an inactive variant, S195A, denaturant unfolding and extrapolating the data to zero denaturant, the unfolding rate is determined to be $1.8 \times 10^{-8} \text{ s}^{-1}$ at 4°C or an unfolding barrier of 26 kcal/mol ($t_{1/2} = 1.2$ years) (Sohl, Jaswal et al. 1998).

The ratio of the slower folding rate and the unfolding rate establishes an equilibrium free energy that favors Int over N. In fact, Int is only marginally stable compared to the denatured unfolded state (U), placing N to be less stable than either U or Int.

α LP has coevolved an N-terminal Pro region to surmount the extraordinary folding barrier and the thermodynamic instability of the native state. Pro can assist folding when supplied either in *cis* or *trans*. The Pro-catalyzed folding rate is measured by adding Pro to Int, which results in rapid folding to the native state with a slow and fast phase. The slow phase is thought to be related to proline isomeration, while the Pro dependent fast phase, when measured as a function of Pro concentration yields a catalyzed folding rate of 0.037 s^{-1} and an affinity for Int of $13 (\pm 5) \text{ }\mu\text{M}$.

Pro acts as a foldase, a folding catalyst, by binding tightly to the transition state (TS), lowering the folding barrier by 18.2 kcal/mol and accelerating the rate of α LP folding by 3×10^9 (Sohl, Jaswal et al. 1998). In addition to TS, Pro also binds tightly to the native state with an inhibition constant $K_i = 3 \times 10^{-10} \text{ M}$ making the Pro•N complex the thermodynamically favored state over the Pro•Int complex. As such, Pro is the tightest binding inhibitor of N (Peters, Shiau et al. 1998) (Baker, Silen et al. 1992) and this results in Pro being a single-turnover catalyst. To make active N, the Pro region is released from the inhibitory complex by proteolysis. Upon the degradation of Pro the metastable N is kinetically trapped, uncoupling it from the thermodynamically favored Int and U.

The unfolding of α LP is a largely entropic transition as determined by Eyring analysis (Jaswal, Truhlar et al. 2005): While the enthalpy of unfolding is -2 kcal/mol, the entropic contribution to the unfolding free energy is 28 kcal/mol. The strong unfavorable entropy suggests a significant increase in TS compared to the native state. While the $\Delta C_p = 1.3$ kcal/molK is large compared to the 13 studied thermodynamically stable proteins it is a fraction of the total calculated heat capacity between the U and the N states (Jaswal 2000) (see below).

The native-like TS and current models of TS

The fractional ΔC_p and m-values provide a measure of fraction non-polar and total surface area exposed in TS (Myers, Pace et al. 1995) and are calculated to be 0.61 and 0.82 for α LP, respectively, which suggests a very native-like TS. On the other hand, the high entropy of unfolding suggests a very ordered TS. In addition, a screen for faster folding mutants identified R138H/G183S, for which the uncatalyzed and Pro-catalyzed folding were similarly affected, despite a 10^9 difference in rate, suggesting that the TS is very resilient (Derman and Agard 2000).

Mutational studies on salt bridges across the N- and C-terminal domains suggested a separation along the domain-domain interface (Kelch, Eagen et al. 2007), which would also necessitate the involvement of the domain bridge, a beta-hairpin

connecting the two domains. Moreover, the height of unfolding barrier at 70 °C for Pro-dependent proteases correlates strongly with the surface area buried by the folding of the domain bridge (Kelch and Agard 2007).

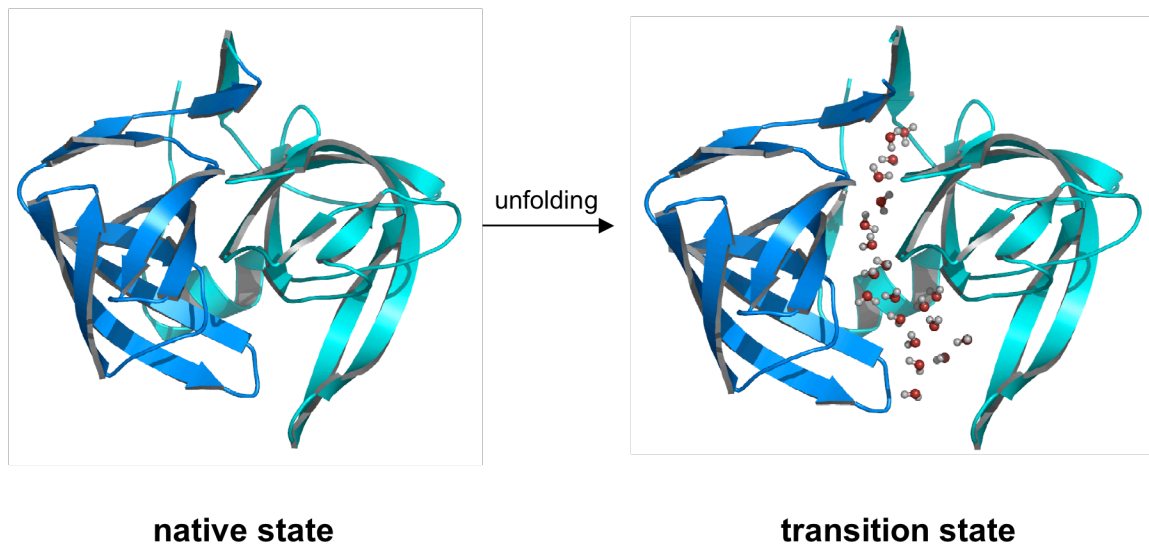


Figure 1.2 Proposed Unfolding TS Model for α LP. The N- and C-terminal β -barrel domains retain much of the native structure, but separate from each other allowing for solvent exposure of the domain-domain interface. Reprinted with permission from (Kelch, Eagen et al. 2007).

Taking all this evidence into account, a model that was first suggested by Sheila Jaswal, (Figure 1.2) was improved by Brian Kelch: In TS the N- and C- domains separate, yet remain intact themselves. Upon this separation ordering of water decreases the entropy in the TS. The domain bridge breaks many key contacts holding these two domains together (Figure 1.3).

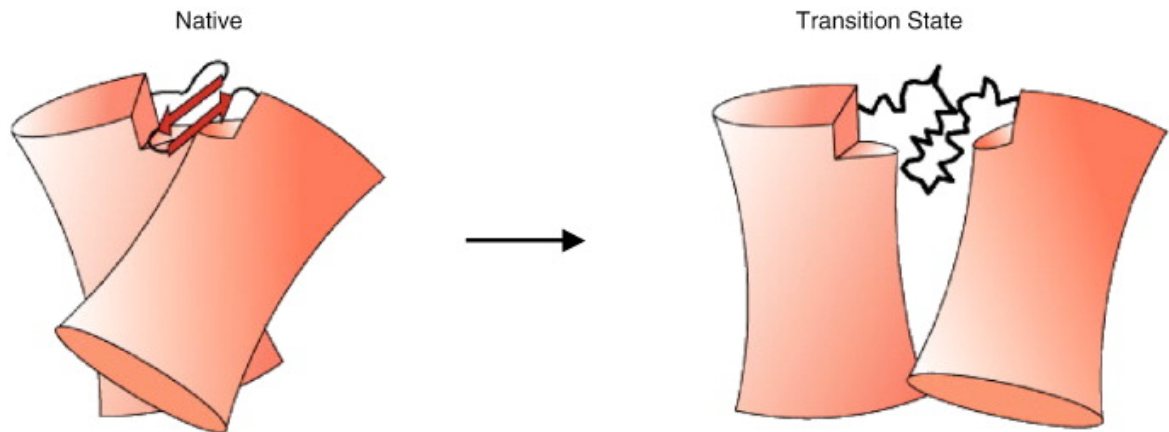


Figure 1.3 Model for the unfolding transition state for the Pro-meidated proteases. In the TS, the domain bridge breaks many key contacts holding the N- and C- terminal domains, which remain intact, but separate from each other and allow for water to flow in between. Reprinted with permission from (Kelch and Agard 2007).

Cooperative unfolding of α LP

During α LP unfolding, the secondary and tertiary structures are lost at the same rate, when monitored via CD or Trp fluorescence, respectively (Sohl Jaswal 1998), pointing to a single rate-limiting step. A survival assay was developed to compare local and global unfolding, which involved incubating several proteases and following their activity over time. α LP survival assay rate of degradation matches its global unfolding rate implying a perfectly cooperative unfolding (Jaswal, Sohl et al. 2002). Moreover, in addition to unusually low crystallographic B-factors (Fuhrmann, Kelch et al. 2004), α LP's native state has protection factors on the order of $>10^9$ for >40 core amides (Jaswal, Sohl et al. 2002). Not only are these of a magnitude never observed before, they are also distributed over a large region of the structure, as opposed to central core regions

observed in other proteins. Thus, the α LP native state is very rigid and local breathing motions are extremely suppressed. All of this evidence point to an extremely cooperative single step unfolding transition. The extreme cooperativity of the unfolding transition is clearly very important to increase the functional life time of α LP in harsh, proteolytic, extracellular environments, however very little is known about the structural basis that establishes it.

While the TS models previously proposed predict the separation of the two domains to yield unfolding, the extent of this separation is unknown. α LP could be exposing all of the domain-domain interface at once or separation of the two domains along a fraction of this interface might be sufficient for unfolding. In Chapter 2, I investigate the structural properties of TS utilizing inter-domain salt bridge mutants as reporters of unfolding. In addition, I look into the cooperative effect of placing multiple salt bridges in the domain-domain interface by analyzing the pair wise energetic coupling of multiple salt bridges. The results indicate that in the TS, the domain separation to expose the domain-domain interface happens mostly around one of the three salt bridges involving Arg64 and Glu182. While the other two inter-domain salt bridges do not significantly contribute to the height of the unfolding barrier, α LP uses combinations of the three salt bridges to make the unfolding transition more cooperative and combines their cooperative effects efficiently. By placing these three salt bridges across the two domains α LP optimizes its unfolding cooperativity compared to its acidophile homolog NAPase, but the choice for these electrostatic interactions renders α LP unfolding barrier susceptible to acid. Moreover, the analysis allows for a quantitative assessment of the

degree of unfolding cooperativity contributed by each of the structural elements that is associated with the salt bridge coupling energy.

Characterization of the Intermediate state

In vitro refolding of chemically denatured α LP in the absence of Pro results in an inactive molten globule-like intermediate (Int) upon dilution from denaturant (Baker, Sohl et al. 1992). Int is monomeric and expanded in size relative to the native state. CD spectroscopy of Int indicated it possesses substantial secondary structure, however Trp fluorescence studies showed that it lacks stable tertiary interactions required to bury the Trp residues. Urea denaturation studies showed that Int is marginally stable (<1 kcal/mol) relative to the U state (Sohl, Jaswal et al. 1998). Under native conditions, Int exists without appreciable maturation to N for months.

Calorimetry of Int showed that the origin of its stability over N is through a strong entropy term (Sohl, Jaswal et al. 1998): N is enthalpically favored over Int by 18 kcal/mol, but entropically disfavored by -22 kcal/mol. The high entropy in Int could be originating from the increased flexibility due to the high abundance of glycines in α LP's sequence. The Pro dependent α LP subfamily of proteases contain 16-20% glycine, while trypsin family mammalian digestive proteases do only 7-11% (Jaswal 2000). The higher glycine content would be expected to stabilize Int by about 7 kcal/mol at 4°C (D'Aquino, Gomez et al. 1996).

In addition, the large entropic free energy component could be a result of the extreme rigidity in the native state demonstrated by survival assays, low crystallographic B-factors and wide spread, extremely high protection factors (see above). Contributing to this rigidity, α LP's loops are generally shorter and thus likely to be less flexible than those found in chymotrypsin.

The folding transition of Int to TS has a small $\Delta C_p = -0.17$ kcal/molK and hence a shallow dependency of folding parameters to temperature. α LP has a completely different behavior compared to 13 thermodynamically stable proteins from literature: Its entire free energy curve is shifted left to lower temperatures. The folding free energy minimum is reached at around 100K, nearly 200K below the next free energy minimum in the temperature dimension. Thus, in the physiological range, the free energy of folding is monotonically increasing with temperature.

In order to dissect the folding transition of Int, the N- and C- domains of α LP were purified and their folding studied (Cunningham and Agard 2003). These studies showed that the two domains fold interdependently, as opposed to the independent folding of the two domains in trypsin and chymotrypsin (Higaki and Light 1986) (Light and al-Obeidi 1991). When mixed, the two domains rapidly dock onto each other, however the rate limiting step is a cooperative rearrangement to form the native state (Cunningham and Agard 2003), which suggests that the domain interface plays an important role in folding, as well.

While a study of Int and its folding transition in greater detail would yield valuable information about the nature of the folding transition, it comes with its own challenges. Int aggregates at concentrations over 10 μM and at higher temperatures out-ruling the use of NMR to study its structure (Jaswal 2000). In addition, uncatalyzed folding requires very long experiments, making them extremely impractical in usual laboratory settings.

In Chapter 4, I discuss the development of a method for the characterization of nonnative structures like αLP Int. The method development includes the design and synthesis of tri-functional cross-linkers that can best map the protein structure providing distance constraints by intramolecular cross-linking of the protein followed by the identification of cross-linked peptides via mass spectrometry. The designed cross-linkers provide two means to enhance the detection of cross-linked peptides: with a desthiobiotin affinity tag on the cross-linker and a characteristic fragmentation pattern of the same tag in the mass spectrometer. While the individual steps of the method are optimized, challenges related to the low abundance of each molecular species due to unspecific cross-linking chemistry renders the detection of each cross-linked peptide beyond the limits of current instrumentation. Improvements in the signal to noise levels and detection limits of mass spectrometers may allow for the use of this method for structural assessment of native and nonnative proteins in the future.

α LP catalytic mechanisms

In addition to being an exemplary folding model, α LP has been a useful model to understand enzyme catalysis (Bone, Shenvi et al. 1987), (Fuhrmann, Daugherty et al. 2006) and substrate specificity (Bone, Frank et al. 1989), (Bone, Fujishige et al. 1991), (Bone, Sampson et al. 1991), (Bone, Silen et al. 1989), (Mace and Agard 1995; Mace, Wilk et al. 1995), (Ota and Agard 2001). α LP has preference for small hydrophobic residues in the S1 position. Two residues lining the substrate walls, M190 and M213, when mutated to alanine, will result in loss of enzyme specificity (Bone, Silen et al. 1989). Further investigations showed that there is conformational coupling between the two substrate walls and the alanine substitutions result in the loss of this coupling (Bone, Fujishige et al. 1991), (Bone, Silen et al. 1989), (Ota and Agard 2001), (Miller and Agard 1999).

Cynthia Fuhrmann studied enzyme catalysis of α LP using ultra-high resolution crystal structures complexed with boronic acid inhibitors or in apo form. With a particular interest in the effect of strong hydrogen bonds, she crystallized and solved α LP+boroVal(gol). A short ionic hydrogen bond between His 57 N ϵ 2 and boroValLP1 O2, the mimic of the substrate leaving group implies a low energy barrier for proton transfer, which would promote the TI₁ \rightarrow EA step of the reaction (Fuhrmann, Daugherty et al. 2006). In addition to this hydrogen bond, there are two others contacting the boronic acid directly, which as a group were among the shortest of all hydrogen bonds in the structure. Three other very short hydrogen bonds help position His57 and Asp102 in the

catalytic triad. On the other hand, the three hydrogen bonds connecting P1 and P2 positions to the enzyme have donor-acceptor distances close to (or larger) than the average distance for N...O hydrogen bonds, suggesting that perhaps the transition state stabilization is favored over substrate binding. In fact, kinetic studies have found a strong dependence of k_{cat} to the length of the substrate rather than K_m , as well as a strong dependence of K_i to peptidyl boronic acid inhibitor length. An explanation of this may be the optimization of binding site dynamics upon peptide binding.

In Chapter 3, I try to answer this question using crystal structures of α LP+peptidyl boronic acid inhibitors of varying length. Since the crystal structures had excellent data to parameter ratio, they were refined with anisotropic B-factors that provide great detail about the amplitude and direction of protein dynamics at atomic detail. Neither the positioning of the atoms nor their dynamics at the active site or the binding pocket show inhibitor length dependent tendencies to allow us predict the origins of the k_{cat} effect.

Specific aims

- What are the structural rearrangements in the folding transition of Int to Transition State?
- What is the structure in the transition state?
- What are the physical origins for cooperativity?
- How is cooperativity quantified?
- How are proteins evolved to adapt to extreme conditions under the control of a cooperative kinetic stability?
- What are the origins of the substrate length dependent k_{cat} effect for α LP enzyme catalysis?
- What is the role of disulfide bridges in α LP folding and unfolding?

Chapter 2. Quantifying protein unfolding cooperativity with acid sensitive probes: inter-domain salt bridge contributions to unfolding cooperativity are combined efficiently in α -Lytic Protease

Preface

The kinetic studies in this work are performed by me. The crystal structures were solved by Chris A. Waddling. I provided the protein for the crystals and helped setting up the crystals. An expanded version of this chapter will be submitted for publication with me as the first author, as I have contributed experimentally and intellectually the majority of this work.

Abstract

α -Lytic protease (α LP) is a kinetically stable protein with a very high barrier to unfolding, extraordinary unfolding cooperativity and unusually suppressed native state dynamics (Sohl, Jaswal et al. 1998), (Jaswal, Sohl et al. 2002). Together, these properties result in a remarkable resistance to proteolytic destruction. Unlike its acid stable homolog NAPase, α LP rapidly unfolds at low pH. While the two proteins have the same number of acid titratable groups, α LP has three salt bridges that span the N- and C-terminal domain interface while NAPase has none (Kelch, Eagen et al. 2007). This differential distribution raises an important question about the role of electrostatic interactions in the unfolding transition and provides a powerful means of probing energetic contributions, structural rearrangements and unfolding cooperativity along the domain interface. By mutation and pH titration, we examined the contribution of individual, pairs and all three salt bridges to α LP unfolding kinetics. Paralleling α LP's extreme rigidity, the pK_a values determined for the native state (<1.23 , <1.57 , <0.84) were amongst the lowest reported. Of the three inter-domain salt bridges, salt bridge 3 (Arg64/Glu182) had the most significant contribution to the height of the unfolding barrier, indicating an involvement in structural rearrangements in the unfolding transition. In combination, each salt bridge caused a reciprocal transition state (TS) pK_a shift on the others, the greatest of which caused by salt bridge 1 (Glu8/Arg105) followed by salt bridge 2 (Arg78/C-term). This suggests that α LP inter-domain salt bridges contribute to the unfolding cooperativity by decreasing the number of ways for passage into the TS ensemble. α LP combines the

cooperativity contributions from multiple sites efficiently, seizing all but 12% of the total attainable coupling energy. Although not as adaptive to acidic conditions as NAPase, α LP instead utilizes the inter-domain placement of three salt bridges to optimize its unfolding cooperativity at neutral pH.

Introduction

The extracellular bacterial protease, α -lytic protease (α LP) has an extremely high unfolding free energy barrier ($t_{1/2} = 1800$ years) which separates the native (N) state from the thermodynamically more stable (by ~ 4 kcal/mol) intermediate (I) and unfolded (U) states (Sohl, Jaswal et al. 1998). α LP maintains its functional folded state by an extremely cooperative unfolding transition as evidenced by hydrogen exchange protection factors of greater than 10^9 throughout a large hydrophobic core that spans both N- and C-terminal domains (Jaswal, Sohl et al. 2002). Under harsh or highly proteolytic conditions, this all-or-none unfolding affords α LP a greatly enhanced survival time compared to its mammalian homologues trypsin and chymotrypsin (Jaswal, Sohl et al. 2002).

To bias the thermodynamic stability towards the N state, as well as facilitate folding on a biologically relevant timescale, α LP has co-evolved a covalently linked pro-region that both catalyzes folding by a factor of $\sim 10^9$ and stabilizes the N state sufficiently to shift the thermodynamic equilibrium in favor of folding (Peters, Shiau et al. 1998), (Sauter, Mau et al. 1998), (Sohl, Jaswal et al. 1998). The pro-region is proteolyzed upon folding, yielding the kinetically trapped native α LP (Cunningham and Agard 2004). A consequence of the large kinetic barriers is that the N state energetics are uncoupled from those of the unfolded states, allowing thermodynamically unfavorable yet functionally favorable choices to be made in sequence and geometry. Thus the relevant energetics for stability is not between folded and unfolded states but between the

folded state and the highly structured unfolding transition. Moreover, an unusually high entropy of unfolding combined with a native-like TS supports a TS model with relatively intact N- and C-terminal domains intercalated by water, exposing hydrophobic surface area (Figure 2.1 (a)). Tuning the interaction between domains is thus an ideal strategy for optimizing beneficial properties of the unfolding transition.

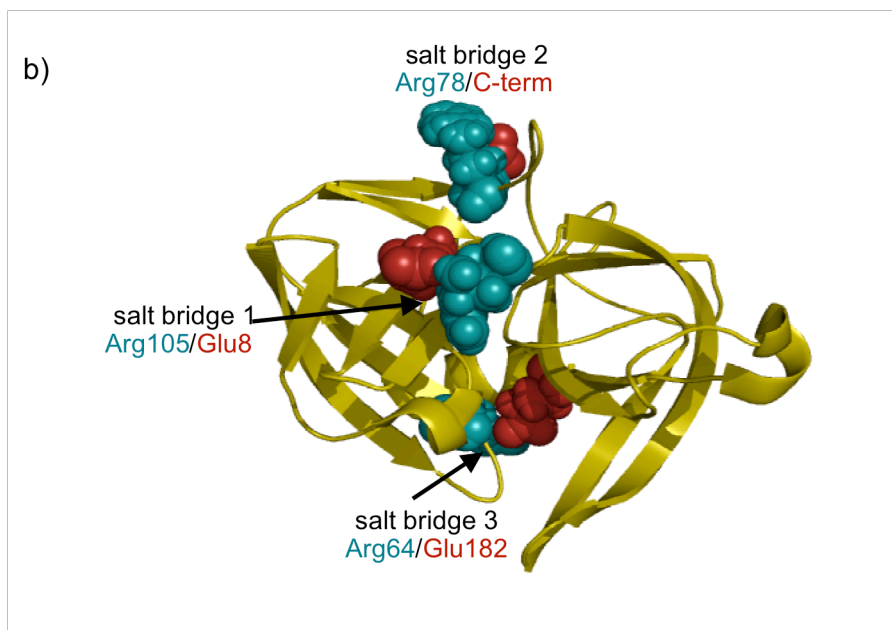
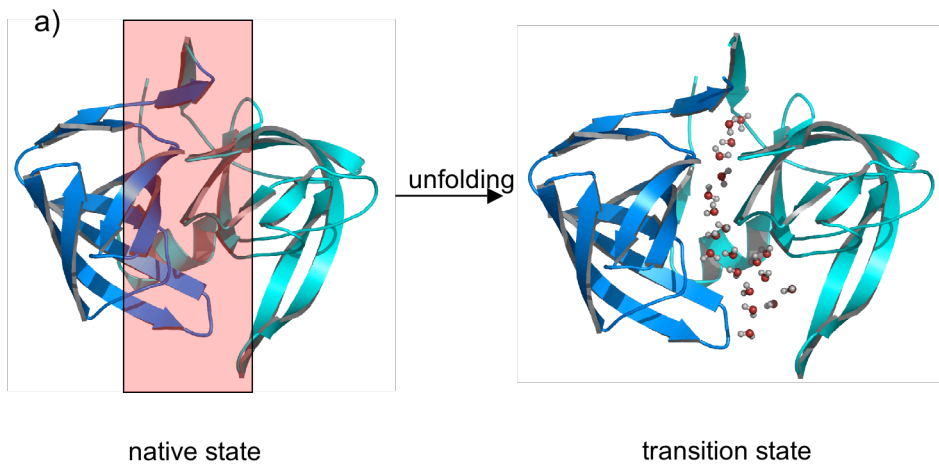


Figure 2.1 Salt bridges across α LP's two domains correlate with the TS model. (a) Proposed model for α LP TS: Water intercalates in the domain-domain interface (shaded pink), while the N- and C-terminal domains maintain a high degree of structural integrity. Modified and reprinted with permission from (Kelch, Eagen et al. 2007) (b) The changes made in α LP to generate the salt bridge mutants. Salt bridges are eliminated by mutating them to their NAPase counterparts: Arg105 \rightarrow Ser and Glu8 \rightarrow Ala for salt bridge 1, Arg78 \rightarrow Leu for salt bridge 2, and Arg64 \rightarrow Ala and Glu182 \rightarrow Gln for salt bridge 3. The double salt bridge mutants SB2+3, SB1+3 and SB1+2 have the salt bridge 1, 2, or 3 eliminated, respectively. SB1, SB2, and SB3 include the salt bridges 1, 2, and 3, respectively and have the other two eliminated.

α LP displays ~400-fold increase in unfolding rate at pH 2.5 relative to pH 5, while its acid-resistant homolog *Nocardiopsis alba* Protease A (NAPase) only accelerates its unfolding by ~5-fold over the same pH range. (Kelch, Eagen et al. 2007) While most previously studied acidophiles utilize a strategy of lowering the pI and minimizing electrostatic repulsion by increasing the number of carboxylates or of eliminating carboxylates to reduce the magnitude of the overall electrostatic effect (Cooper, Khan et al. 1990), (Fushinobu, Ito et al. 1998), (Bonisch, Schmidt et al. 2002), (Kashiwagi, Kunishima et al. 1997), (Schafer, Magnusson et al. 2004), (Settembre, Chittuluru et al. 2004), (Goto and Nishikiori 1991), (Chen, You et al. 1991), (Ionescu and Eftink 1997), (Flanagan, Garcia-Moreno et al. 1983), (Hughson and Baldwin 1989), (Geierstanger, Jamin et al. 1998), (Barrick and Baldwin 1993). NAPase uses neither strategy. While the number of acid titratable groups is identical and the number of salt bridges is very similar, α LP and NAPase distribute their salt bridges quite differently: α LP has three salt bridges spanning the interface between its N- and C- terminal β -barrels, while NAPase has none (Kelch, Eagen et al. 2007) (Figure 2.1 (b)). Previous work showed that relocation of an α LP inter-domain salt bridge (SB1, see below) to an intra-domain location by mutation to residues found in NAPase (Relo1) decreased the sensitivity to low pH unfolding by ~15-fold. The rate at pH 2.5 was ~27 faster than at pH 5 (Kelch, Eagen et al. 2007). This observation directly supports the hypothesis of domain separation in the unfolding transition state.

Here, we study the energetic contributions of inter-domain salt bridges as probes to examine the spatial distribution of transition state determinants as well as quantify

origins of cooperativity in the unfolding transition. We determine the pK_a of each of the salt bridges as an indicator for the strength of structure around the salt bridges in N state, but more importantly in TS. Moreover, we look at the combined effect of distant loci, utilizing mutants with combinations of salt bridges to study the cooperative effect to the unfolding transition in the presence of multiple interactions spread around the protein. We report the magnitude of this cooperativity in a quantitative free energy measure, in addition to a model for the structure in the TS.

Results

pH dependent unfolding profiles of single salt bridge mutants

To investigate the contribution of salt bridges at these domain-domain interface locations in isolation, we first created three mutants (SB1, SB2, or SB3), each having the residues in the other two salt bridges switched to their counterparts in NAPase, leaving only one salt bridge present in the domain-domain interface (Figure 2.1 (b)). The SB1 and SB2 mutants unfold significantly faster than wt α LP at pH 5 (~13-fold and ~48-fold, respectively), indicating a lowered overall unfolding free energy barrier. Despite overall kinetic destabilization, these mutants show decreased acid sensitivity: While wt α LP unfolds ~27-fold faster at pH 2.5 than at pH 5, SB1 is accelerated ~19-fold and SB2 is only accelerated by ~ 6-fold. By contrast, SB3 unfolding rates remain essentially unchanged from α LP at either pH (Figure 2.2 (a), Supplemental Table 2.1). This makes SB2 ~4.1 times as acid stable as wt, SB1 ~1.7 and SB3 merely ~1.1 times (acid stability= $(k_{low\ pH, WT} / k_{neutral\ pH5, WT}) / (k_{low\ pH, mutant} / k_{neutral\ pH5, mutant})$) averaged between pH=2.0-2.67; Figure 2.2(b), Supplemental Table 2.1).

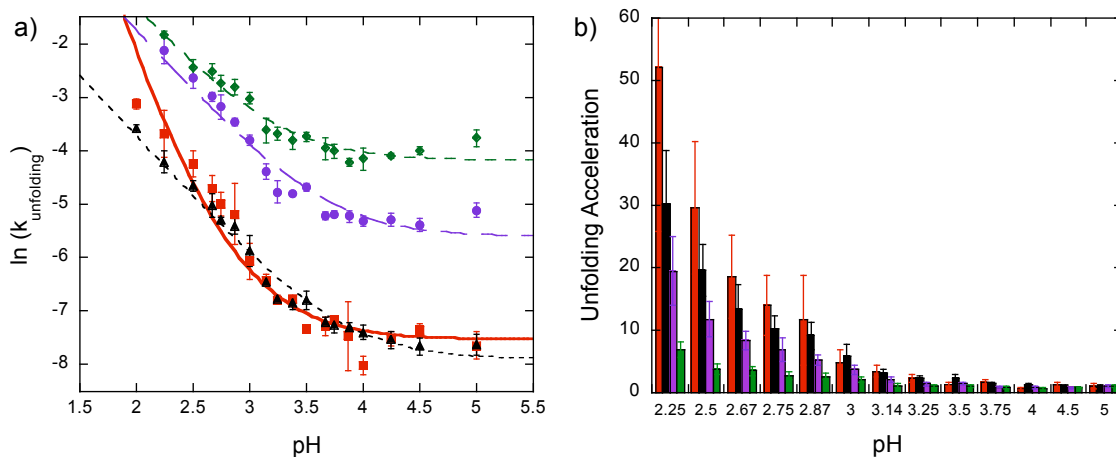


Figure 2.2 pH-dependent unfolding of single salt bridge mutants. (a) Natural logs of SB1 (purple, solid circles), SB2 (green, solid diamonds), SB3 (black, solid triangles), and wt α LP (red, solid squares) unfolding rates are plotted as a function of pH at 60 °C. (b) Fold increase of unfolding rates from neutral pH baseline. SB1 (purple), SB2 (green), SB3 (black) and wt α LP (red).

According to Eyring reaction rate theory (Eyring 1935), the overall unfolding rate is taken to be proportional to the relative populations of the transition state (TS) and native (N) ensembles, $k_u \propto [TS]/[N]$. Under conditions where equilibrium of the acidic residue protonation is established faster than the unfolding rates, the effect of broken salt bridges is reflected in the population shifts of each of the states.

$$k_u \propto \frac{[TS] + [TS \cdot H]}{[N] + [N \cdot H]} = \frac{[TS]}{[N]} \cdot \frac{1 + [H]/K_{TS}}{1 + [H]/K_N} \quad (1)$$

or

$$k_u \equiv k_o \frac{1 + [H]/K_{TS}}{1 + [H]/K_N} \quad (2),$$

where k_o is an indicator of the overall height of the unfolding barrier.

In the case of a single salt bridge being involved in the N to TS transition, the K_{TS} and K_N values are directly modulated by the K_a values of the acidic residue in the salt bridge in the TS and N states respectively.

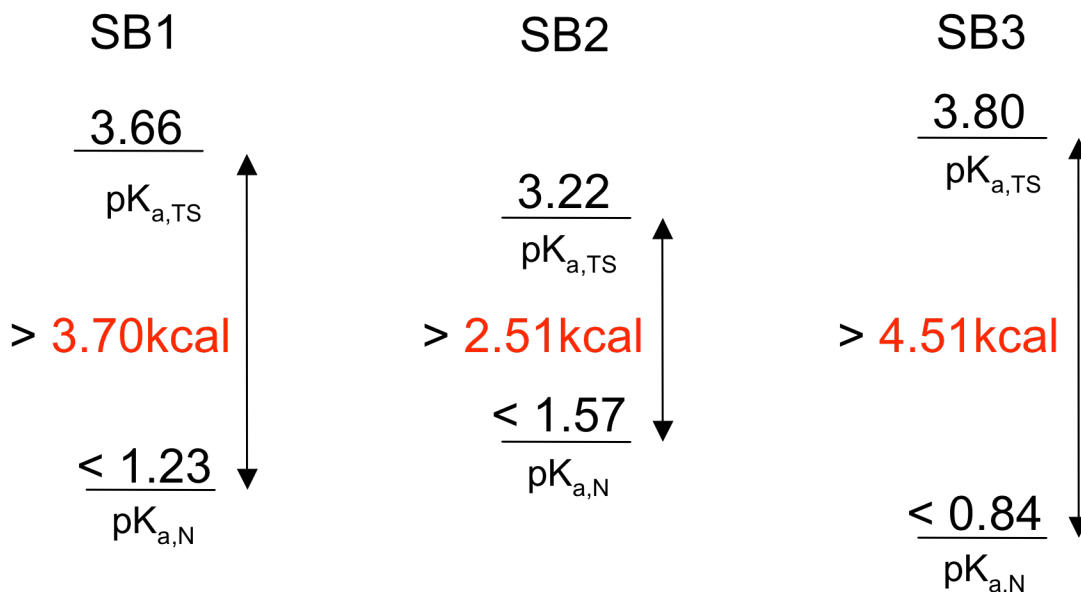


Figure 2.3 Salt bridge energetics in the single salt bridge mutants. The TS and N state pK_a 's associated with salt bridges 1, 2, and 3 in single salt bridge mutants SB1, SB2, and SB3, respectively. The shift in pK_a 's from TS to N state is translated to a free energy contribution of each salt bridge to the unfolding transition.

When the pH-dependent kinetic unfolding data for the SB1, SB2 and SB3 are fit with Eqn. 2 the pK_N values appear to be lower than the pH range of our data collection making accurate determination difficult. However it is practical to determine upper limits for the pK_N values by fitting the pK_{TS} using a range of fixed pK_N values and determining when the error of the fit worsens significantly (data not shown). Maximum allowable

value for the pK_N values of the acidic residues in the three salt bridges are thus determined to be $pK_{N1} < 1.23$, $pK_{N2} < 1.57$, and $pK_{N3} < 0.84$. Corresponding to two glutamate residues and a C-terminal carboxyl group, these values are much lower than the range previously measured in various proteins (Wisz and Hellinga 2003), and are most likely correlated with the extreme rigidity of the α LP native state (Jaswal, Sohl et al. 2002).

The pK_{TS} values are thus determined according to

$$k_u \equiv k_o \frac{1 + [H]/K_{TS}}{1 + [H]/K_N} \approx k_o(1 + [H]/K_{TS}) \quad (3)$$

to be $3.66(\pm 0.09)$, $3.22(\pm 0.06)$ and $3.80(\pm 0.07)$ for SB1, SB2 and SB3, respectively. We determine that each of the salt bridges in the SB1, SB2 and SB3 mutants would contribute a

$$\begin{aligned} \Delta\Delta G &= -2.303RT(pK_{a,TS} - pK_{a,N}) \\ &= -1.52 \text{ kcal/mol}(pK_{a,TS} - pK_{a,N}) \end{aligned} \quad (4)$$

of at least 3.7 kcal/mol, 2.61 kcal/mol and 4.51 kcal/mol to the unfolding transition at 60°C (Tanford 1970), (Yang and Honig 1992).

In the fitting process, we sought to determine if our data indicated whether titratable groups other than the three salt bridges were contributing to the K_{TS} . Based on this analysis (see Methods) we conclude that no additional groups contribute to the N to TS transition over our observable pH range.

Typical pK_a values determined for glutamic acids in solution peptides are 4.25-4.4 while those for a free C-terminal carboxylate are 3.67-3.8 (Thurlkill, Grimsley et al. 2006), (Nozaki and Tanford 1967). Thus, the TS pK_a values for the SB1 and SB3 glutamates (3.66 and 3.8, respectively) and the C-terminal carboxylate in SB2 (3.22) are significantly lower than the values determined in free solution, suggesting substantial residual structure in the environment of these salt bridges in the transition state with a free energy contribution at 60°C of at least 0.89 kcal/mol, 0.68 kcal/mol, and 0.68 kcal/mol for SB1, SB2, and SB3 respectively, according to Eqn. 4.

pH dependent unfolding profiles of double salt bridge mutants

If all three of these salt bridges were contributing independently to α LP unfolding, then the pH dependent unfolding profile of wt α LP would then be explained by:

$$k_u = k_o \left(1 + \frac{[H]}{K_{TS,1}}\right) \left(1 + \frac{[H]}{K_{TS,2}}\right) \left(1 + \frac{[H]}{K_{TS,3}}\right) \quad (6),$$

where k_o is an indicator for the overall height of the unfolding barrier and $K_{TS,1}$, $K_{TS,2}$, $K_{TS,3}$ are determined by the pH dependent unfolding profiles of SB1, SB2 and SB3 as described above. Such a model clearly fails to describe the wt α LP data (Supplemental Figure 2.1), indicating that the salt bridges interact with one another.

In order to probe the coupled behavior, we generated three mutants that each contained a different pair wise combination of the three salt bridges (Figure 2.1 (b)). At

high pH, the overall kinetic stability of the SB2+3 and SB1+3 mutants remain very similar to that of wt α LP, while the SB1+2 mutant unfolds ~ 10 times faster than the wt (Figure 2.4 (a)). This, combined with the wt-like behavior of SB3 as opposed to SB1 and SB2 (Figure 2.2(a), Supplemental Table 2.1) indicates the dominating importance of the salt bridge 3 in the overall kinetic stability of the protein.

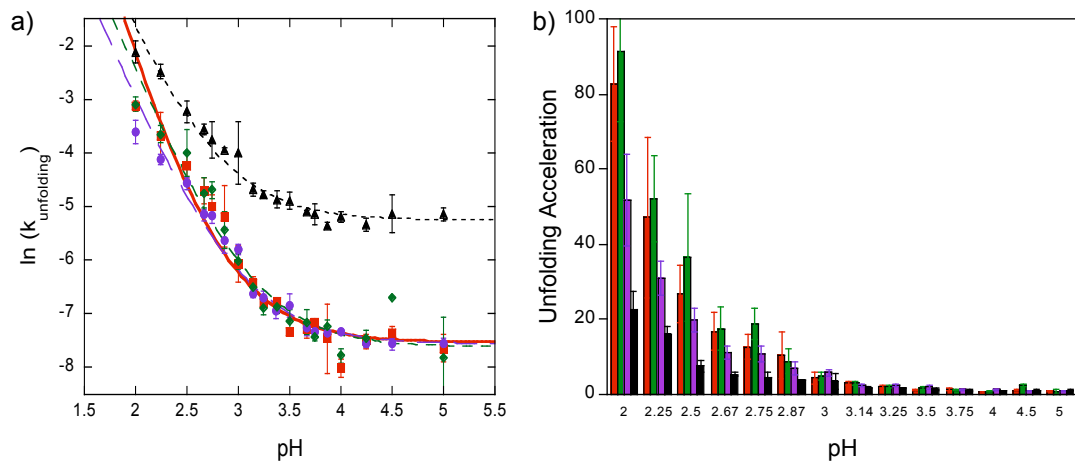


Figure 2.4 pH-dependent unfolding of double salt bridge mutants. (a) Natural logs of SB2+3 (purple, solid circles), SB1+3 (green, solid diamonds), SB1+2 (black, solid triangles), and wt α LP (red, solid squares) unfolding rates are plotted as a function of pH at 60 °C. (b) Fold increase of unfolding rates from neutral pH baseline. SB2+3 (purple), SB1+3 (green), SB1+2 (black) and wt α LP (red).

While the acid sensitivity of the SB1+3 and SB2+3 mutants are about the same as wt α LP within error, the SB1+2 mutant has decreased acid sensitivity relative to wt α LP with its unfolding rate at pH 2.5 ~ 8 -fold relative to pH 5 (Figure 2.4 (b) and

Supplemental Table 2.1), demonstrating an acid stability of 3.3-times compared to wt α LP.

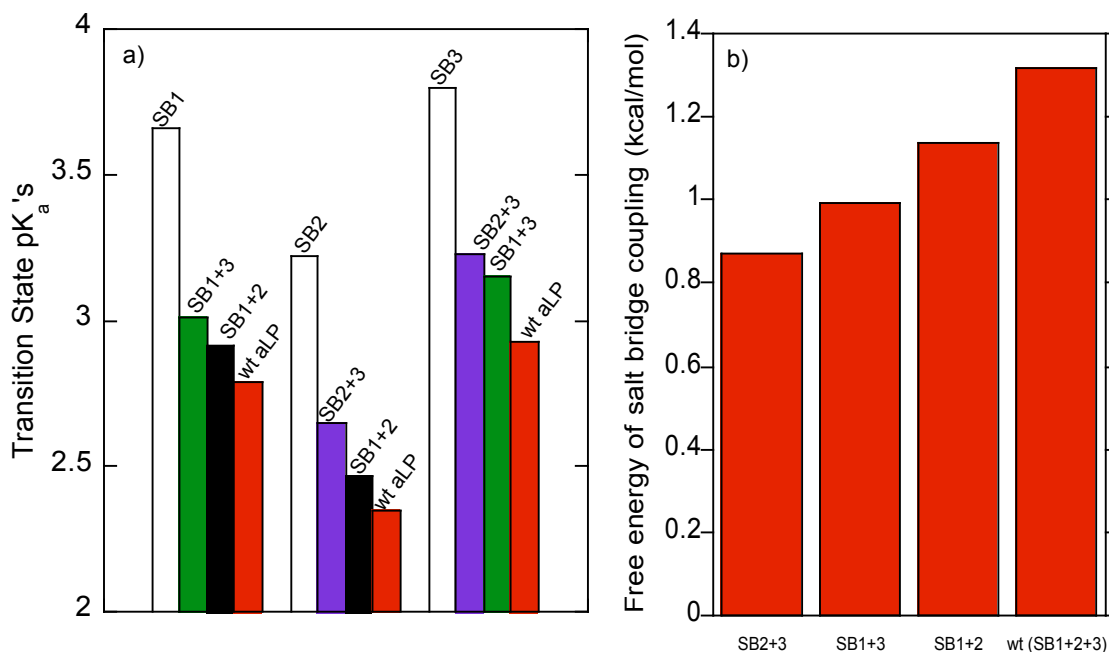


Figure 2.5 Energetic coupling of salt bridge pairs result in pK_a shifts. (a) TS pK_a 's of each salt bridge in the context of the salt bridge mutants, SB1, SB2 and SB3 (clear), SB2+3 (purple), SB1+3 (green), SB1+2 (black) and wt α LP (red). (b) Coupling energies corresponding to pK_a shifts in double salt bridge mutants and wt α LP associated with the reciprocal interactions of the salt bridges in the context of each protein.

Crystal structures of SB2+3 and SB1+2

In order to determine whether the mutation of one salt bridge affects another structurally, we solved crystal structures of SB2+3 and SB1+2 (Table 2.1) and compared

them to a previous ultra high-resolution crystal structure of wt α LP (Fuhrmann, Kelch et al. 2004).

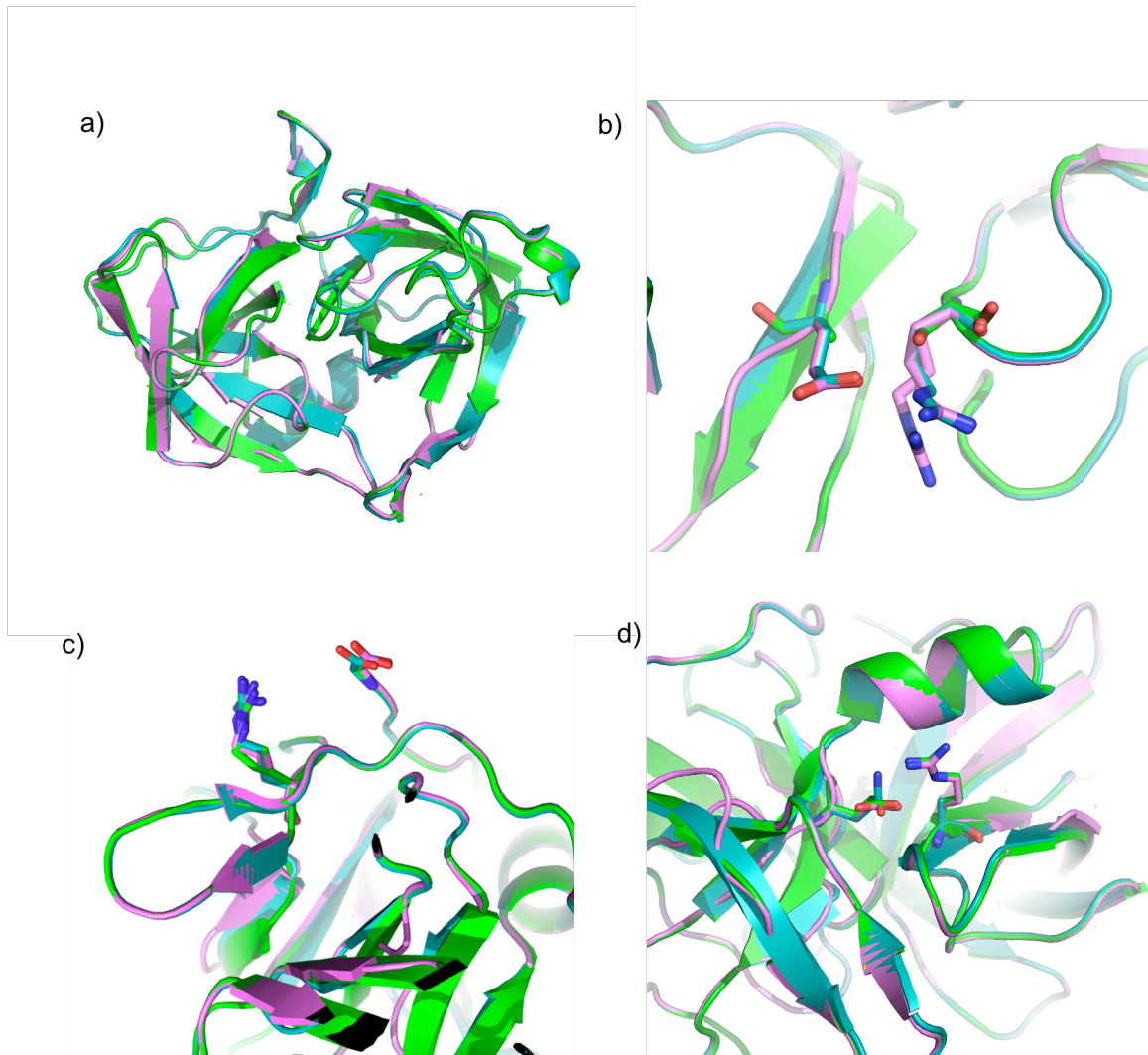


Figure 2.6 Crystal structures of salt bridge mutants. (a) Overlay of the wt α LP (pink), SB1+2 (teal), and SB2+3 (green) crystal structures. (b) Salt bridge 1 in wt α LP (pink), SB1+2 (teal) and corresponding NAPase residues in SB2+3 (green). (c) Salt bridge 2 in wt α LP (pink), SB1+2 (teal), and in SB2+3 (green). (d) Salt bridge 3 in wt α LP (pink), SB2+3 (green) and corresponding NAPase residues in SB1+2 (teal).

Table 2.1 Data collection and refinement statistics.

	SB 2+3	SB 1+2
Data Collection		
Space group	P 3 ₂ 21	P 3 ₂ 21
Cell dimensions a, c (Å)	66.015, 79.749	65.396, 79.856
Resolution (Å)	1.55	1.13
R _{fac}	0.10	0.13
Completeness (%)	97.40	99.59
Redundancy	5.5	19.4
Refinement		
Resolution (Å)	33.01-1.55	46.20-1.13
No. of unique reflections	28905	74012
R _{work} /R _{free} ^a	15.54/20.10	14.65/17.04
No. of atoms		
Protein	1434	2887
Glycerol	6	N/A
Sulfate	25	20
Peptide	N/A	69
Water	452	885
Avg B-factors, Å ²		
Protein	10.4	11.0
Glycerol	14.4	N/A
Sulfate	19.8	21.8
Peptide	N/A	29.1
Water	31.9	32.6
rmsd		
Bond lengths, Å	0.04	0.008
Bond angles, °	0.933	1.258

^a R_{free} is calculated from 5% of reflections chosen randomly.

^b Number of atoms includes hydrogen atoms

The crystal of the SB1+2 mutant resulted in a structure with an average backbone rmsd of 0.158Å from that of the wt (Figure 2.6 (a)). Mutated residues displayed rmsd's of 0.20 and 0.17Å, respectively, both well under 1σ from the average value. A bound peptide, which was determined to be a self-proteolysis product comprised of residues 184-187, with a sequence of LQPI, was refined to a group occupancy of 0.90. The

largest rmsd's were observed in protein residues within a 3Å radius of the bound peptide (Supplementary Figure 2.3), which were shifted by up to 0.63 Å (up to 4x above the average rmsd). Despite this, thermal parameters for protein residues near the bound peptide were essentially the same as those seen in the wt structure, in which there was no peptide.

The SB2+3 mutant resulted in a structure with an average backbone rmsd of 0.103Å from that of the wt (Figure 2.6 (a)). Mutated residues displayed rmsd's of 0.10 and 0.21Å, respectively, and two sidechain and backbone conformations were observed for the nearby Ser105. Residue Gly141, which is within a 3Å radius of the R105S site, exhibits an rmsd from its wt position of 0.3Å (3x the average).

The absence of the salt bridge 1 or salt bridge 3 caused only minor local perturbations and did not cause structural alterations at the remaining salt bridge locations (Figure 2.6 (b), (c), and (d)), excluding the possibility of grand pK_a shifts for these salt bridges due to altered electrostatic interactions.

Cooperativity through non-local interactions between salt bridges

Much like with wt α LP, a model using the K_{TS} values determined from single salt bridge pH-unfolding profiles together with a fit k_o as an indicator for the overall height of the unfolding barrier, fails to describe the pH dependent unfolding profiles of the mutants containing two salt bridges, suggestive of a cooperative behavior between the salt

bridges. This indicates that individual K_{TS} values are different in the presence of other salt bridges. Since no structural changes in the mutants that would directly perturb the other salt bridges were observed in the SB1+2 and SB2+3 crystal structures, we wanted to see if the pH-unfolding profiles of the double salt bridges could be explained by a coupled change in the pK_a values associated with each salt bridge. Thus, for SBi+j

$$k_{u,i+j} = k_{o,i+j} \left(1 + \frac{[H]}{\xi_{i+j} K_{a,TS,i}}\right) \left(1 + \frac{[H]}{\xi_{i+j} K_{a,TS,j}}\right), \dots, i \neq j \quad (7)$$

and

$$\Delta pK_{a,i+j} = -\log(\xi_{i+j}), \dots, i \neq j \quad (8),$$

where $k_{o,i+j}$ is an indicator of the overall barrier height for the N to TS transition, $K_{a,TS,i}$ and $K_{a,TS,j}$ are determined by the SB1 and SB2 data as described above and ξ_{i+j} is the coupling factor for the cooperative dynamics in SBi+j.

pK_a shifts of $\Delta pK_{a,2+3} = -0.57 (\pm 0.07)$, $\Delta pK_{a,1+3} = -0.65 (\pm 0.10)$ and $\Delta pK_{a,1+2} = -0.75 (\pm 0.06)$ were determined, corresponding to free energies of 0.87 (\pm) kcal/mol, 0.99 (\pm) kcal/mol and 1.14 (\pm) kcal/mol according to Eqn. (4) for SB2+3, SB1+3 and SB1+2 respectively (Figure 2.5 (a) and Figure 2.3). As with the single salt bridge mutants, introducing a third $K_{a,TS}$ representing an additional group did not improve the fits over the ones with two coupled or fixed K_{TS} 's. The cooperative behavior manifesting itself in these coupled pK_a changes best interprets the pH dependent unfolding profiles of these double salt bridge mutants.

The pH unfolding profile of wt α LP with all three salt bridges present is similarly fit using coupled pK_a values, with

$$k_{u,wt} = k_{o,wt} \left(1 + \frac{[H]}{\xi_{wt} K_{a,TS,1}}\right) \left(1 + \frac{[H]}{\xi_{wt} K_{a,TS,2}}\right) \left(1 + \frac{[H]}{\xi_{wt} K_{a,TS,3}}\right) \quad (9),$$

where $k_{o,wt}$ defines the overall barrier height for the N to TS transition and ξ_{wt} is an average coupling factor for wt α LP. The average wt pK_a shift of $\Delta pK_{a,wt} = -0.87 (\pm 0.06)$, corresponding to a free energy of 1.32 (± 0.09) kcal/mol (Figure 2.5 (a), (b) and Supplementary Figure 2.1) and is larger than any of the shifts in the double salt bridge mutants, suggesting that the introduction of a third salt bridge to any of these mutants results in additional cooperative stability to each of the salt bridges. While it is formally possible that a drastically reduced coupling between one pair of salt bridges is paired with a significantly increased coupling between another pair, however, the very similar values obtained for all the pairs of salt bridges in the double salt bridge mutants makes this unlikely.

Number of salt bridges

To see whether the total number of salt bridges contributed to the overall stability of the mutants at pH 5, we introduced an additional salt bridge to the double salt bridge mutants recreating a NAPase intra-domain ion pair and maintaining the same number of salt bridges as wt α LP. We tested the double salt bridge mutant SB2+3 against Relo1 (Asn2 \rightarrow Asp, Arg105 \rightarrow Ser, Glu8 \rightarrow Ala) (Kelch, Eagen et al. 2007), SB1+3 against Relo2 (Val196 \rightarrow Arg, Arg78 \rightarrow Leu) and SB1+2 against Relo3 (Ser68 \rightarrow Arg, Ala55 \rightarrow Glu, Arg64 \rightarrow Ala, Glu182 \rightarrow Gln). At pH 5, the corresponding double salt bridge mutants and relocation mutants unfold at the same rate within error in the case of SB1+3 vs. Relo2

and SB1+2 vs. Relo3, while Relo1 is slightly stabilized relative to SB2+3 (data not shown). The relocation mutants also display pH dependent unfolding rate profiles similar to the corresponding double salt bridge mutants. While the total number of salt bridges may be important for overall pI and thermodynamic stability of the protein, the favorable interactions at the most crucial locations for unfolding contribute most markedly to the unfolding barrier.

Discussion

For α LP, the height of the cooperative unfolding barrier is the sole determinant of functional “stability”. α LP achieves kinetic stability with an extremely slow global unfolding and suppressed local breathing motions, thus making the global unfolding event, the only event to access unfolded states. Thus, in α LP, the native state is uncoupled from the unfolded states and hence, changes to the unfolded states are irrelevant for the unfolding transition of α LP.

The domain-domain interface is especially important in α LP’s unfolding as evidenced by the effect of salt bridge distribution in the acid resistant NAPase vs. α LP (Kelch, Eagen et al. 2007). Here we focus on the role of the inter-domain salt bridges in establishing α LP’s high unfolding barrier and its cooperativity. Further, we use these salt bridges as probes to the unfolding events in their environments to have a more detailed structural look into the unfolding transition.

As shown, all three inter-domain salt bridges in α LP have significant roles in unfolding. Of the three, salt bridge 3 dominantly determines the height of the unfolding barrier, thus it must contribute differentially to the stability of the TS and N state. If the separation of the N- and C- domains is a major event in α LP’s unfolding transition, eliminating inter-domain pH-dependent interactions would reduce or obviate pH sensitivity of the unfolding rate as in NAPase and Relo1. Salt bridge 3 contributes most

significantly to the acid sensitivity, as well. Combined, these findings suggest that the unfolding events in its environment are involved in α LP's unfolding transition.

The pH-unfolding profiles of double salt bridge mutants and wt α LP are best explained by a pair wise coupling between salt bridges. All three salt bridges interact with each other pair wise and cause a reciprocal downward shift in each other's TS pK_a 's, alluding to a cooperative behavior associated with spatially large-scale protein dynamics (see below). When the pairing is with salt bridge 1, the greatest degree of pKa shifts are observed, followed by salt bridge 2 and salt bridge 3. With all three together in wt α LP, the coupling is further increased as evidenced by a larger pK_a shift than those in any of the pair wise combinations of salt bridges.

A strategy for broad range pH stability

Of the three inter-domain salt bridges, salt bridge 3 is the dominant contributor to the height of the unfolding barrier. This salt bridge is also associated with a higher pK_a in the transition state suggesting a lower amount of structure in its vicinity in the TS ensemble. Interestingly, an extremely low upper limit is predicted for the native state pK_a associated with this salt bridge, lower than both of the other salt bridges, which suggests that salt bridge 3 contributes the greatest amount of free energy (>4.51 kcal/mol) to the unfolding barrier (Figure 2.3). Of the three single salt bridge mutants, the presence of salt bridge 3 contributes the largest amount of acid sensitivity at ~ 20 -fold (measured at pH

2.5 vs. pH 5). Correspondingly, its absence is associated with the greatest amount of acid stability at ~4-fold in SB1+2 compared to wt α LP (Supplemental Table 1).

A comparison of the NAPase and SB1+2 crystal structures around the salt bridge 3 locus shows no significant differences in terms of main or side chain positions and solvent exposure around the salt bridge (data not shown). Yet, while SB1+2 and SB2 have acid sensitivities similar to that of NAPase at pH 2.5 (~5-fold) (Kelch, Eagen et al. 2007) and SB1 is within 2-fold, the reduced acid sensitivity comes at the expense of destabilizing the proteins at pH 5, as well as across the pH range we tested (Figure 2.2 (a) and Figure 2.4 (a)). In contrast to NAPase, which unfolds an order of magnitude slower than wt α LP at 70 °C, pH 5 (Kelch, Eagen et al. 2007), SB1+2, SB1 and SB2 unfold ~10, 7 and 29 times faster than wt α LP at 60C, pH5 , respectively. In NAPase then, the kinetic stability lost through deletion of the strategically positioned salt bridge 3, is more than compensated for through changes elsewhere in its structure (the NAPase and SB1+2 structures align extremely well in the vicinity of salt bridge 3). In the mutants however, the substitutions R64->A and E182->Q, while being equivalents of these residues in NAPase, do not recapitulate the compensating stability found in NAPase.

Salt bridge 3 is centered around a large hydrophobic complex, which includes the N- and C-terminal domains, the domain bridge, as well as the α -helix and seems to act as a gatekeeper for further downstream unfolding events in computational studies on the separation of the N- and C-terminal domains (Bosco Ho, unpublished results). Comparison of the α LP and NAPase structures for differences suggests that it is likely

that enhanced packing interactions of a strengthened ‘domain bridge’ comprising residues 120A to 120L in NAPase relative to α LP (chymotrypsin numbering), are contributing to its large unfolding barrier while providing thermostability (Kelch and Agard 2007). If true, this would mean that NAPase replaces interactions that are broken in TS to ones that are less susceptible to pH in order to attain stability at a broad range of pH values. This strategy is in contrast to previously reported strategies employed by acidophiles: lowering of pI by an increased number of carboxylates (Cooper, Khan et al. 1990), (Fushinobu, Ito et al. 1998), (Bonisch, Schmidt et al. 2002), (Kashiwagi, Kunishima et al. 1997), or reducing the magnitude of the electrostatic effect all together by eliminating them (Schafer, Magnusson et al. 2004), (Settembre, Chittuluru et al. 2004).

Cooperative contributions of salt bridges

Even with the recent advances in understanding protein folding, the complex tertiary structures, as well as highly cooperative dynamics of globular proteins are hard to dissect energetically. While concerted folding and unfolding of globular proteins provides an uncomplicated, “two-state” folding landscape simplifying the analysis kinetic data, it prevents dissection of energetics to different structural elements and interactions between them, which is required to construct energy landscapes and cooperativity networks for real proteins (Kloss, Courtemanche et al. 2008).

The reciprocal TS pK_a shifts originating from the pair wise coupling of structurally distant salt bridges, provide a means to investigate the dampened long range

cooperative dynamics of α LP, as evidenced by hydrogen exchange protection factors of greater than 10^9 throughout a large hydrophobic core that spans both N- and C-terminal domains (Jaswal, Sohl et al. 2002).

α LP TS has been found to be compact, native-like, highly-specific and resilient. Catalyzed and uncatalyzed folding reactions of α LP point mutants were affected equivalently, even though the catalyzed reaction is more than nine orders of magnitude faster (Jaswal, Sohl et al. 2002), (Truhlar 2004), (Derman and Agard 2000), (Kelch, Eagen et al. 2007). Thus, an electrostatic interaction between pairs of salt bridges in the TS is extremely unlikely, as it is in the N state, due to the large distance separating them and the high structural similarity of TS to the N state (Jaswal 2000).

Instead, a useful way of thinking about this mostly entropic coupling free energy is through altered protein dynamics and the number of available states in the TS ensemble. Accordingly, the free energy associated with the pK_a shifts can be largely understood as arising from effects on the number of states available in the TS (entropy).

$$\Delta\Delta G = \Delta\Delta H - T\Delta\Delta S \approx -T k_B N_A \ln \Omega \quad (10),$$

$$\Omega \approx e^{-\frac{\Delta\Delta G}{T k_B N_A}} \quad (11),$$

where k_B is the Boltzman constant, N_A is the Avogadro constant and Ω is a count of total number of states in the ensemble representing a single molecule. Using Eqn. 4 and Eqn. 11, we can thus calculate the factor by which the number of states in the TS ensemble is reduced upon introduction of additional salt bridges (Figure 2.7). The more restricted the

TS ensemble, the harder for the protein to get to it, since it requires a rare concerted transition, i.e. cooperativity.

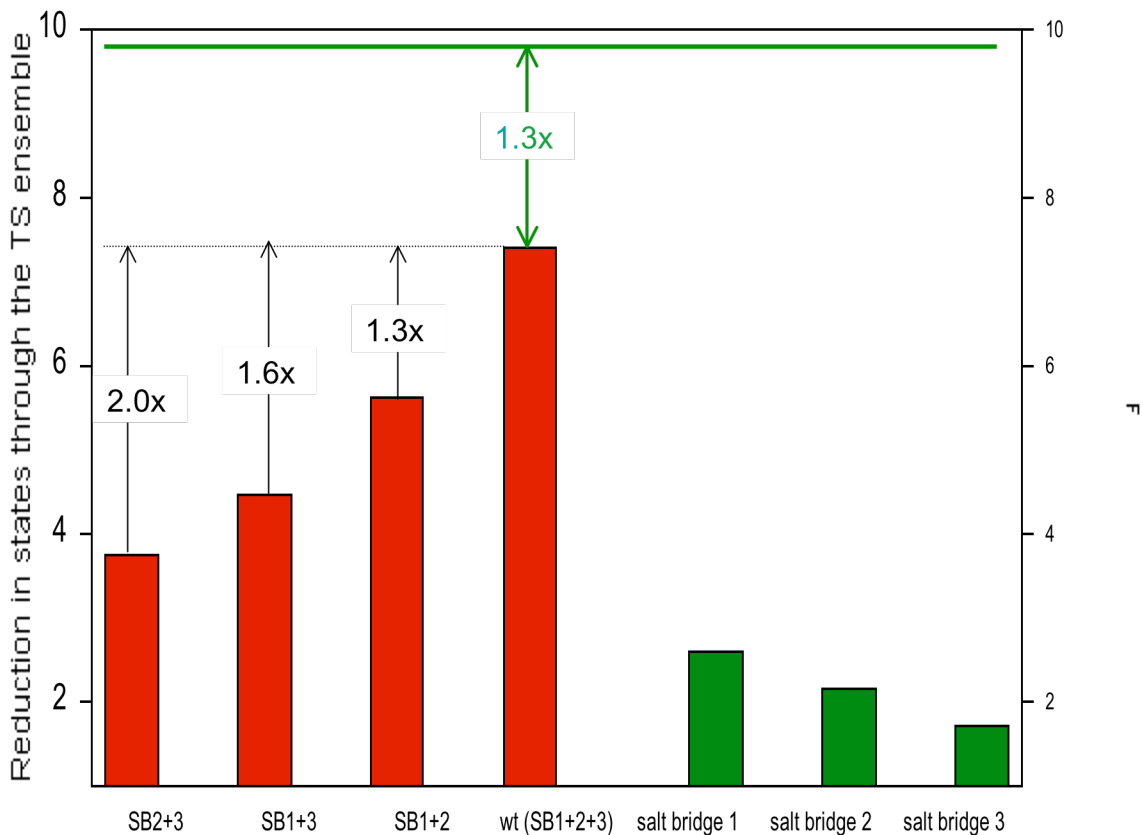


Figure 2.7 Cooperativity measured by reduction in number of states. In red, the reduction in number of states in the TS ensemble calculated for each mutant and wt α LP. In green, contribution of each salt bridge to the reduction in number of states is calculated using linear combinations of individual contributions describing the coupling energies in the double salt bridge mutants. In dashed arrows, fold difference in number of states in TS between double salt bridge mutants and wt α LP. Green line indicates total attainable reduction in number of states.

The order introduced by the presence of another salt bridge is highest for salt bridge 1, then salt bridge 2 and lowest for salt bridge 3. Together in SB1+2, these two salt bridges induce the greatest drop in pK_a on each other and hence a reduction of number of states of ~ 5.6 fold, compared to the other double salt bridge mutants involving salt bridge 3 (Figure 2.5 (a) and Figure 2.7). Put another way, addition of salt bridge 3 to SB1+2 contributes the smallest reduction in the number of states (~ 1.3) amongst all salt bridge additions to double salt bridge mutants. The lower structural freedom in the TS would make the unfolding transition more cooperative, since there are fewer possibilities for the protein to reach it.

SB1+2 unfolding barrier is drastically lowered at all pH's (Figure 2.4 (a)), which suggests that unfolding around salt bridge 3 occurs more frequently in the TS. While salt bridges 1 and 2 are not significantly contributing to the height of the unfolding barrier, the reciprocal effects caused by all of salt bridges on each other result in a highly concerted transition making the role of these two salt bridges extremely relevant for cooperativity, an essential component of kinetic stability. Thus, while only salt bridge 3 contributes to the overall height of the unfolding barrier significantly, all three have pronounced contributions to the cooperativity and kinetic stability of α LP.

The coupling energies are calculated from the reciprocal TS pK_a shifts according to

$$\begin{aligned}\Delta G_{coupling} &= -2.303RT\Delta pK_a \\ &= -1.52kcal/mol\Delta pK_a\end{aligned}\quad (12).$$

We express the TS coupling energies in the double salt bridge mutants as linear combinations of the individual salt bridge contributions, assuming the latter to be independent of each other, and calculate individual contributions of 0.63, 0.51, and 0.36 kcal/mol to the coupling energies for salt bridges 1, 2, and 3 respectively. These energies correspond to a reduction of ~ 2.59 , 2.17 and 1.73 fold in the number of states in the TS, respectively. All three together would provide a 9.72 fold reduction in states or 1.5 kcal/mol total coupling energy or 0.99 pK_a units (Figure 2.7). The pH titration data of wt α LP unfolding rates, on the other hand, predicts a 1.32 kcal/mol coupling energy, which corresponds to a 7.40 fold reduction in the number of states. α LP therefore seems very efficient in stacking up additional cooperative effects fairly linearly, failing to harness only 12% of the total attainable coupling energy (or 24% of total reduction in states). The difference likely indicates a slight overlap in what states are being frozen with the addition of individual salt bridges to the different positions in the domain-domain interface of α LP. The use of salt bridges to adjust characteristics of the unfolding transition, thus, seems very optimized in that the additions of each salt bridge provides cooperativity contributions close to the maximum in each context. The overlap in the frozen states is limited and the cooperativity in wt α LP is maximized by optimal coupling interactions.

In contrast, NAPase unfolding cooperativity is reduced compared to α LP as evidenced by its survival rate in the presence of other proteases (Kelch, Eagen et al. 2007). In order to attain acid resistance, NAPase avoids the use of acid sensitive interactions in the domain-domain interface and instead utilizes hydrophobic interactions

to keep its unfolding barrier high. Thus, NAPase's acid resistance comes at the expense of much lower cooperativity at neutral pH compared to wt α LP (Kelch, Eagen et al. 2007). While hydrophobic interactions can provide a broad area of interaction, electrostatic interactions in salt bridges can be more precisely positioned and can have more drastic effects in loop entropy. Whether the failure to attain as high cooperativity in NAPase is a result of the precise positioning of its inter-domain interactions or their different nature remains to be seen.

Structure in the TS ensemble and the cooperative unfolding transition

Previous studies reported that α LP has an unusually large entropic barrier for unfolding, suggesting exposure of hydrophobic surfaces without significant structural changes (Jaswal 2000). In addition, α LP possesses a large amount of structural similarity in the TS compared to the native state (Jaswal 2000). Original studies focusing on the role of salt bridges in the domain-domain interface suggested an involvement of changes in the domain-domain interface in unfolding (Kelch, Eagen et al. 2007).

Here, SB1, SB2 and SB1+2 show reduced sensitivities to acid (Figure 2.2 (b) and 4 (b)), which suggests that low pH affects the TS and N state similarly. In contrast, SB3 maintains a wt-like acid sensitivity indicating a differential effect of acid to salt bridge 3

in TS and N states and suggesting a structural transformation of the salt bridge 3 environment in TS.

It is certainly noteworthy that the salt bridges 1 and 2 have a profound effect on the cooperativity of the unfolding transition even though they don't undergo a substantial structural transformation. As such, weakening of the relatively native-like structures in the TS ensemble, through absence of the salt bridges 1 and 2, have rather global effects on the unfolding transition, by allowing a multiplicity of less concerted entries to the TS ensemble.

In α LP, a kinetically stable protein, alterations to structural elements in the U state are inconsequential for stability and longevity. Thus, α LP and its kinetically stable homologues can alter their response to acid by migrating acid sensitive elements that are unfolded in the TS, as well as those that determine the cooperative nature of this transition. Overall stability can be restored by strengthening the acid independent elements, such as the hydrophobic interactions between the "domain bridge" and the core of the protein. While migration of susceptible elements in the unfolding transition is studied here in the context of acid, kinetically stable proteins can easily generalize this strategy to reduce sensitivity to other environmental perturbations such as temperature, chemical denaturants and alkali. Evolutionarily, kinetically stable proteins have a rapid and flexible access to these alterations, since they are void of consequences to the U state.

Extraordinarily low pK_a 's linked to high protection factors

Previous studies found solvent accessible salt bridges to contribute modest amount of energy to the thermodynamic stability (<1kcal/mol) and unfolding transition (~5 fold) (Marti and Bosshard 2003), (Lyu, Gans et al. 1992), (Garcia-Arribas, Mateo et al. 2007), (Dao-pin, Sauer et al. 1991), (de Prat Gay, Johnson et al. 1994), (Horovitz, Serrano et al. 1990), (Sali, Bycroft et al. 1991), (Serrano, Horovitz et al. 1990), (Zubillaga, Garcia-Hernandez et al. 2006). Modest drops of less than 1 in native state pK_a relative to solution or TS pK_a are associated with such contributions. While the TS pK_a 's exhibit changes of this order relative to solution pK_a 's (Thurkill, Grimsley et al. 2006), we are only able to estimate an upper limit for the N state pK_a of the acidic groups at 1.57 for the C-terminal carboxylate in salt bridge 2 and 1.23 and 0.84 for the glutamate side chains in salt bridge 1 and salt bridge 3, respectively. These extremely low pK_a values are most likely linked to the extraordinary rigidity in α LP native state. The protection factors for α LP are among the highest measured in a protein and the most protected residues are spread throughout both domains instead of being concentrated in a 'core' region. (Jaswal, Sohl et al. 2002), (Huyghues-Despointes, Scholtz et al. 1999) This finding is further in line with the idea that the coupling free energies in TS between salt bridges are largely attributed to concerted reduction of the number of states in this ensemble.

In comparison, some of the lowest recorded pK_a 's are <2.1 and 2.2 for barnase Glu73 and insulin Glu13B for a glutamate and 2.4 and 2.38 for the C-terminal carboxylates in RNase and insulin Chain-B (Wisz and Hellinga 2003), (Li, Robertson et

al. 2005). Notably, two very low pK_a values (0.5 and 0.9) have been reported for buried aspartate side chains involved in polar interactions for Asp76 in RNase T1 and for Asp66 in hen-egg white lysozyme, respectively (Pace, Huyghues-Despointes et al. 2002), (Bartik, Redfield et al. 1994). Nevertheless, these low pK_a values in the native state are among the lowest for solvent exposed groups and can be attributed to the extremely high level of conformational rigidity demonstrated by the high protection factors.

Overall conformational rigidity is demonstrated by comparison of the crystal structures of α LP and its double mutants. Thermal parameters of wt α LP (PDB ID 1TAL) refined at 1.5Å and the two mutants, SB1+2 and SB2+3, when refined to 1.55Å (chosen because this was the highest resolution data available for the SB2+3 mutant) shows very low average all-atom B-factors of 6.73, 8.43, and 10.33, respectively. The B-factors are uncorrelated with the height of the unfolding barriers for these mutants, but are likely related to the quality of the crystals. Interestingly, local thermal motion of the mutated residues R64A/E182Q is reduced - as indicated by nearly 2x reduction in B-factors - relative to the protein as a whole, in the SB1+2 mutant compared to wt, while that of the E8A/R105S mutation in SB2+3 is increased 1.5x. Regardless of any variation in thermal parameters, rmsd's of the mutated residues in both structures are within 0.21Å of those seen in the wt structure. This strong alignment of salt bridges in wt and mutant structures indicates that it is extremely unlikely that the native state pK_a 's are higher in the double salt bridge mutants or wt α LP and this could not in any way be interfering with our data modeling. A further downward shift in native state pK_a 's is plausible going

from single to double salt bridge mutants and to wt α LP in parallel with the increased order the coupling of these salt bridges caused in the TS.

Materials and Methods

Crystallization and structure determination

Crystals were grown by the hanging-drop, vapor-diffusion method in 4 μ l drops containing 2 μ l of concentrated protein (15–20 mg/(Fuhrmann, Kelch et al. 2004)ml) and 2 μ l of precipitant solution (1.3 M lithium sulfate and 20 mM Tris–sulfate, pH 8.0). Hanging drops were equilibrated over 1 ml of precipitant solution for 24–72 hours prior to seeding. A microseed suspension was prepared, using the Hampton Seed Bead Kit™, from one small crystal in 50 μ l of precipitant solution. One to three crystals grew in drops seeded with 1 μ l of a 1:4,000 or 1:16,000 (v/v) dilution of the microseed suspension. Crystals grew to their maximum size following incubation for about five weeks at room temperature for SB1+2 and eight weeks at 4C for SB2+3.

Data were collected using a wavelength of 1.11Å at Beamline 8.3.1 at the Advanced Light Source (Lawrence Berkeley National Laboratory). Data were processed using HKL2000 (Otwinowski and Minor 1997). Molecular replacement was performed with Phaser (McCoy, Grosse-Kunstleve et al. 2007) using the 0.83Å X-ray structure of α LPha-lytic protease (PDB code: 1SSX) (Fuhrmann, Kelch et al. 2004) as a search model. Iterative cycles of refinement and model building were performed using Phenix (Adams, Afonine et al.) and Coot (Emsley and Cowtan 2004). Riding hydrogen atoms and anisotropic displacement parameters for all atoms were generated and written to the output in the SB1+2 structure, while only anisotropic displacement parameters for the protein atoms were generated and written to the output in the SB2+3 structure. The

structures were assessed using MolProbity (Davis, Leaver-Fay et al. 2007). Data processing and refinement statistics are summarized in Table 2.1.

Protein production

The mutants were created using a combination of the QuikChange Multi Site-Directed Mutagenesis and QuikChange Site-Directed Mutagenesis kits (Stratagene, La Jolla, CA). The SB2+3 mutant (Glu8→Ala, Arg105→Ser), SB1+2 mutant (Arg64→Ala, Glu182→Gln) and the SB1+3 mutant (Arg78→Leu) were created using the wt α LP plasmid as a template with primers, 5'-CATCGTCGGCGGCATC GCA TACTCGATCAACAACGC-3' (P11) and 5'- GTGTGCCGCTCGGGC AGC ACCACCGGTTACCAGTGCG-3' (P12), 5'- GTTCCCCGGCAACGAC GCC GCCTGGGTCAGCC-3' (P31) and 5'- CGCTCGAGCCTGTTC CAG CGTCTGCAGCCGATC-3' (P32), and 5'- GCAGACTCTGCTGCCG CTC GTGGCCAACGGCAGC-3' (P2), respectively. The SB2 (Glu8→Ala, Arg105→Ser, Glu182→Gln, Arg64→Ala) and SB3 (Glu8→Ala, Arg105→Ser, Arg78→Leu) mutants were created using the SB2+3 mutant plasmid as a template with primers P31 and P32, and P2, respectively. The SB1 (Glu182→Gln, Arg64→Ala, Arg78→Leu) mutant was created using SB1+2 mutant plasmid as a template with the primer P2. Mutant expression and purification were performed as described previously (Kelch, Eagen et al. 2007).

Unfolding

α LP and salt bridge mutant unfolding was monitored by fluorescence as previously described (Kelch, Eagen et al. 2007) (excitation 283 nm, emission 322 nm). The buffers (all at 10 mM) utilized were glycine (pH 2.0, 2.25, 2.5, 2.67, 2.75, 2.87 and 3.0), malate (pH 3.14, 3.25, 3.37, 3.5, 3.67, 3.75, 3.87, 4.0, 4.25, 4.5 and 5.0). All experiments were carried out at 60°C. The cuvette temperatures were checked using a thermocoupler. Unfolding rate constants were obtained by fitting data to a single exponential equation or with a single exponential plus a linear term.

Acid sensitivity and stability

Acid sensitivity was calculated as a ratio of a mutants unfolding rate at low pH with respect to its k_o , representing a baseline unfolding rate at neutral pH and obtained from fitting pH titration data (Supplementary Table 2.1).

$$\text{acid sensitivity} = k_{\text{low pH, mutant}} / k_{o, \text{mutant}}$$

Acid stability is calculated as a ratio of wt to mutant acid sensitivity.

$$\text{acid stability} = \text{acid sensitivity}_{\text{WT}} / \text{acid sensitivity}_{\text{mutant}} = (k_{\text{low pH, WT}} / k_{o, \text{WT}}) / (k_{\text{low pH, mutant}} / k_{o, \text{mutant}}).$$

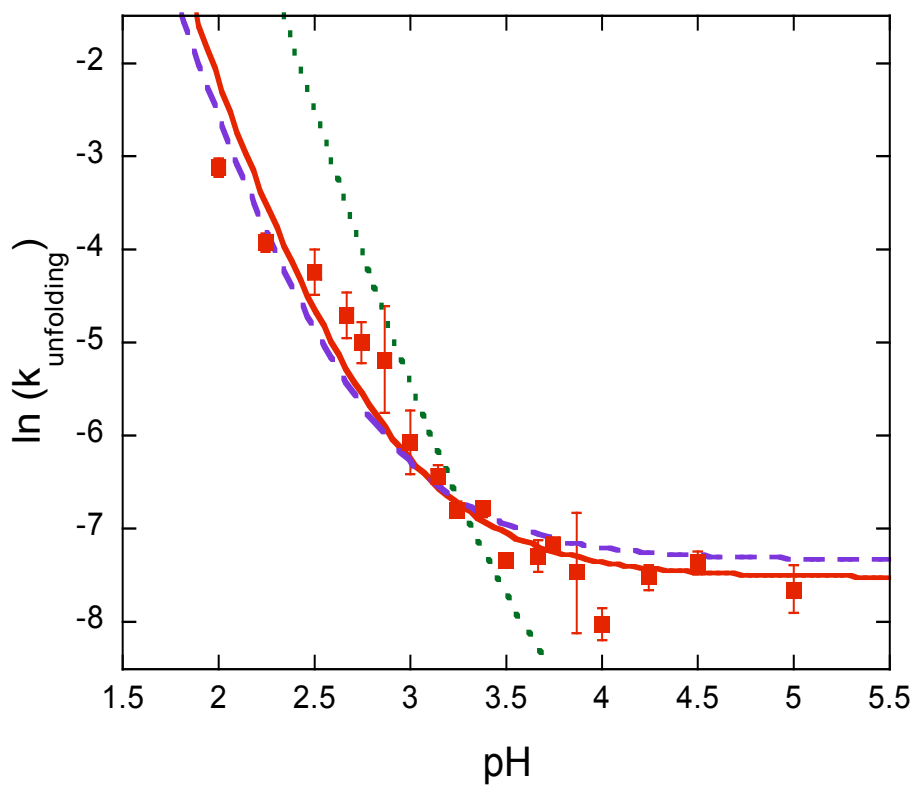
Testing for additional acid titratable groups

For the single salt bridge mutants, the presence of two acid titratable groups was assumed to be contributing to the N to TS transition titrating at disparate pK_a values

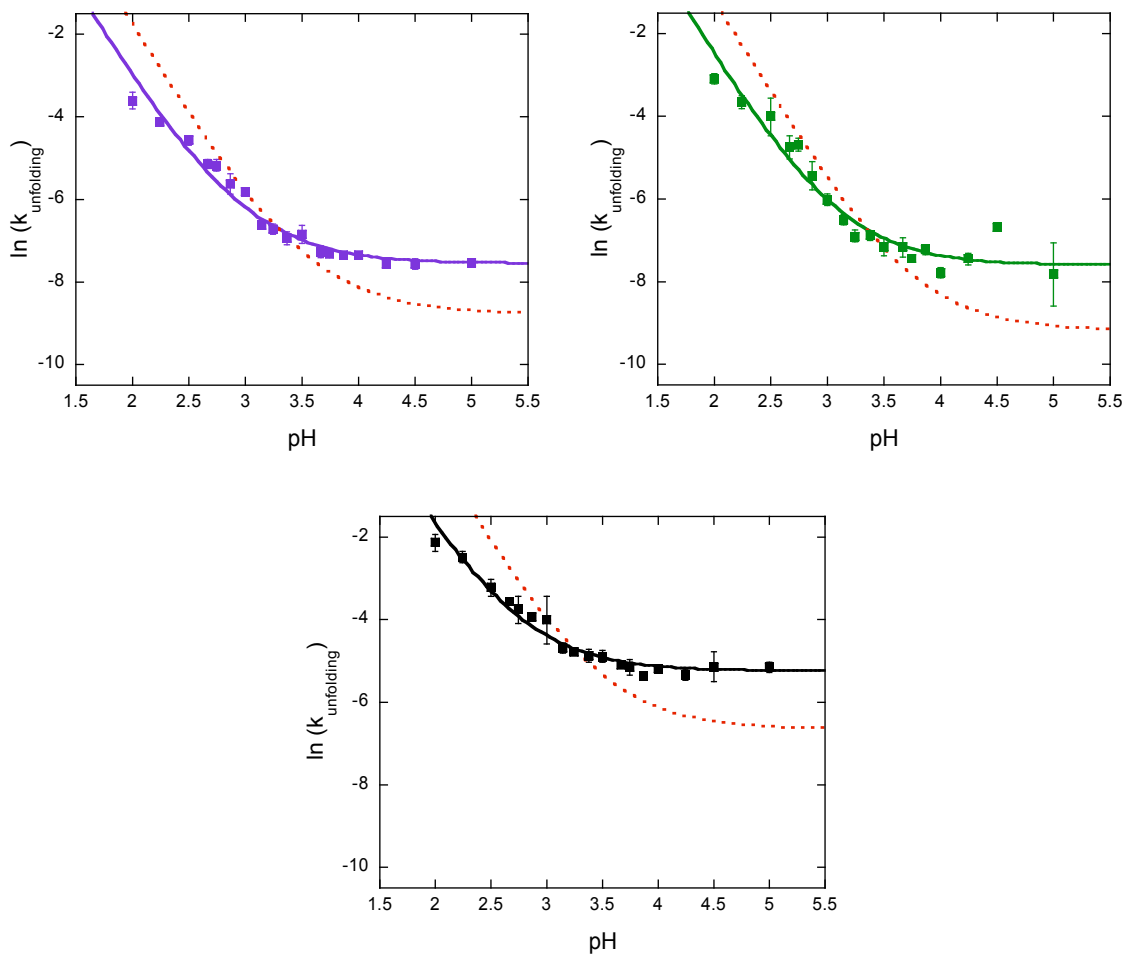
instead of only one for the single salt bridge in the domain-domain interface. The pH-dependent unfolding profiles were modeled according to

$$k_u \equiv k_o \left(1 + \frac{[H]}{K_{TS,1}} + \frac{[H]}{K_{TS,2}} + \frac{[H]^2}{K_{TS,1}K_{TS,2}} \right) \quad (5).$$

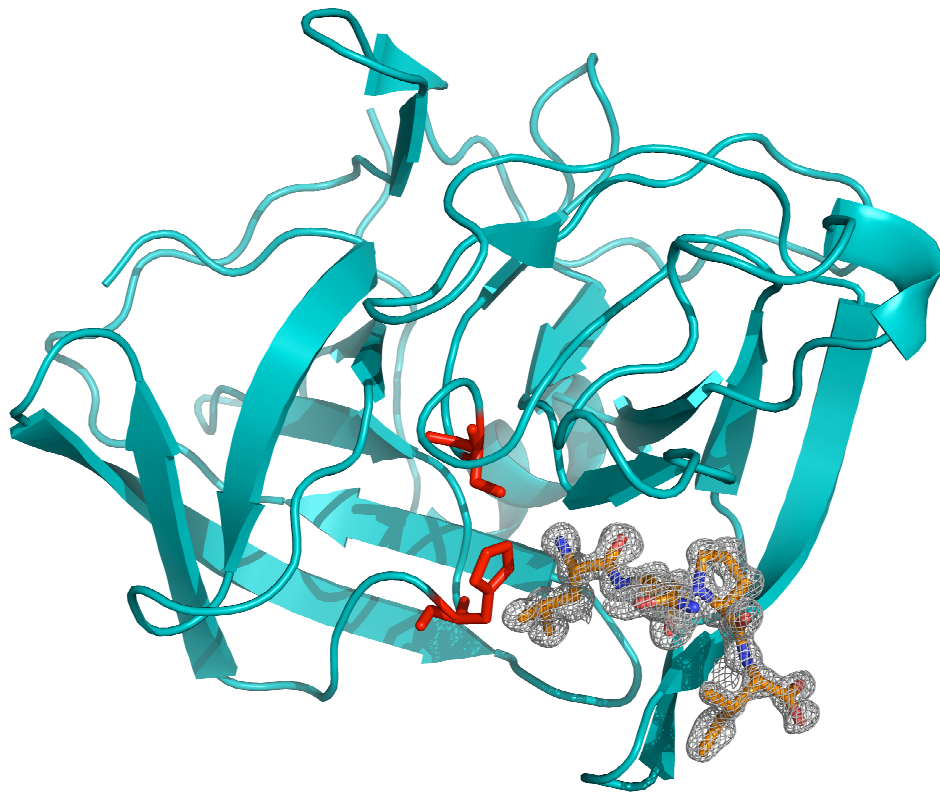
Of the two calculated pK_{TS} values from these fits, one for each mutant was within error the same as the one obtained by using Eqn. 3 (Set 1) and the errors, as well as the errors of fit were equivalent (data not shown). The second pK_{TS} was determined to be vastly different for each mutant: 2.42 (± 0.29), 1.43 (± 1.08), 1.72 (± 0.29), for SB1, SB2 and SB3, respectively (Set 2). The Set 2 values were assumed be the ones corresponding to a residual titratable group common to all single salt bridge mutants. The great variance in the Set 2 values for each mutant indicates that the double- K_{TS} model for the single salt bridge mutants is an inadequate interpretation and that there is only one acid titratable group (one of the three inter-domain salt bridges studied) contributing to the unfolding transition in these mutants. In addition, the values in Set 2 mostly outside the pH range of our data.



Supplemental Figure 2.1 Coupled pK_a fits of wt α LP pH-dependent unfolding rates. wt α LP pH titration of unfolding rates fit by either coupled TS pK_a 's (red, solid), fixed shifted TS pK_a 's determined using calculated cooperativity contributions of each salt bridge from double salt bridge mutant fits (purple, dashed), or fixed TS pK_a 's from single salt bridge mutant fits (green, dotted).



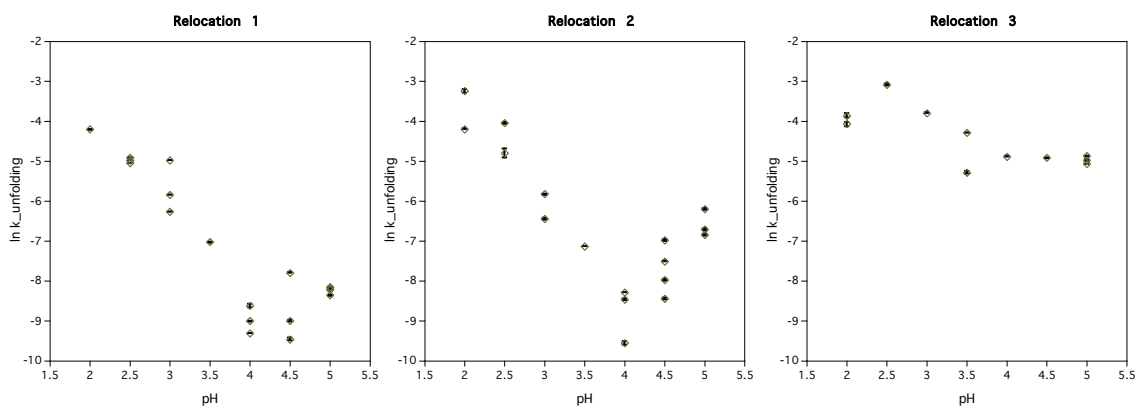
Supplemental Figure 2.2 Coupled pK_a fits of double titration mutants' pH-dependent unfolding rates. Double salt bridge mutant pH titration of unfolding rates fit by either coupled TS pK_a 's (solid), or fixed TS pK_a 's from single salt bridge mutant fits (dotted). (a) SB2+3, (b) SB1+3, and (c) SB1+2.



Supplemental Figure 2.3 Autolysis product in SB1+2 crystal structure. Bound peptide in the SB1+2 crystal structure matches the protein sequence between residues 184-187.

Supplemental Table 2.1 Overall unfolding rate and pK_{TS} values.

Mutants	k_o	$pK_{TS,1}$	$pK_{TS,2}$	$pK_{TS,3}$	Acid Sensitivity	Acid Stability
SB1	$3.69 (\pm 0.50) \times 10^{-3}$	3.66	NA	NA	11.8 (± 2.7)	2.5 (± 1.1)
SB2	$1.52 (\pm 0.12) \times 10^{-2}$	NA	3.22	NA	3.7 (± 0.8)	7.9 (± 3.3)
SB3	$3.71 (\pm 0.42) \times 10^{-4}$	NA	NA	3.8	19.5 (± 4.2)	1.5 (± 0.6)
SB2+3	$5.23 (\pm 0.53) \times 10^{-4}$	NA	2.65	3.23	19.5 (± 2.8)	1.5 (± 0.6)
SB1+3	$4.95 (\pm 0.76) \times 10^{-4}$	3.01	NA	3.15	45.0 (± 39.7)	0.7 (± 0.6)
SB1+2	$5.24 (\pm 0.37) \times 10^{-3}$	2.91	2.47	NA	6.8 (± 1.6)	4.4 (± 1.9)
wt α LP	$5.32 (\pm 0.81) \times 10^{-4}$	2.79	2.35	2.93	29.7 (± 10.6)	



Supplemental Figure 2.4 pH-dependent unfolding rates of relocation mutants.

Chapter 3. Investigating the role of substrate wall dynamics in

α LP k_{cat} effect

Preface

The crystal structures of α LP with boronic acid inhibitors were grown and solved by Cynthia Fuhrmann. The structures are listed on pg. 205-207 of her dissertation (Fuhrmann 2005).

Introduction

α LP is a chymotrypsin-like serine protease with a catalytic triad of Ser195, His 57 and Asp102 (Chymotrypsin numbering). The catalytic mechanism for chymotrypsin-like serine proteases involves two tetrahedral intermediates that are high energy and short lived (Fersht 1999); (Figure 3.1).

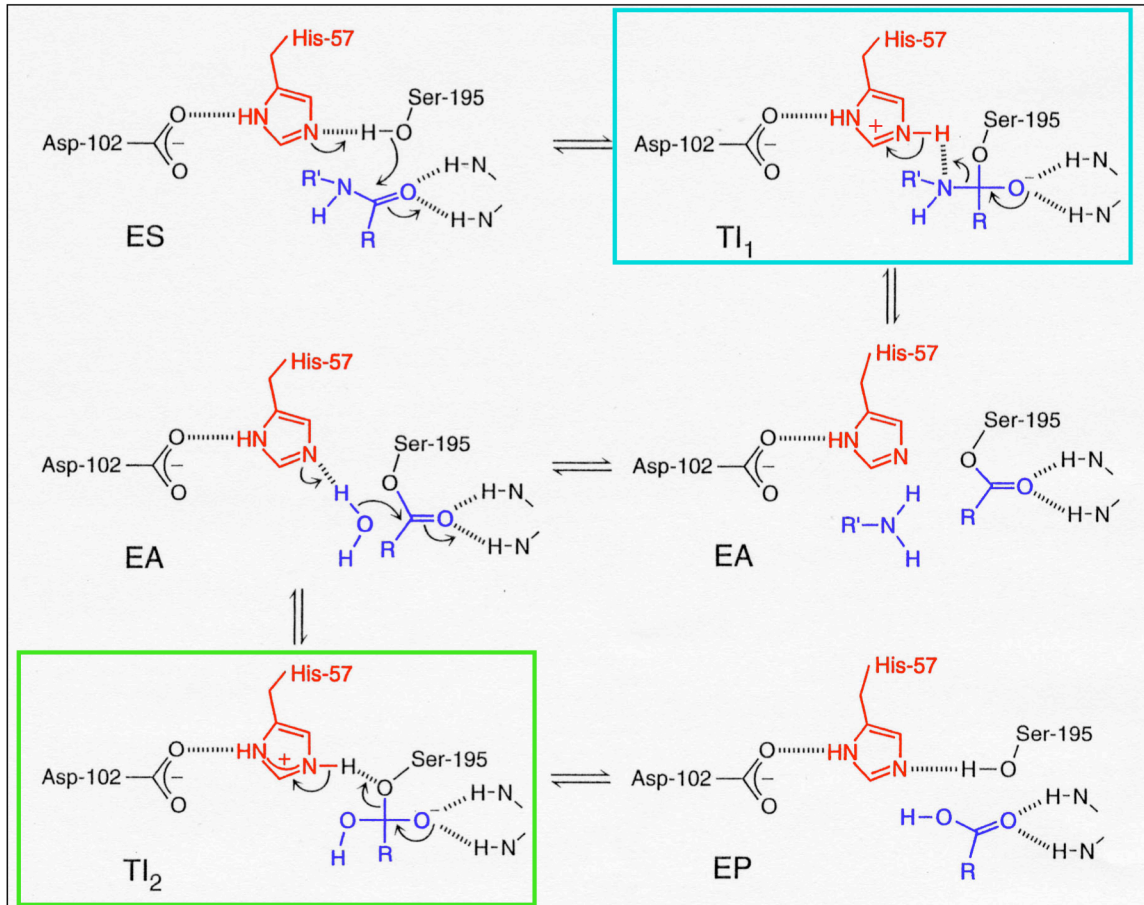


Figure 3.1 The general mechanism for catalysis by chymotrypsin-like serine proteases. In the first stage (acylation) active site serine – deprotonated by His- attacks the carbonyl for the substrate's scissile bond forming a metastable tetrahedral intermediate (TI₁). His donates the proton to the amide of the scissile bond allowing the release of the C-terminal substrate. During the second stage, a water molecule becomes the active nucleophile. Another tetrahedral intermediate (TI₂) forms prior to the release of the product. Figure modified from ((Fersht 1999)).

Peptidyl boronic acid inhibitor complexes were previously shown to mimic the tetrahedral intermediate 1 (Bone, Silen et al. 1989); (Figure 3.2).

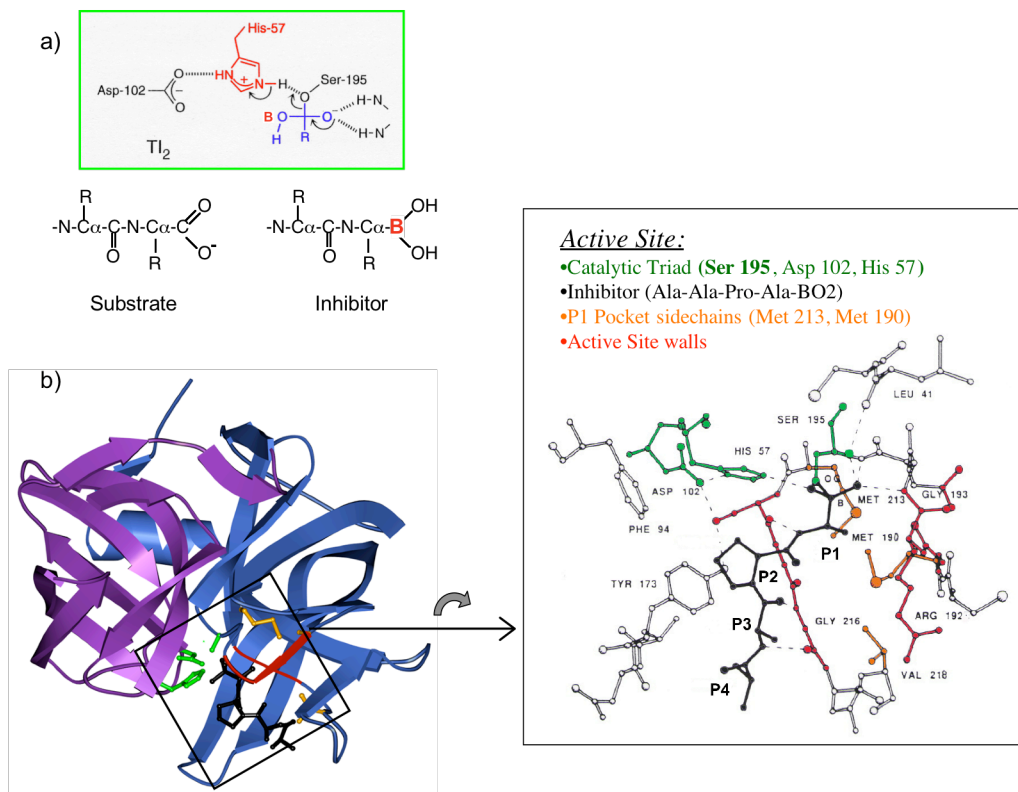


Figure 3.2 Structure of the α LP+boronic acid inhibitor complex. a) Boronic acid inhibitors are mimics of the tetrahedral intermediate TI_2 . b) Structure of the boronic acid inhibitor Ala-Ala-Pro-Ala-BO₂ in α LP active site. Reprinted with permission from (Bone, Silen et al. 1989).

In α LP the increase in catalytic efficiency (k_{cat}/K_m) upon lengthening of substrate is primarily due to a large increase in k_{cat} and not a large decrease in K_m , as might be expected (Bone, Ketner, and Agard, unpublished results, Table 3.1) The longer substrates of α LP display differences of several orders of magnitude k_{cat} but K_m values are only slightly affected. The K_i values for boronic acid inhibitors, which mimic high energy intermediates in the catalytic mechanism, further correlate with the length of the inhibitor, suggesting a k_{cat} effect. This raises the question: Does the binding of additional

residues alter the dynamic properties of the substrate-enzyme adduct, stabilizing the transition state?

Table 3.1 Kinetic parameters for substrates of varying chain length (α LP)

Kinetic Parameters for Substrates of Varying Chain Length (αLP)			
Substrate	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m
P ₃ P ₂ P ₁ ↓ P ₁ '			
suc-Ala-p-Na	0.013 ± 0.001	44 ± 4	0.28 ± 0.003
suc-Ala-Ala-p-Na	0.39 ± 0.02	35 ± 2	11.3 ± 0.1
suc-Ala-Ala-Ala-p-Na	32 ± 1.5	15 ± 0.7	2100 ± 32
Inhibitor	K_i (nM)		
P ₄ P ₃ P ₂ P ₁			
Acetyl-boroVal	722000	} Ultra-high resolution crystallography	
Acetyl-Pro-boroVal	3300		
Acetyl-Ala-Pro-boroVal	24		
m-Ala-Ala-Pro-boroVal	6.4		

(Bone, R., Kettner, C. A., and Agard, D. A. (Unpublished Results).

In chymotrypsin A_{α} , the increased substrate lengths were similarly explained by k_{cat} effects, rather than K_m (Bauer, Thompson et al. 1976), (Hill and Tomalin 1981), although the correlation is not as pronounced as in the case of α LP over the entire range of substrate lengths. By studying k_{cat} as a function of temperature, the activation enthalpies and entropies can be extracted (Hill and Tomalin 1981). The trend in ΔG^\ddagger is dominated by ΔS^\ddagger , rather than ΔH^\ddagger , therefore any interpretations of enzyme-substrate interactions must take into account the entropy term. Differential terms are calculated by

subtracting from each substrates energetic activation parameters the ones from a substrate that is one residue shorter. Addition of an interaction in subsite S3 in Ac-Gly_Phe-chymotrypsin A_α contributes a hydrogen-bond donated from NH of Gly-216 to CO of substrate acetyl group, and assists the catalytic process not by lowering ΔH^\ddagger , but presumably by ordering the acyl enzyme in such a way that only a small amount of further ordering is required to form the transition state. However, the additional ΔS^\ddagger and ΔH^\ddagger terms in chymotrypsin A_α do not have an orderly trend at each subsite. In combination with the weaker dependence of k_{cat} values in this enzyme to substrate length indicates that additional factors may complicate the interpretation of the data. In fact, the energetic parameters for substrate lengthening to fill further subsites have more meaningful trends for lengthening the leaving group positions of the substrate at P_1 , P_2 , and P_3 subsites in chymotrypsin A_α . In contrast, in α LP, the k_{cat} dependence on substrate length at P_1 , P_s , and P_3 positions is strong making this protease a great model to study the k_{cat} effect using boronic acid inhibitors.

To investigate how the binding of additional residues alter the dynamic properties of the substrate-enzyme adduct, stabilizing the transition state, we need to understand how the substrate binding energy is converted into transition state stabilization. In order to observe differences in inhibitor binding qualities as a function of inhibitor length we looked at the previously solved structures of three boronic acid inhibitor (Pro-boroVal, Ala-Pro-boroVal and Ala-Ala-Pro-boroVal) complexes with α LP (Fuhrmann 2005). The description of these crystal structures can be found in Cynthia Fuhrmann's dissertation pg.

205-207 (Fuhrmann 2005). The structural positioning of the inhibitor atoms around the active site were very close in all structures making an investigation on the protein dynamics necessary and meaningful (data not shown). The ultra-high resolution structures studied had excellent data to parameter ratio allowing the calculation of anisotropic B-factors.

Materials and Methods

For all residues with multiple conformations, the most abundant conformation was chosen. All three structures were of high enough resolution and thus had been refined with anisotropic B-factors.

$$f = f_0 \exp(-2\pi^2 h^T U h)$$

with

$$U = \begin{pmatrix} U^{11} & U^{12} & U^{13} \\ U^{21} & U^{22} & U^{23} \\ U^{31} & U^{32} & U^{33} \end{pmatrix},$$

where U is symmetric.

Since the vibrational modes of bonded atoms are not isotropic, the U tensor is a more accurate approximation to the actual scattering behavior of the atoms in a protein. The 6 independent components of the anisotropic U tensor define a probability distribution for the density. The contour of the distribution at a certain probability value is an ellipsoid. The eigenvalues of the U tensor are proportional to the axis of the thermal ellipsoids describing the dynamic behavior of the individual atoms. Anisotropy is a scalar descriptor related to the eccentricity of this ellipsoid.

$$\text{anisotropy} = \text{shortest axis} / \text{longest axis}$$

Since different protein crystals may have different static disorder and different internal vibration modes to compare anisotropy and thermal motion of corresponding atoms in two different structures, we have to normalize for overall protein anisotropy. An average difference in each six anisotropic factors was calculated for pairs of crystal structures.

$$\Delta B_{\text{avg}} = 1/N \sum_{i=\text{atoms}} (B_{1,i} - B_{2,i}),$$

where $B_{1,i}$ and $B_{2,i}$ are vectors of six anisotropic B-factors for each atom in crystal structures 1 and 2. The anisotropic B-factors were scaled by adding the difference to achieve equal average B-factors for all structures.

Anisotropy was calculated for each as the ratio of the smallest and largest eigenvalues of the U tensor. An “ellipsoid volume” was calculated as a descriptor of overall disorder at each atom from the product of the three eigenvalues.

All calculations were done in MATLAB.

Results and Discussion

The anisotropy distribution centered around 0.446, 0.431, and 0.396 for Ala-Ala-Pro-boroVal, Ala-Pro-boroVal, and Pro-boroVal complexed structures, respectively. (Figure 3.3) Only protein atoms were analyzed and hydrogens were excluded. The anisotropy distributions seemed relatively symmetric, with the Ala-Ala-Pro-boroVal structure slightly right shifted and the Pro-boro-Val structure slightly left. A comparison with anisotropy distributions from other anisotropically refined structures in the Protein Data Bank (Merritt 1999) (Figure 3.4) suggest a narrower distribution for the inhibitor bound structures.

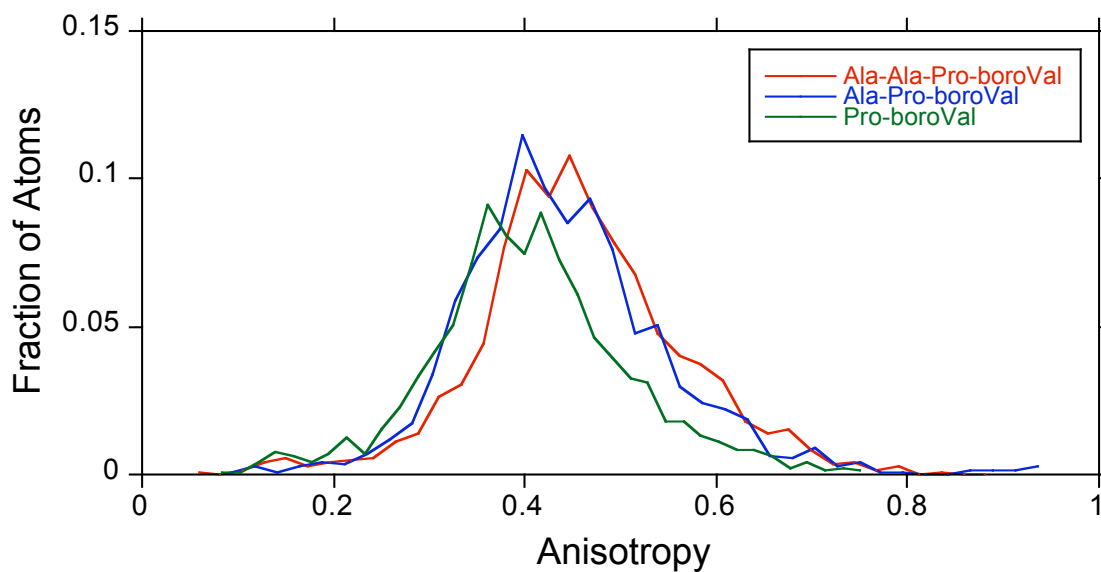


Figure 3.3 Anisotropy distributions in crystal structures with boronic acid inhibitors of varying length.

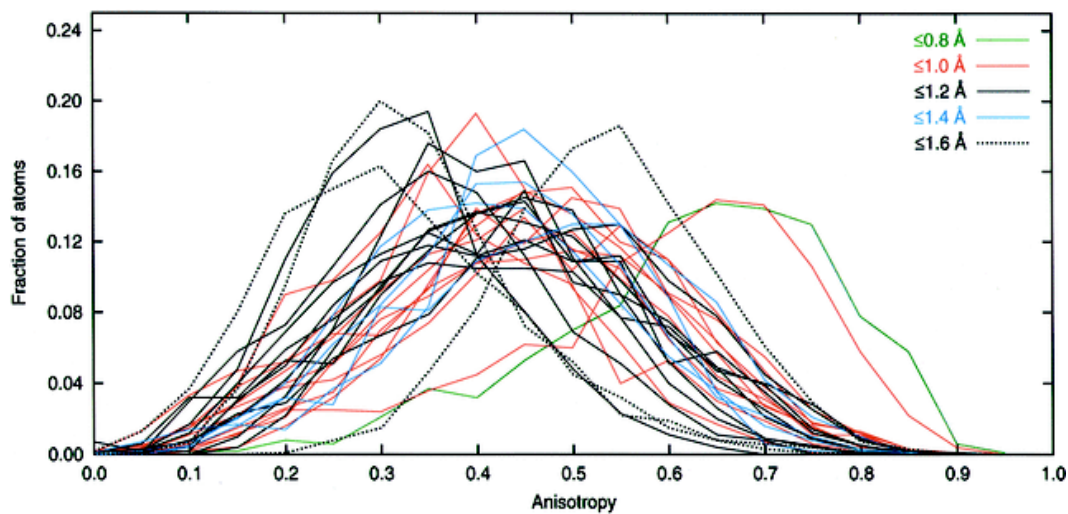


Figure 3.4 Anisotropy distributions of various structures with anisotropic B-factors in the Protein Data Bank. Figure reproduced from (Merritt 1999).

Overall, no significant correlation between anisotropy and increased thermal motion was observed in any of the structures, excluding the biggest outliers in anisotropy and ellipsoid volume. To test, the scalar products inverse anisotropy and “ellipsoid volume” vectors containing data for all atoms were calculated to be 0.493, 0.581 and 0.635 for the Ala-Ala-Pro-boroVal, Ala-Pro-boroVal, and Pro-boroVal structures, respectively. In contrast the products of the averages were 0.635, 0.560, and 0.927. The differences are less than a fraction of the standard deviation in the products of individual atoms (2.002, 1.450, and 4.506). The majority of the extremely high “ellipsoid volume” atoms had lower anisotropy values. Next, the atoms with inverse anisotropies or “ellipsoid volumes” more than 3σ over the protein average were excluded and the scalar products are recalculated. The scalar products inverse anisotropy and “ellipsoid volume” vectors were calculated to be 0.316, 0.322 and 0.336 for the Ala-Ala-Pro-boroVal, Ala-Pro-boroVal, and Pro-boroVal structures, respectively. In contrast the products of the averages were 0.320, 0.329, and 0.335. The differences are less than a fraction of the standard deviation in the products of individual atoms (0.240, 0.242, and 0.319) and the deviation is significantly smaller without the largest outliers.

Inverse anisotropy values were calculated for each of the three inhibitors within their respective structures, where larger values indicate further deviations from perfectly spherical dynamics (Figure 3.5). These values did not show a significant increase for atoms closer to the active site of the protein and did not vary significantly among the different length inhibitors excluding the possibility of length-dependent directional effect on the dynamics of the inhibitor itself. The atomic “ellipsoid volumes” showed a

significant and systematic increase with decreasing inhibitor length (Figure 3.6), suggesting an overall all-atom ordering with increasing inhibitor length. All three inhibitors had smaller “ellipsoid volumes” for atoms closer to the active site. There were no significant changes to the directionality of the ellipsoids around the boron (Figure 3.7).

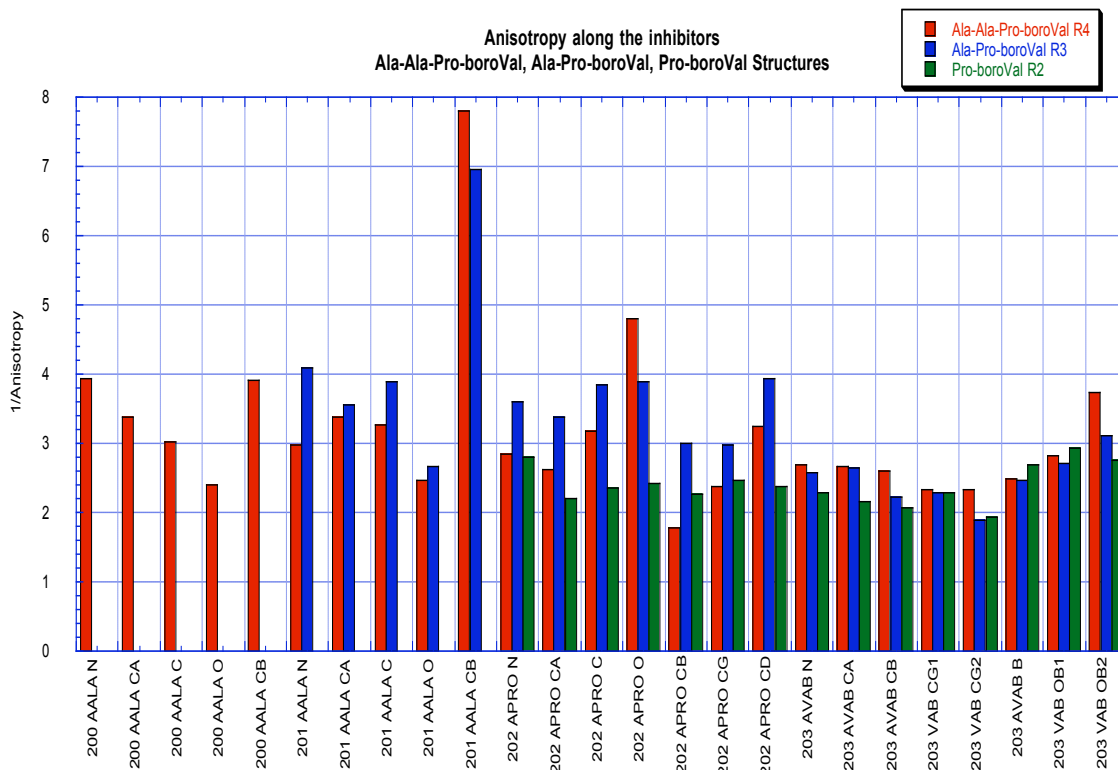


Figure 3.5 Inhibitor anisotropy plotted as inverse anisotropy.

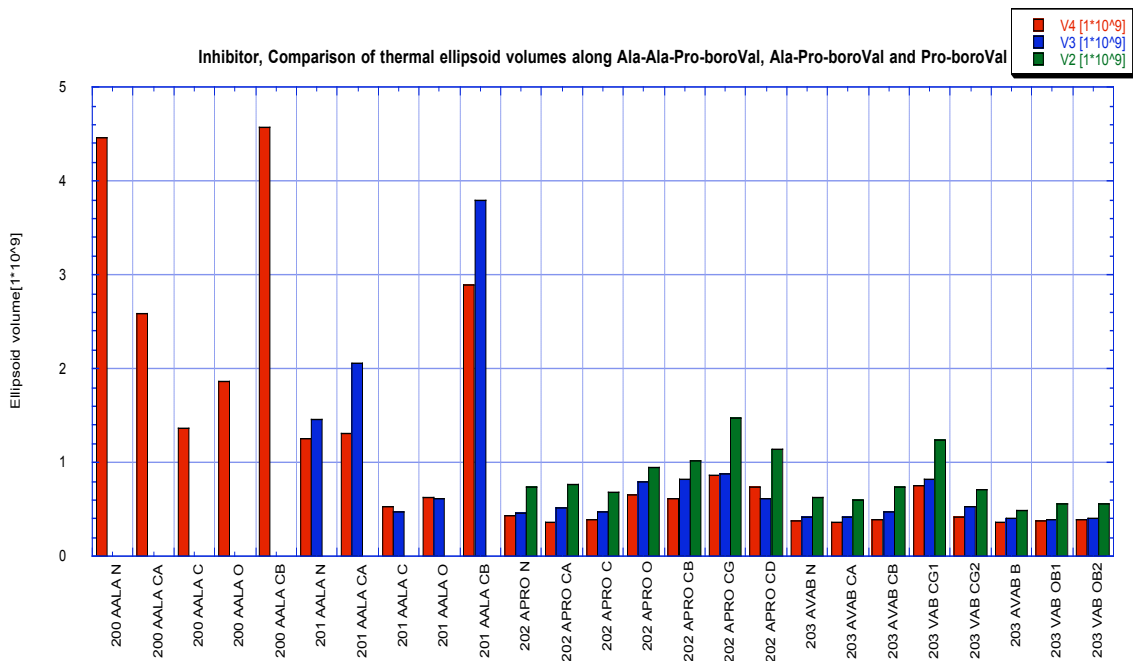


Figure 3.6 Anisotropic B-factor ellipsoid volumes representing the magnitude of thermal motion for boronic acid inhibitors of varying length.

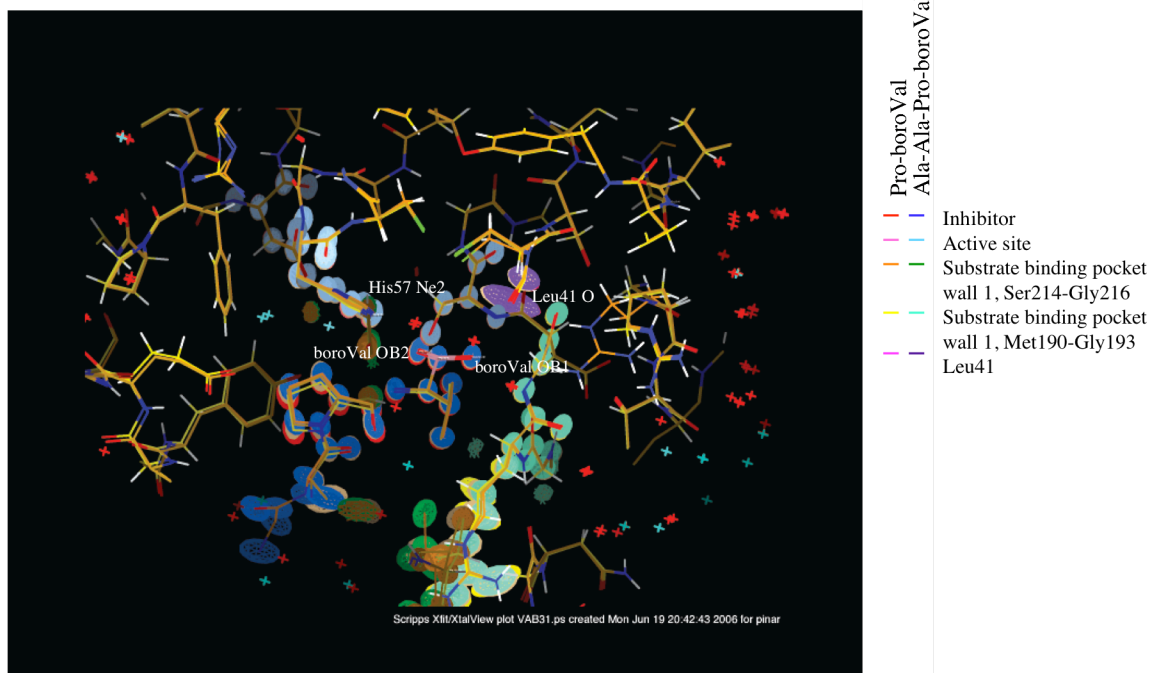


Figure 3.7 Thermal ellipsoids around the boronic acid do not show significant changes in size and directionality with inhibitors of varying length.

α LP was shown to have a concerted motion along the two walls of the peptide binding pocket and mutations affecting the peptide specificity were found to affect this concerted motion (Ota and Agard 2001). The first of these two walls comprises residues 214-219B (159-165 sequential). For the Pro-boroVal structure, a jump in asymmetric motion is observed at the point where the inhibitor ends (Figure 3.8). The raised value in inverse anisotropy is more than 2σ sigma above the average for this structure. This behavior might contribute to the extremely low k_{cat} value for the shortest peptide, but it doesn't set a length-correlated trend for longer peptides/inhibitors. The other wall is on the other side of the substrate binding pocket comprising residues 184-193 (135-141 sequential). It also includes Gly 193, which forms the oxyanion hole with Ser 195. The oxyanion coordinates the boron oxygen (carboxyl oxygen) of the boronic acid inhibitor

(peptide) (Figure 3.1, Figure 3.2). Arg 192 side chain aligns well with the tail of the Pro-boroVal and has a significantly increased asymmetry in the Pro-boroVal structure. The other structures seem to be prone for increased asymmetry at this position, but have this behavior restrained to under 2σ levels.

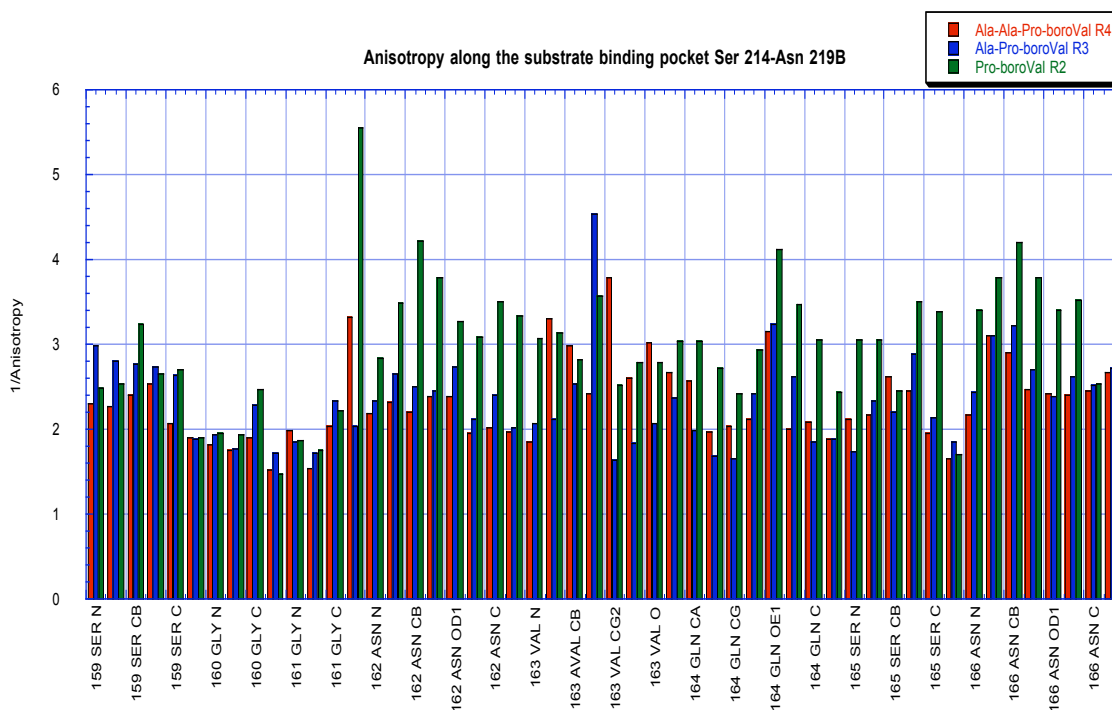


Figure 3.8 Anisotropy along the substrate binding pocket, first wall.

As opposed to the substrate pocket and the inhibitor, the catalytic triad does not sustain major changes in dynamics among the three structures (Figure 3.9). The “ellipsoid volumes” are lower than average for all three structures, which is in line with a previously identified hydrogen bond network stabilizing the catalytic triad (Fuhrmann,

Daugherty et al. 2006). Interestingly, catalytic triad residues display slightly lower values for the shortest inhibitor suggesting increased order, however the drop doesn't exceed the average for the entire protein by more than 2σ (Figure 3.10).

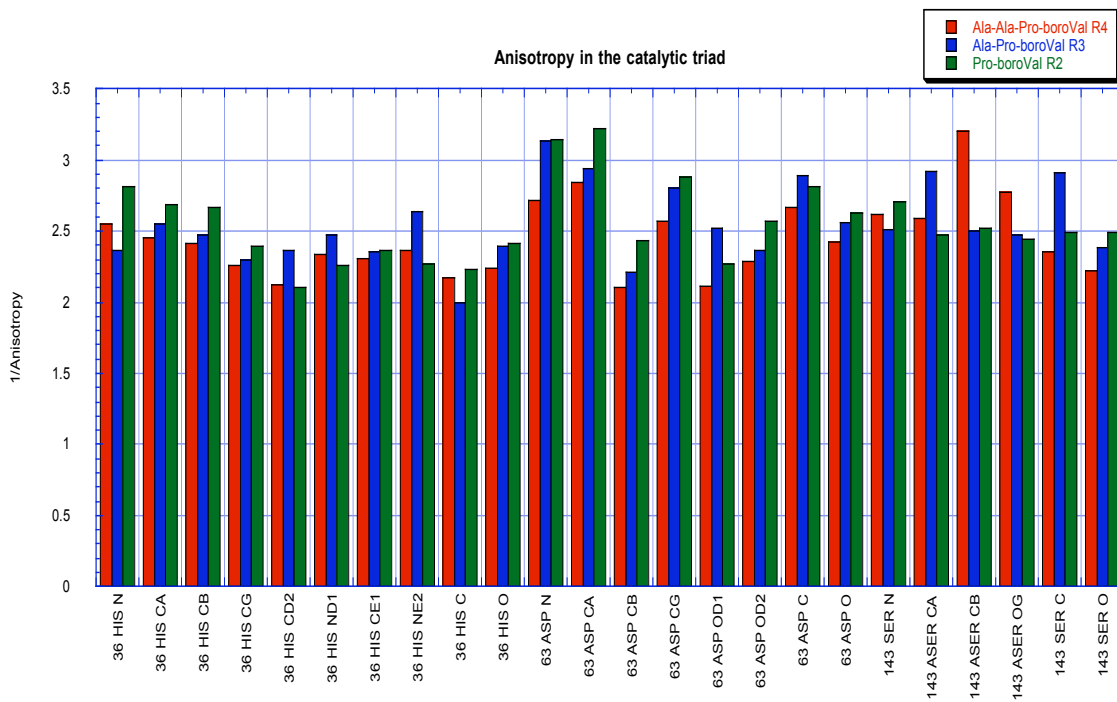


Figure 3.9 Anisotropy in the catalytic triad.

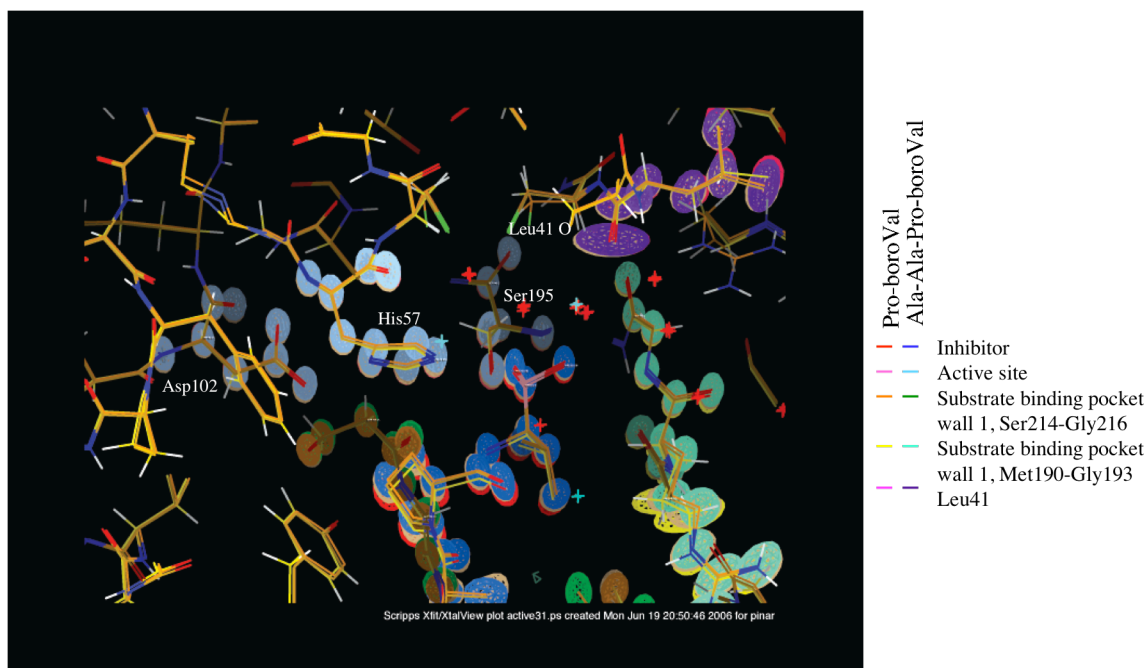


Figure 3.10 Thermal ellipsoids for the catalytic triad residues show reduced motion compared to the rest of the structure. Thermal ellipsoids of the Leu16 (41 in chymotrypsin numbering) side chain extending to the catalytic triad are anomalously large.

To explore the overall increase in thermal motion for the catalytically important atoms I calculated the product of “ellipsoid volumes” for the backbone nitrogen of Ser195 and Gly193 forming the oxanion hole, the side chain oxygen of the catalytic Ser195, the two His57 ring nitrogens, the two side chain carboxylate oxygens in Asp102 and boron and the two oxygens in the inhibitor in each structure. A higher product would indicate a higher overall disorder in the placement of the catalytically important atoms. The differences between structures were slight, yet in the contrary of a hypothesis of decreased disorder with increasing inhibitor length. In fact, the Ala-Ala-Pro-boroVal structure had 4.5 times the combined disorder for these atoms relative to the Pro-boroVal

structure. The Ala-Pro-boroVal structure had a 3.2 fold increase. In contrast, subdividing the catalytic and substrate binding pocket atoms into two groups reveals significant trends among the three structures. While, the product of the ellipsoid volumes for the binding pocket atoms in the Pro-boroVal structure is 81% greater than the Ala-Ala-Pro-boroVal structure, the product of the catalytically important atoms is 18% of that of Ala-Ala-Pro-boroVal, a 5 fold decrease.

The greatest anomaly both in anisotropy and in “ellipsoid volume” in the vicinity of the inhibitor is observed at Leu41 (Figure 3.10). For example, the C-delta atom has values about three times that of the protein average. This deviation is more than 2σ larger than the average for the overall protein and is unusual for the vicinity of the active site.

Conclusions

Can we explain the k_{cat} effect by increasingly ordered positioning of the inhibitor in the substrate binding pocket with inhibitor length? Some of the data indicated that there may be some differences in the ordering of the atoms of and around the inhibitor when the Pro-boroVal structure was compared to the other two structures. However, these effects did not exhibit a strong correlation with the inhibitor size beyond Pro-boroVal. The combined disorder for the catalytically most important atoms also did not exhibit a huge variation and had a contrary trend with inhibitor size. It is possible to have some ordering effect of the protein by the inhibitor in a size dependent manner for the smallest inhibitors slightly passed the oxyanion hole in the substrate binding pocket. However, at least one other mechanism must be playing a role for the k_{cat} effects for longer peptides.

The only significant trends were observed by studying in two groups the catalytically important atoms and the atoms important in placing the boron in the substrate binding pocket. A correlation was found with increased thermal motion in catalytic atoms and decreased thermal motion in the substrate binding pocket, and increasing inhibitor length. The latter with the higher overall disorder may provide an explanation for the k_{cat} effect with an increased order in catalytic transition states in the presence of longer substrates.

It is important to note that the crystal structures that were used in this study can be improved: The three structures that are discussed in this chapter were noted to have multiple conformations at Ser195-boroValLP1, which is likely to be due to radiation damage (Fuhrmann 2005). Thus their refinement was stopped at $R/R_{\text{free}} = 12.6/14.2$, $15.0/17.0$, and $11.6/13.0$, for the Pro-boroVal, Ala-Pro-boroVal, and Ala-Ala-Pro-boroVal complex structures, respectively. It is possible that any dynamic effects we would have seen in this study were hidden by the quality of the data. Collecting data with fresh crystal and utmost care to radiation damage might unveil further differences in the dynamics of the substrate binding wall.

**Chapter 4. A method to characterize nonnative states of
proteins like α LP Int using cross-linkers**

Preface

The design and synthesis of the in house cross-linker in this work is a combined effort of I and Naoaki Fujii. Naoaki Fujii was responsible for the synthesis of the carboxylic acid form of the cross-linker, which lacks the primary amine reactive NHS-ester group. Since the efficiency of this group was lacking in subsequent experiments, I took over responsibility to optimize this last step of the chemical synthesis of the cross-linker. The carboxylic acid form originally synthesized by Naoaki Fujii was handed to me in powder form..

Introduction

α LP is a serine protease of the chymotrypsin fold secreted by the soil bacterium *Lysobacter enzymogenes*. α LP is a 198 amino acid protein synthesized with a 166 residue pro region that is required for proper folding and secretion of the enzyme (Silen, Frank et al. 1989). This is in contrast to the mammalian homologs utilizing N-terminal zymogen peptides to prevent premature activation of the protease, but not requiring them for folding (Vasquez, Evinin et al. 1989). Refolding of α LP in the absence of Pro results in a stable folding intermediate (Int), which is prevented from reaching the native state by a very high free energy barrier resulting in a half life of $t_{1/2} = 1,800$ years for folding (Sohl, Jaswal et al. 1998). However, it has been shown that Pro dramatically catalyzes the folding reaction. Incubation of Int with different concentrations of Pro has allowed extracting values for k_{cat} and K_M , which indicate a lowered free energy barrier for folding. The energy diagram (Figure 4.1) has been completed by measuring affinity of Pro for the native state by inhibition studies and measuring the unfolding rate (Sohl, Jaswal et al. 1998).

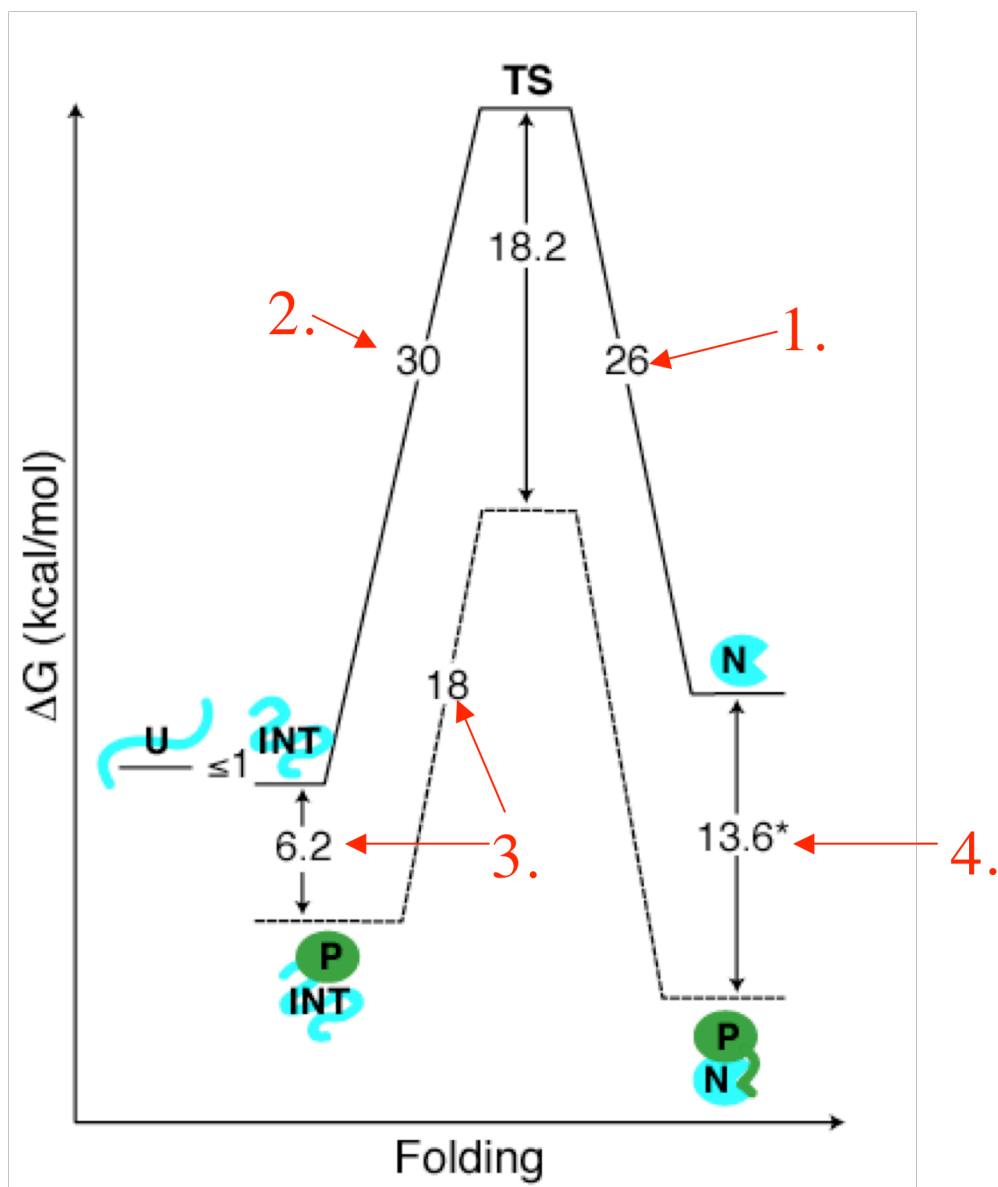


Figure 4.1 Folding landscape of α LP. 1. Rate of unfolding by fluorimetry or CD with GdnHCl and extrapolation to 0 M. 2. Rate of folding with a sensitive substrate. 3. Pro catalyzed folding, k_{cat} and K_M . 4. Inhibition of α LP by Pro, K_i .

Surprisingly, the native state is not at the free energy minimum, it is actually thermodynamically less stable than both fully denatured α LP and Int. Pro catalyzes the folding of α LP by lowering the high free energy barrier and by stabilizing native α LP over

Int. Once Pro is degraded, native α LP is kinetically stable, since the rate of unfolding is also extremely slow. Even though the energetics of catalyzed and uncatalyzed α LP folding are well understood, the reasons for the high energy barrier separating Int and native α LP remain unclear.

Int has structural characteristics between the native and denatured states. Its Stokes radius (28.4 Å) is between that of native (20.4 Å) and denatured protein (37.5 Å). Spectral deconvolution of the peptide region CD spectra of the native, intermediate and GdnHCl denatured states indicate nearly as much β -sheet formation in the Int (70%) as there is in the native (75%), whereas there is none in the denatured state (Baker, Sohl et al. 1992). Aromatic region CD and fluorescence spectroscopy data were not able to detect any tertiary structural organization in Int. There is great enhancement in the ANS fluorescence signal in the presence of Int compared to the unfolded protease in 8M urea and native α LP (Jaswal 2000).

Further characterization of Int showed that it is marginally stable compared to the unfolded state, which is the ensemble obtained by denaturing the protein with urea or GdnHCl (Sohl 1997). Urea denaturation of the Int followed by CD and ANS binding indicate low stabilities for Int (between -0.6 and 0 kcal/mol by ANS binding and $0.4 - 1$ kcal by CD) (Jaswal 2000), (Sohl, Jaswal et al. 1998). The m values for urea denaturation followed by the both methods closely match (1100 cal/mol*M). This would correspond to about burying 50% of the surface exposed in the unfolded state based on the calculated m value for complete unfolding from the native state (Jaswal 2000).

Following the secondary structure content by CD as a function of temperature in the range of 5-25°C shows that Int is stabilized with increasing temperatures indicating a possible role of hydrophobic effect in its stabilization. This observation is also supported by the decreased folding rate at 25°C compared to 5°C (Sohl, Jaswal et al. 1998).

The presence of extensive secondary structure, along with the absence of tertiary organization, as well as expanded hydrodynamic radius and ANS binding characteristics are hallmarks of the molten globule states of proteins often seen under nonnative conditions. Int is however a long-lived molten globule that is stable under native conditions (Baker, Sohl et al. 1992).

Int shows a very low solubility profile. It is shown to be monomeric between 1-10 μM (Sohl, Jaswal et al. 1998), but starts to precipitate at higher concentrations. Int's tendency to aggregate can be suppressed by keeping the protein at low temperature and in low ionic strength buffer (Baker, Sohl et al. 1992).

A current model for αLP folding suggests that αLP 's N- and C-terminal domains associate early in the structuring of Int, independent of Pro foldase activity. The interdependent, concerted folding of the N- and C-terminal domains appear to be the rate limiting step of the catalyzed folding reaction (Cunningham 2003).

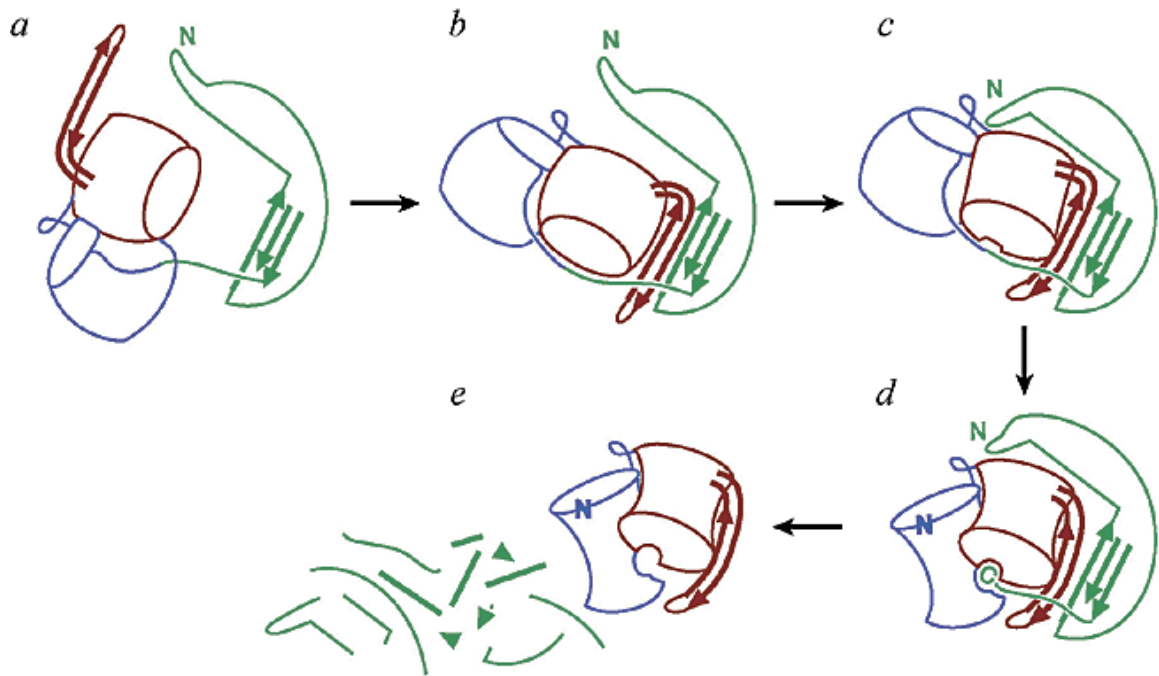


Figure 4.2 Revised model of Pro-catalyzed α LP folding. (a) The pro domain (green) of the Pro- α LP precursor folds separately from the N- and C-terminal domains of α LP (blue and red, respectively), which associate with one another to form substantial local secondary structure. (b) The three-stranded β -sheet from the Pro C-terminal domain pairs with the β -hairpin of C α LP to form a continuous five-stranded β -sheet. (c) Both N- and C-terminal domains of Pro bind to C α LP to help arrange key structural elements, enabling (d) N α LP and C α LP to simultaneously fold. Interdependent folding of the α LP domains completes the protease active site, which can then process the Pro- α LP junction. The new N-terminus of α LP repositions to its native conformation, while the Pro C-terminal tail remains bound to the α LP active site, inhibiting protease activity. (e) Intermolecular cleavage of Pro by α LP, or other exogenous proteases, leads to the degradation of Pro and the release of mature, active α LP (Cunningham and Agard 2003).

In the native α LP \blacklozenge Pro complex a three-stranded β -sheet of Pro interacts with the 118-130 β -hairpin to form a five-stranded β -sheet. Further, the presence of this β -hairpin is directly correlated with the presence of a Pro region (Sauter, Mau et al. 1998).

A two-domain model for Int has been suggested where the domains would each

collapse into separate hydrophobically stabilized molten-globule-like states (Cunningham and Agard 2003). On their own, or mixed in solution, the covalently untethered individual domains show reduced ANS binding and are predominantly random coil, while being able to partially shield tryptophan residues (Cunningham and Agard 2003). Thus, the two domains rely on each other to form significant amounts of structure in Int, drawing attention to the importance of the domain-domain interface for folding. The partial tryptophan protection is consistent with individual cores found in the both domains. However, in native α LP the extremely high protection factors observed in native state hydrogen exchange experiments are not isolated, but are spread across the two domains, unlike most proteins (Jaswal, Sohl et al. 2002), (Li and Woodward 1999). Moreover, cooperative interactions identified by molecular dynamics unfolding simulations show cluster patterns that are connected across domains in α LP, as opposed to isolated clusters in trypsin (Salimi, Ho et al.); (unpublished data). Thus, the concerted nature of α LP's folding and unfolding transitions all point that the two domains should not be considered in isolation and that interactions both within and across the two domains are important.

Covalently untethering the two domains also caused the domain bridge β -hairpin to be disrupted. The domain bridge forms a network of hydrophobic interactions with the both domains that are important for the unfolding transitions (Bosco Ho, unpublished results). Salt bridge 3 studied in Chapter 2 is located near this network and is disrupted upon unfolding. Thermophilic homologs of α LP are shown to have evolved an increased surface area for the interactions between the domain bridge and the rest of the protein (Kelch and Agard 2007). Thus the domain bridge may be important in overcoming Int's challenge of concerted folding.

The degree to which the Int structural elements are native is ambiguous, but it is an

essential piece of information for understanding Int stability. Similarly, it is important to infer what kinds of interactions play a role in the high energy barrier for the uncatalyzed folding reaction from Int to the native state. Structural characterization of Int would help to identify the transitions that must occur during the folding reaction and may help elucidate the basis for the high energy barrier in the folding direction.

The information about the structural features in the Int ensemble is going to be also very valuable considering that α LP Int is the only known example of a monomeric long-lived molten-globule-like state that is more stable than the native state. The recombinant mouse prion protein is shown to fold either to a monomeric α -helical or oligomeric β -sheet-rich isoform (Baskakov, Legname et al. 2001). The native α -helical isoform is less stable than the monomeric β -sheet-rich isoform and hence folding to the native isoform is under kinetic control. In addition, the conformational transition from the α -helical to the β -rich isoform is associated with a high energy barrier, suggesting a high degree of similarity between the β -rich prion isoform and α LP Int.

The low solubility of Int (1-10 μ M) makes the application of higher resolution spectroscopic techniques such as NMR difficult. Thus, here I develop a method for obtaining distance constraints by cross-linkers to map a higher resolution structure to Int. A similar approach has been used for protein fold determination with Fgf-2 (Young, Tang et al. 2000).

The method involves intramolecularly cross-linking primary amines on α LP to other points within a distance determined by the cross-linker length (Figure 4.3), which can be varied to get information at different distance ranges. The non-specific photoactive groups on the other end of the cross-linkers are designed to expansively sample pairs of distances

from lysines to other points throughout the protein. The cross-linked peptides are then analyzed by mass after a proteolytic digest to provide distance constraints.

The wild type α LP sequence contains three primary amines: N-terminus and two lysines. However, the additional 12 arginines can easily be mutated to lysines one at a time to provide additional anchor points.

This method is not without challenges. First, the unselective reaction of photoactive groups also translates to an incredibly long list of molecular ion species with inversely proportional abundance of each species. To overcome this challenge, trifunctional cross-linkers with an affinity tag were designed and synthesized in-house (Figure 4.3). Second, it is extremely important to know how the cross-linker distances correspond to their behavior in real proteins. To validate the method native α LP was used, as well as bovine insulin, which offered a smaller size system to work with. Third, it is crucial that most of the time the cross-linkers react intramolecularly with the protein and not with water. The experiments were designed keeping the protein at a high monomer ratio, whenever the sample was to be in light.

With all its challenges, this technique is likely to be extremely valuable in analyzing folding intermediate structures, which cannot be accomplished by other methods. In addition, it can be used as an effective tool for determining low resolution structures of native states. It is very suitable to the particular case of Int taking into account the additional problem of solubility it poses. On the other hand, being a long-lived molten-globule-like state, Int provides a remarkable system to establish this technique.

A further promise of this technique in general involves the possibility to apply it in a

time-resolved manner to kinetically study the folding/unfolding reactions at a higher resolution. The method is also founded on collecting the distribution of distances instead of average values as it would be the case with spectroscopic techniques and requires.

Results

Use of heterologous bifunctional cross-linkers with chemical and photoreactive groups involves the original labeling of the protein using the chemical group, followed by photoactivation. In order for results to be easy to interpret, it is essential to control the number of cross-linkers per protein molecule. Originally, the labeling reaction was optimized using NHS-PEO4-Biotin, a biotinylation agent with a similar reaction towards primary amines as the cross-linkers used in this study. Interestingly, native α LP was only labeled when 0.5 M GdnHCL was included in the reaction mixture. This finding points to the extreme rigidity of α LP's native state dynamics as observed with high hydrogen exchange protection factors of greater than 10^9 throughout a large hydrophobic core that spans both N- and C-terminal domains (Jaswal, Sohl et al. 2002).

Optimizing a protocol to detect photocross-linked peptides using native α LP proved to be complicated. The LC-ESI-TOF-TOF scan of SFAD cross-linked α LP resulted in an abundance of peaks, identifying 320 parent ions. There were 180 additional low-intensity ions in the no cross-linker control. Attempts to sequence ~20 of these peptides using a variety of collision energies failed due to low signal.

In order to enhance the abundance of cross-linked peptides, while keeping the functionalities of the cross-linking arm, a trifunctional cross-linker was designed (Figure 4.3). The cross-linker has a sulfonated NHS for increased solubility and a tetrafluoro-arylazide for efficient photoactivation. The design features a third functional arm with

desthiobiotin for affinity purification. Desthiobiotin has a lower affinity for avidin relative to biotin to allow for easy elution upon purification. Isotopic labeling between the cross-linking moieties can easily be incorporated and the linker length can be adjusted by the addition of ether groups, which in turn lowers hydrophobicity for better solubility in aqueous buffers. The calculated ClogP for the molecule is 2.09.

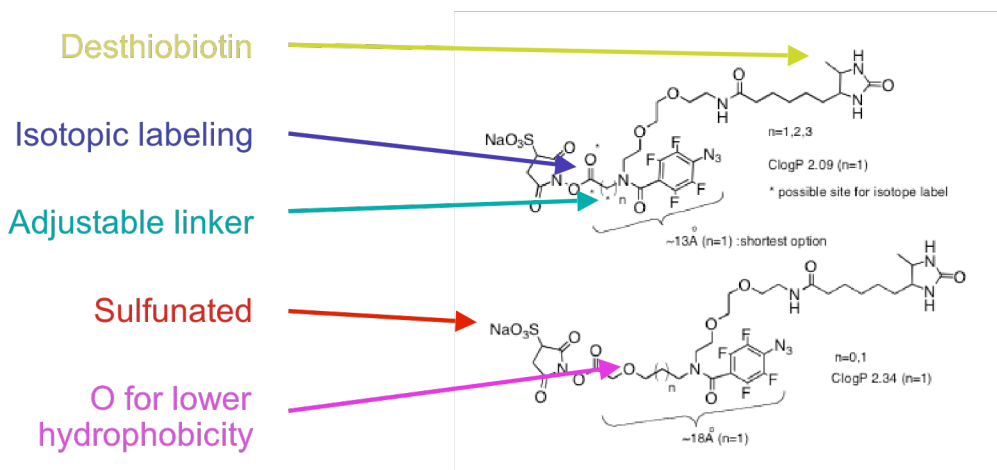


Figure 4.3 Cross-linker design. Cross-linkers that are best suited are specific on one end, non-specific on the other end, soluble, rigid, have good cross-linking yield, can be enriched for and are easily detectable. Desthiobiotin has lower than biotin affinity for strepdavidin for ease of purification. Isotopic labeling allows better mass spectrometric detection. The linker length can be adjusted to span different distances. Sulfonated NHS_ester and hydrophilic groups in the linker increase solubility.

It also proved necessary to optimize the use of this new cross-linker in a less complicated system. Bovine insulin was our choice for its small size and commercial availability.

A 2-3 day protocol on Zn-free bovine insulin including the last step of cross-linker synthesis was developed addressing the following crucial requirements: Efficient

primary amine labeling, elimination of excess cross-linker, monomeric protein at the time of photoactivation, efficient photoactivation, digest with small peptide size and purification of cross-linker labeled peptides.

Cross-linker activation

Since the final form cross-linker is unstable, its synthesis was carried out until the last step and the intermediate product was stored as a powder. The final step of EDC activation was then performed immediately before the use of the cross-linker to increase shelf life. While commercial NHS-esters readily labeled insulin (Figure 4.4 (a)), our in-house cross-linker failed in significant protein labeling, even at a 40X molar excess in DMF (Figure 4.4 (b)). To improve the abundance of active cross-linker the EDC activation step was optimized. A time course of activation followed by Insulin (1.35 mg/ml) labeling with 35 molar equivalents of cross-linker in DMF showed that there was a significant amount of active cross-linker after 1 hr of EDC activation (Figure 4.5).

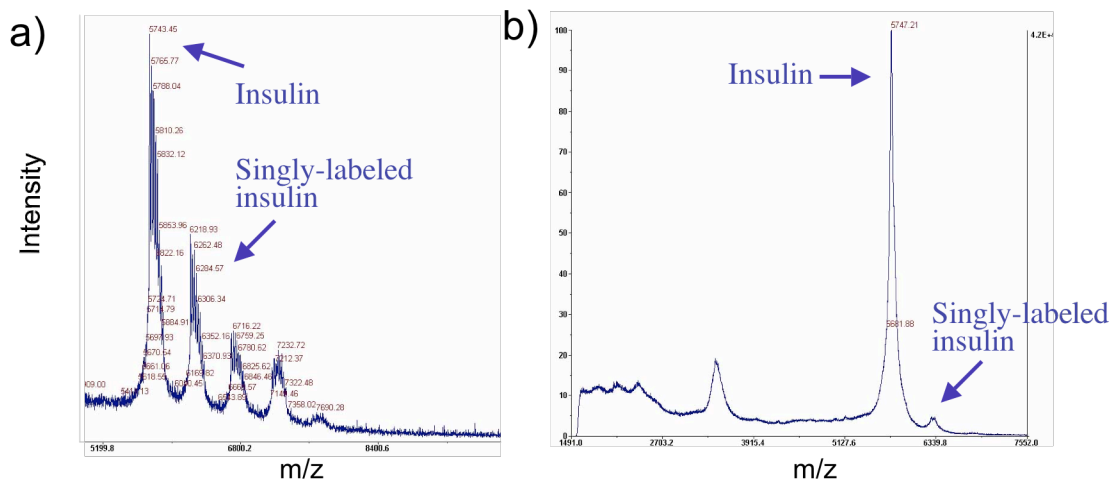


Figure 4.4 Labeling of bovine insulin with NHS-esters. a) Labeling with NHS-PEO4-Biotin in 10 mM Na Carbonate buffer, pH 9/150mM NaCl at RT was stopped with 100 mM Tris, pH 8.5 after 1-2 hrs. b) 3 mg/ml Insulin (0.5mM) was mixed with 21 mM (estimated) cross-linker and 3.5 mM triethylamine in DMF and stirred 1-2 hrs at RT in dark followed by addition of acetic acid to 1M to stop the reaction.

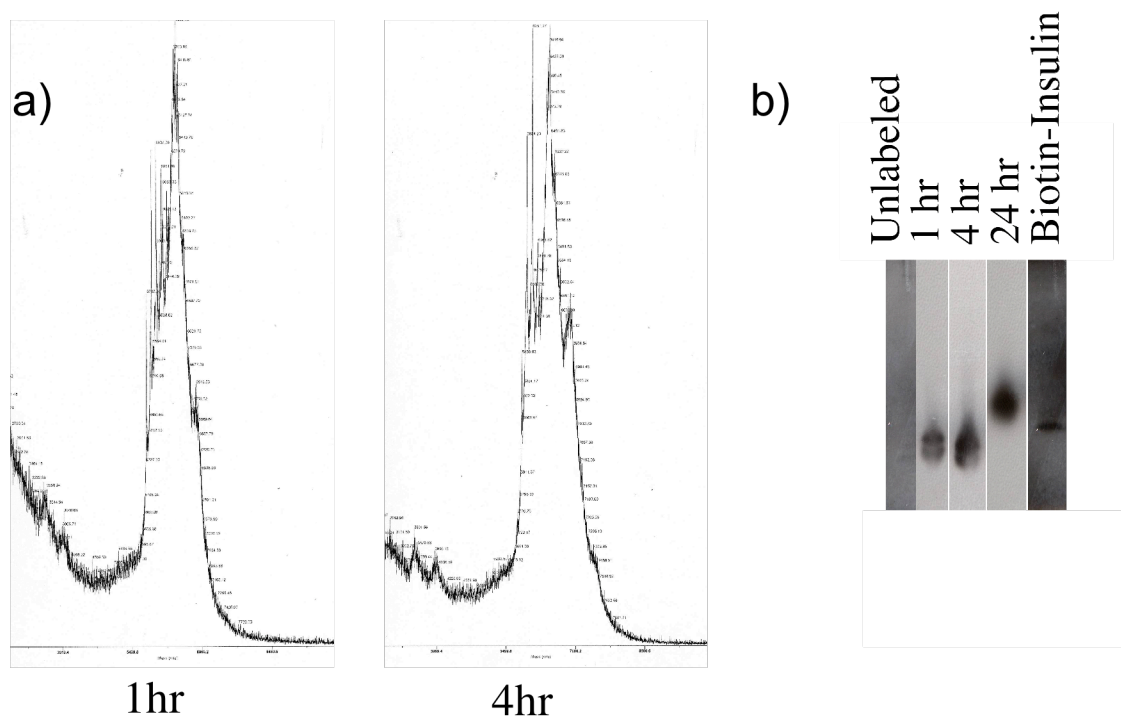


Figure 4.5 EDC activation of Compound 1. Time course followed by Insulin labeling with 35 molar equivalents of estimated cross-linker in DMF. a) MALDI mass spectra show significant mass shifts indicating multiple cross-linkers per Insulin molecule. b) Westerns with streptavidin-HRP showed significant amounts of cross-linker labeling activity after 1hr of EDC activation. Biotinylated insulin control was prepared with NHS-PEO4-Biotin.

Labeling primary amines: lysines and N-terminus

It is necessary to perform the labeling of protein in aqueous buffers for the subsequent structural analysis to be relevant. A pH range of 7-9 is recommended for the reaction (Hermanson 1996). Our labeling buffer (10mM NaHCO₃ / 150mM NaCl / pH9) that showed efficient labeling with NHS-PEO4-Biotin in previous experiments was used

with the cross-linker and cross-linker to protein ratio was varied for optimal labeling. Since the EDC activated cross-linker is in organic solvent, the effect of increasing numbers of cross-linker was not obvious. Labeling efficiency was significant over a range of cross-linker to protein ratio in our labeling buffer, even though such concentrated reactions had significant amounts of organic solvent (Figure 4.6). More modest amounts of insulin labeling at 0.5mg/ml protein concentration was also achieved with 68 molar equivalents of cross-linker, limiting the fraction of organic solvent in the reaction (data not shown).

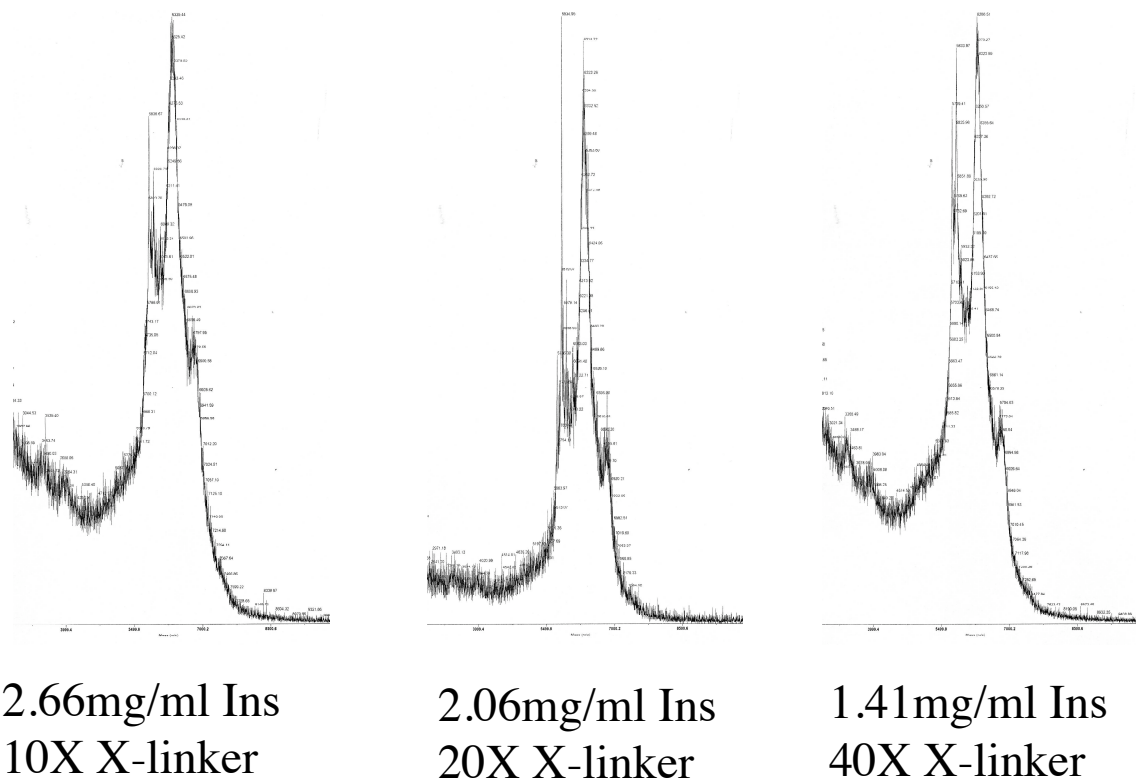


Figure 4.6 Cross-linker to protein ratio. Varying molar equivalents of cross-linker was mixed with insulin and the labeling efficiency was followed with MALDI spectra.

Eliminating excess cross-linker

To control the extent of labeling is extremely important. To stop this process, the labeled protein was purified from free cross-linkers using a SMART System Mono Q column in a buffer system of ammonia/formic acid at pH 8.2. A 1M salt elution concentrated the loaded protein ~7-fold, which was beneficial for the buffer exchange step.

Concentration and buffer exchange

It is absolutely necessary to perform the photoactivation in conditions where the protein is most soluble and monomeric. For insulin a low ionic strength at pH 10.5 offers optimizes these requirements (Kadima, Ogendal et al. 1993). An S75 SMART System column was used to buffer exchange into the photoactivation buffer 5mM CAPS, pH 10.5.

Dilution into photoactivation buffer

To reduce intermolecular encounters in solution, it is crucial that the protein concentration is low enough to minimize diffusion controlled collisions within the time frame of photoactivation. Intramolecular cross-linking of Cytochrome c and ribonuclease

A was previously achieved at 5 μ M protein concentration in a time frame of 1-2hrs with disuccinimidyladipate cross-linkers (Pearson, Pannell et al. 2002). Insulin was diluted to 5 μ M protein concentration subsequent to exchange into the photoactivation buffer.

Photoactivation

The efficiency of photoactivation was compared using a hand-held UV lamp at long UV, an HP spec and a 1 kW laser for 30 min, 10 min and 100x 1 sec pulses, respectively. The cross-linkers Sulfo-SANPAH, Sulfo-SFAD and Sulfo-SBED (Pierce) with photoactive nitro-phenylazide, tetrafluoroarylazide, and phenylazide groups, respectively, were used as described in their manuals without diluting protein to monomeric conditions and the intermolecular cross-linking was used as a probe of the photoactivation efficiency (Figure 4.7). The latter two reagents have reducible disulfide groups in their linker. As evidenced by a shift of molecular weight, the photoactivation by laser performed the best. Also, the photoactive nitro-phenylazide and the tetrafluoro-arylazide groups in Sulfo-SANPAH and Sulfo-SFAD were found to be more efficient than the arylazide in Sulfo-SBED, supporting the choice for a substituted ring in the cross-linker design.

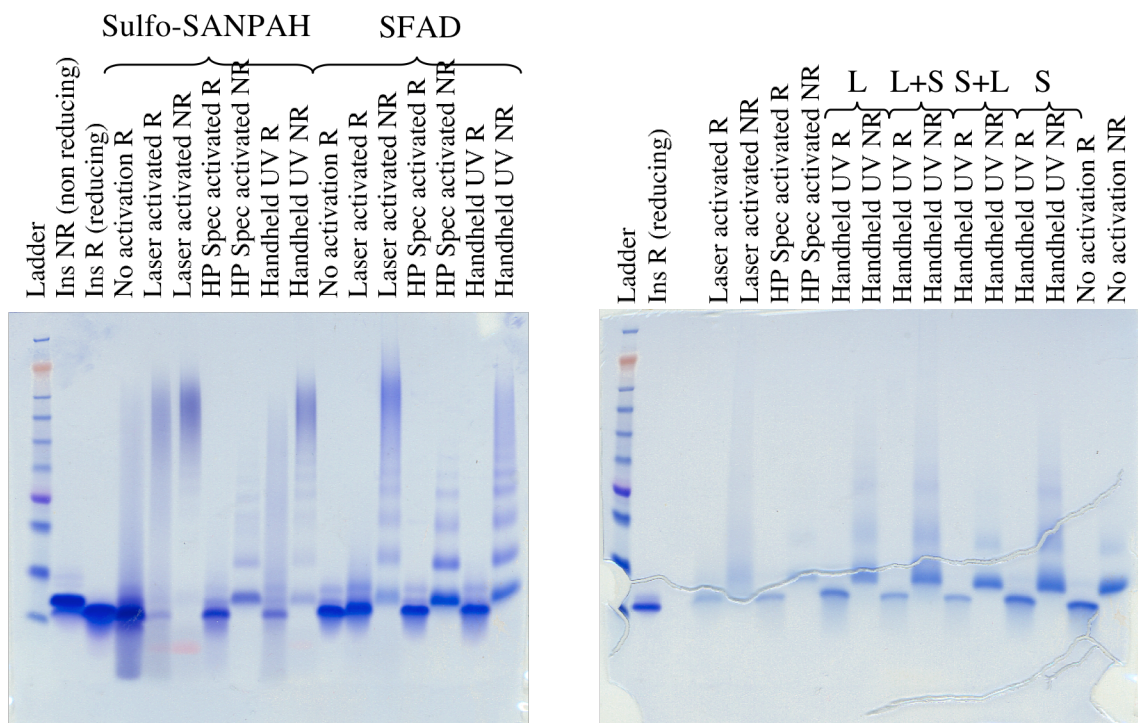


Figure 4.7 Comparison of photoactivation methods. Intermolecular cross-linking as an indicator of photoactivation with a handheld UV lamp (long (L) and short (S) UV), an HP spectrometer deuterium lamp and a 1kW laser with a 300-400 nm pass filter. Samples were run in SDS-PAGE gels with reducing (R) and non-reducing (NR) sample buffers.

Purification of monomers

To eliminate any insulin polymers that would result in false positive peptides for intramolecular cross-linking, the photoactivated sample was first concentrated with a MonoQ SMART System column using the ammonia/formic acid system at pH 8.2. To eliminate any dimers the protein was purified with an S75 column in 20% acetic acid, pH2, where it was characterized to be monomeric up to 20mg/ml (Zhang, Whittingham et al. 2002). A load of 0.4 mg/ml insulin eluted as a monomer and up to 10mg/ml insulin

was separately observed to elute as a monomer in a 20% acetic acid (pH 1.75) on a Superdex 200 column in an AKTA system. To prepare for the digest step, it was once again buffer exchanged to 0.1 M Tris, pH 7.8 using the S75 column.

Proteolytic Digest

The average molecular ion size increases drastically as the peptides get cross-linked. The added molecular weight of a cross-linker and another peptide can more than triple the molecular weight or more, if multiple peptides are linked. Therefore, it is crucial to perform the protein digest with a frequent enough cutter to result in a low average molecular weight for peptides. The conventional choice for protein digests for mass spectrometry analysis is Trypsin, since it produces peptides with positively charged Lys or Arg residues that have good ionization characteristics. However, Insulin has few Lys and Arg residues and Lys residues that are cross-linker labeled would be harder to digest.

In contrast, the specificity of V8 would result in 6 peptides of insulin; 3 of each chain: A1-4, A5-17, A18-21, B1-13, B14-21, B22-30 (Figure 4.8). MALDI analysis after the V8 digest showed the larger five of these peptides and no larger ones suggesting the digest was complete.

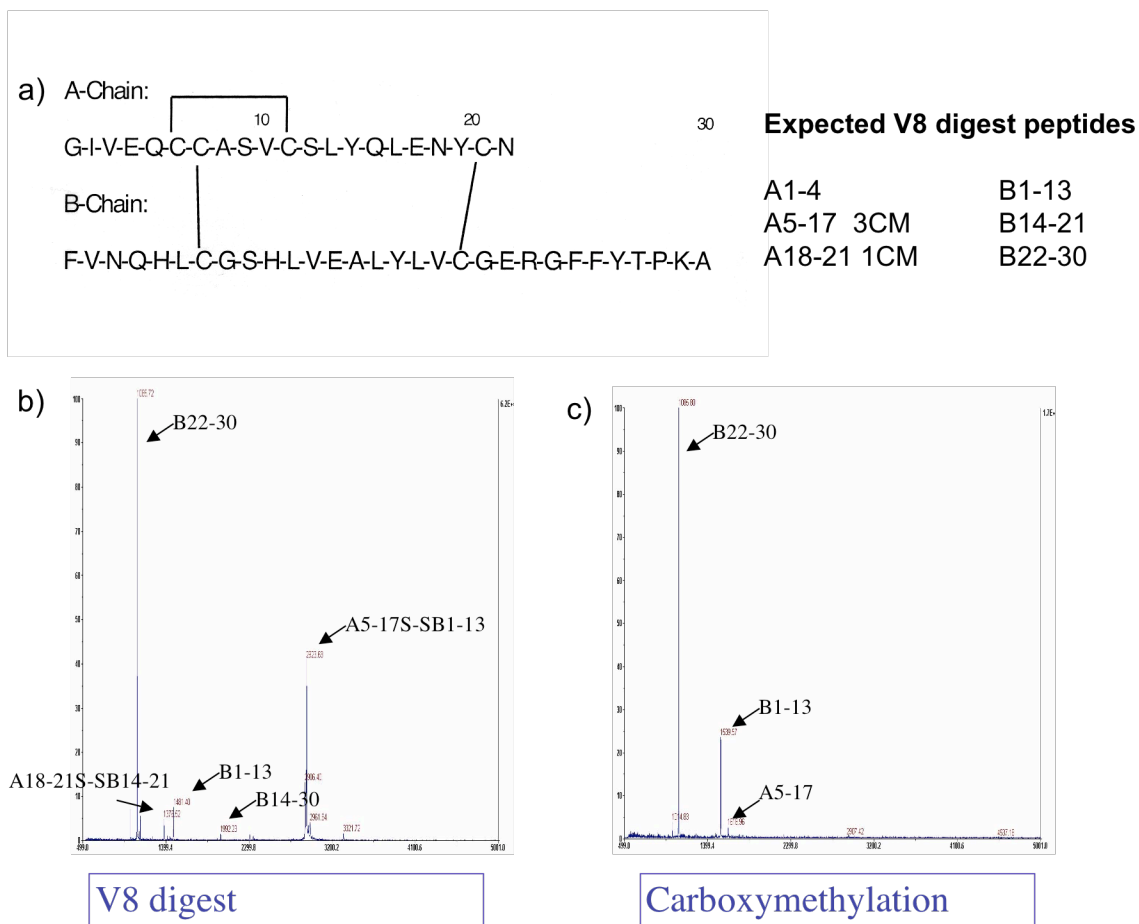


Figure 4.8 V8 digest of insulin. V8 is able to generate reasonably sized peptides from insulin. a) The list of expected peptides with V8 specificity. b) Insulin disulfide bridges keep peptides linked after digest in MALDI spectra. c) Disulfide linked peptides are separated upon reduction and carboxymethylation.

Purification cross-linker labeled peptides

In order to enhance the digest sample with cross-linked or labeled peptides, a desthiobiotin moiety was incorporated into the cross-linker. In order to optimize streptavidin binding and elution unmodified insulin was labeled with desthiobiotin-X-

succinimidyl ester (desthiobiotin-X-SE). The amount of DSB-X labeled insulin in the supernatant was followed with streptavidin-HRP Western gels. The labeled protein bound to the streptavidin plate both at RT after 2 hours. Even with 1mM biotin the acetonitrile/formic acid eluent resulted in much better elution than TBS. This eluent is clearly too harsh for native protein, which indicates that any similar avidin purification cannot be performed prior to photocross-linking to determine native state structural information for difficult folders.

Reducing disulfides and carboxymethylate

Disulfide bonded peptides have large molecular weights. Moreover, the disulfides are known to shuffle during ionization for mass spectrometry analysis. These factors complicate the analysis of these peptides. In order to circumvent this, the disulfides are reduced and capped with iodoacetic acid (IAA) to permanently eliminate disulfide bonds. The carboxymethylated peptide masses were checked using MALDI (Figure 4.8 (b)). The mass shifts for peptides B1-13 and A5-17 corresponded to the carboxymethylation of one and three cysteines, respectively.

LC-MS with precursor ion scanning

In a Q-TRAP mass spectrometer (Applied Biosystems) a precursor ion scan provides information about parent ions that fragment into a specific size upon collision-induced dissociation (CID) (Figure 4.9).

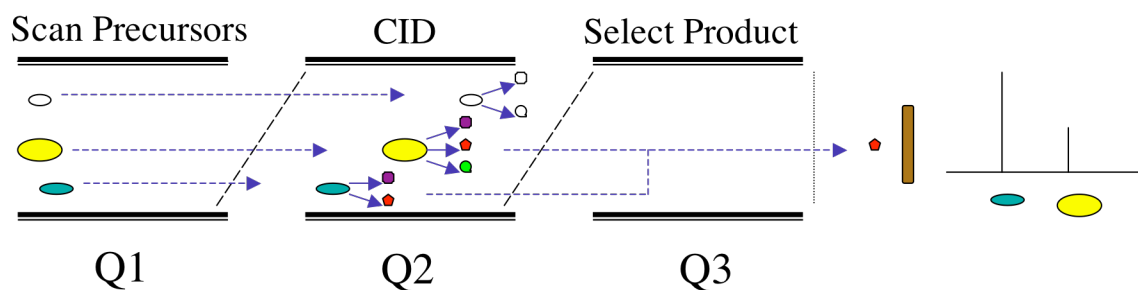


Figure 4.9 Precursor ion scan. In a QTRAP mass spectrometer precursor ion scan Q1 sweeps a given mass range and Q3 is fixed on mass of diagnostic fragments. Only the ions that pass through Q1 and produce a diagnostic fragment will produce signal.

Since the desthiobiotin arm of our cross-linker does not go through any chemical reactions it has the same chemical structure regardless of what happens to the other two arms. Therefore, a frequent CID fragmentation at a specific spot would result in a signature fragment with reproducible molecular weight. Indeed, our cross-linker readily fragments at two locations indicated by the pink and teal arrows (Figure 4.10), giving rise to fragments of sizes 197 and 240 Da.

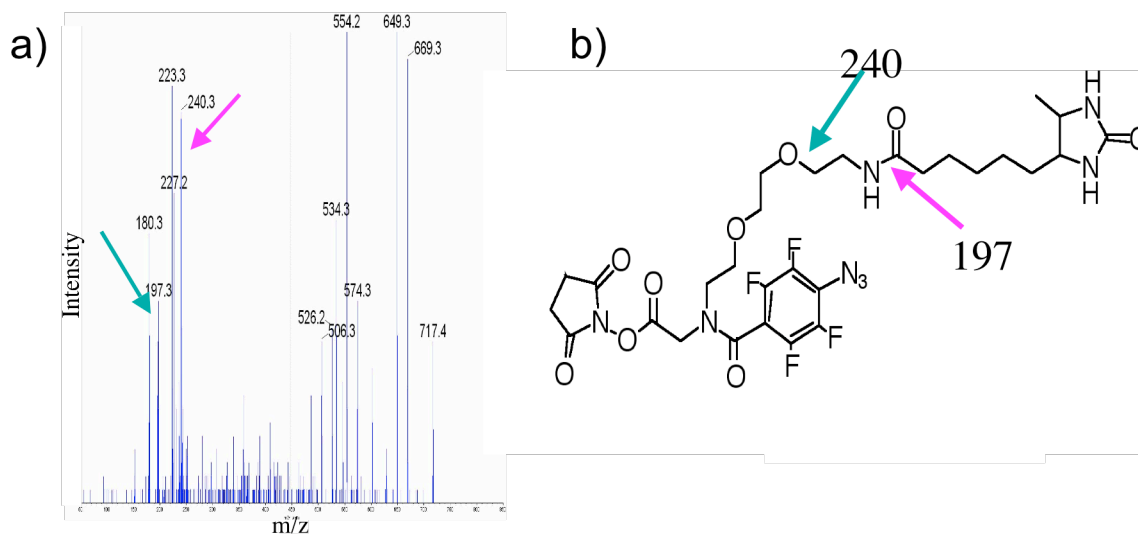


Figure 4.10 Cross-linker fragmentation pattern. a) Fragmentation of the cross-linker gives ions corresponding to fragments from the desthiobiotin arm. b) The fragmentation susceptible bonds on the cross-linker.

The fragmentation pattern was in line with previously observed patterns of lower energy fragmentation sites (Borisov, Goshe et al. 2002). Precursor ion scanning with the fragment ions of $m/z = 240$ and 197 returns the hydrolyzed cross-linker in the spectrum (Figure 4.11).

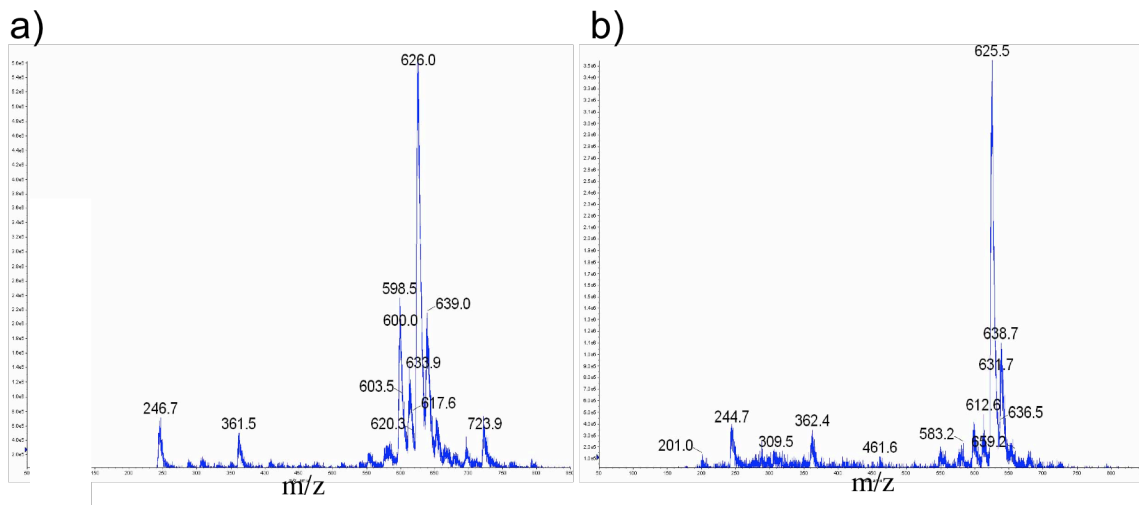


Figure 4.11 Precursor ion scans of the cross-linker. a) Precursor ion scanning with the fragment ion $m/z=240$ returns the hydrolyzed cross-linker in the spectrum. b) Precursor ion scanning with the fragment ion $m/z=197$ returns the hydrolyzed cross-linker in the spectrum.

Fragmentation at the corresponding sites of the NHS-PEO4-Biotin cross-linker is calculated to result in fragments of $m/z=227$ and 270 . To test the efficiency of precursor ion scanning via signature fragments, NHS-PEO4-Biotin labeled Insulin V8 digest was run through an LC-MS-MS protocol on a Q-TRAP with a capillary C18 column (Figure 4.12). Four of the six insulin peptides were pulled out using the $m/z = 270$ fragment in the precursor ion scan (Table 4.1).

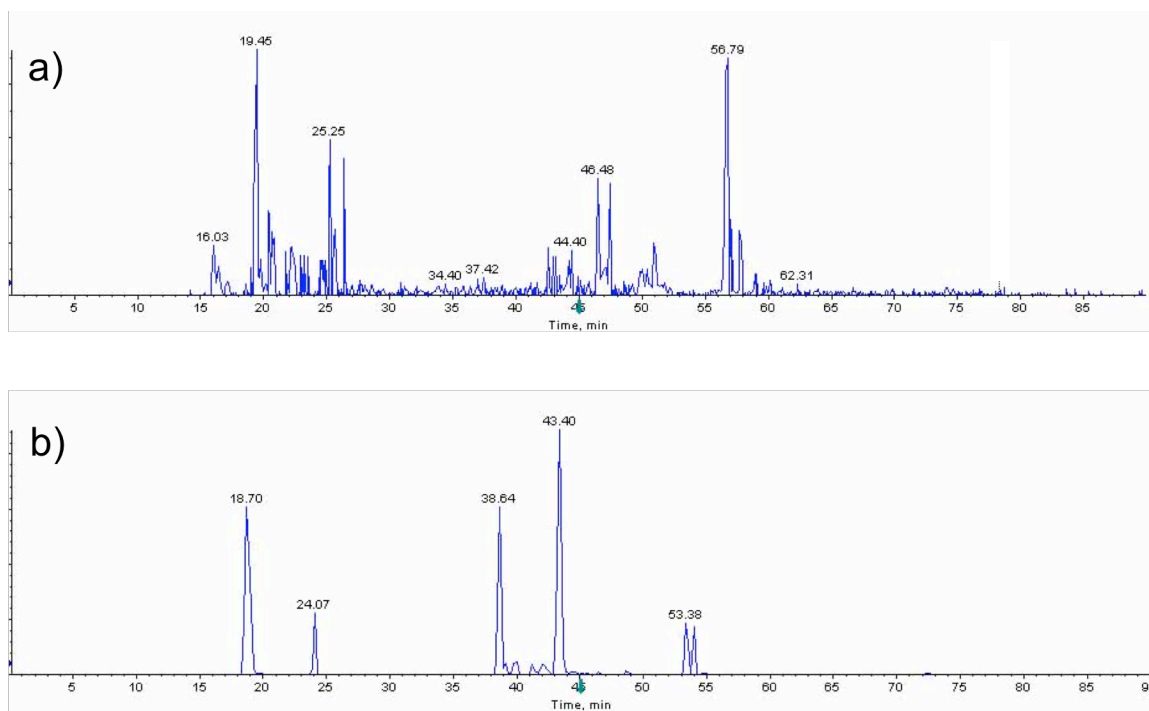


Figure 4.12 Precursor ion scan of biotinylated insulin. Biotinylated insulin was prepared with NHS-PEO4-Biotin. a) Total ion count chromatogram using **E**nhan**C** Multiple Charge (EMC) as survey scan. b) Total ion count chromatogram using Precursor Ion Scan (270; 50) as survey scan.

Table 4.1 Recovery of insulin peptides in mass spectra. CM: Number of carboxymethyls after reducing and alkylating cysteins.

Peptide	Sequence	CM	MW [Da]	MALDI	LC-ESI-MS	Precursor Ion
A1-4	GIVE	0	417	X	(√)	√
A5-17	QCCASVCSLYQLE	3	1620	√	√	√ (once)
A18-21	NYCN	1	571	√ (w/ other peptides)	X (low score √ w/ bigger peptides)	X
B1-13	FVNQHLCGSHLVE	1	1540	√	√	√
B14-21	ALYLVCGE	1	925	√ (w/ other peptides)	X (low score √ w/ other peptides)	X
B22-30	RGFFYTPKA	0	1086	√	√	√

Conclusions

The detection of a large group of cross-linked peptides mapping distance constraints in to the protein structure is an intrinsically difficult task. The trifunctional cross-linker developed for this study provided valuable means to address these difficulties: The abundance of the cross-linked peptides can be enhanced by affinity purification using a desthiobiotin affinity tag on the cross-linker. Also, the characteristic fragmentation of this tag in CID allows for the detection of molecular species carrying this tag in the mass spectrometer.

Despite the many approaches to enhance the abundance of cross-linked peptides, challenges related to the low abundance of each molecular species due to unspecific cross-linking chemistry renders the detection of each cross-linked peptide beyond the limits of current instrumentation. A more specific cross-linking chemistry might be unavoidable in order to decrease the number of cross-linked peptide species and increase the abundance of each. In the future, improvements in the signal to noise levels and detection limits of mass spectrometers may allow for the use of this method as it is for structural assessment of native and nonnative proteins. Unfortunately, with current instrumentation, the use of this method for structural assessment in α LP Int did not succeed.

Materials and Methods

α LP experiments

α LP was prepped as described previously (Kelch, Eagen et al. 2007). For primary amine labeling 230 μ M native α LP was incubated with 25 mM NHS-tetrafluorobenzoate or SFAD (Pierce), 0.5 M Gdn in 10 mM Carbonate at pH 8.7 for 30 min on ice, in dark, vortexing frequently. The labeling reaction was stopped by adding 1 M glycine at pH 3. The reaction was dialysed against 10 mM potassium acetate, pH 5 for 2 hours in the cold room, in dark, to get rid of excess cross-linker. The reaction was diluted down to 5 μ M α LP concentration to prevent intermolecular cross-linking. The cross-linker was photoactivated with a handheld UV lamp in the long UV setting for 30 min. The protein was denatured in 6 M glycine at pH 2 at room temperature overnight. To break the disulfides the sample was reduced with 1000X DTT at 50 °C for 4 hrs and carboxymethylated with 2200X iodoacetic acid in 1 M NaOH + 3:4 (v/v) 1 M NaOH:1 M iodoacetic acid (IAA) to adjust pH to ~8, dark, for 30 min at room temperature (sufficient amount of 3:4 (v/v) 1 M NaOH:1 M IAA was added to adjust pH to ~8). The carboxymethylation was stopped w/ 4400X DTT. The sample was concentrated to ~55 μ M α LP using a centrifugal device with a 5 kDa MWCO. The sample was digested with pepsin or modified trypsin (Sigma) at 37 °C, overnight. These conditions for α LP cross-linking were optimized using Sulfo-NHS-LC-Biotin (Pierce) prior to the use of cross-linker. The cross-linked digests were run on a QSTAR using a 60min LC-ESI-TOF-TOF scan. Zinc-free insulin was prepared as previously described (Yu and Caspar 1998). All subsequent experiments were done with Zinc-free insulin.

Unless otherwise noted protein labeling with NHS-esters at primary amines was done in a labeling buffer of 10 mM Na Carbonate buffer, pH 9/150 mM NaCl at room temperature and stopped with 100 mM Tris, pH 8.5.

Cross-linker synthesis

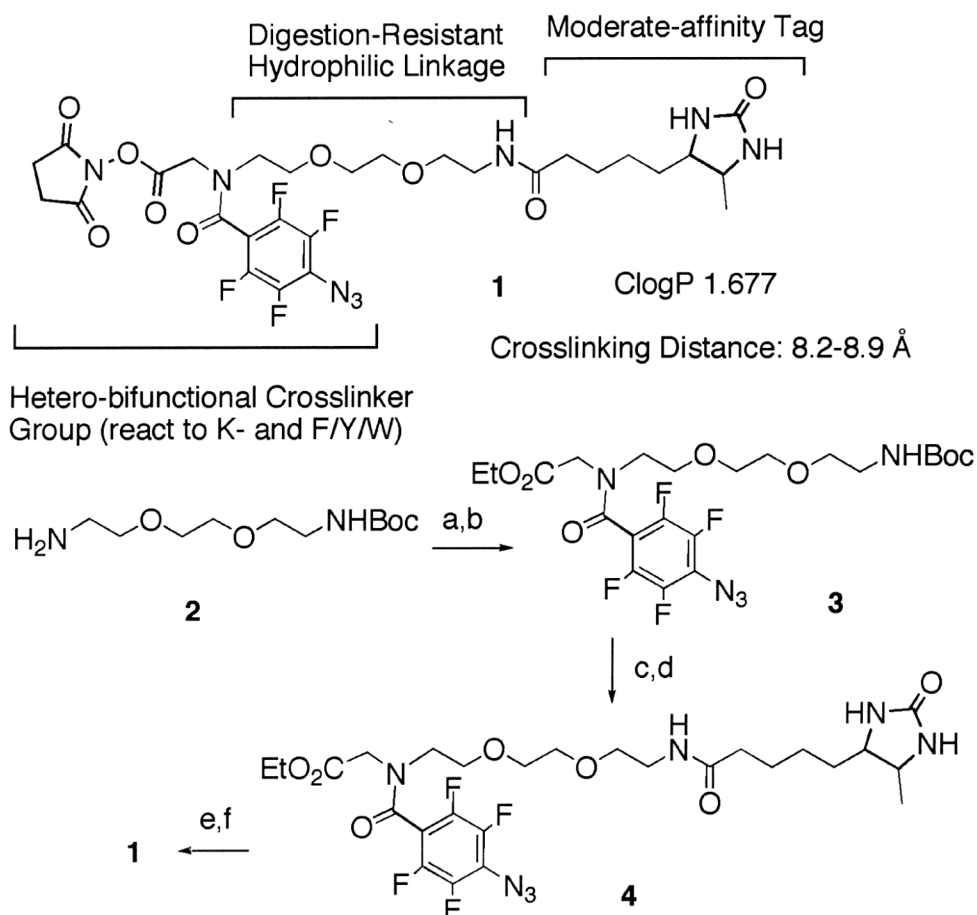


Figure 4.13 Synthesis of cross-linking reagent 1. a) Ethyl bromoacetate (1.0 eq), K_2CO_3 (2.0 eq), rt 1.5 h. b) 4-azido-2,3,5,6-tetrafluorobenzoic acid (**5**, 0.30 eq.), 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) (0.33 eq.), diisopropylethylamine (DIPEA) (0.9 eq.), N,N-dimethylformamide (DMF), rt, 1.0 h, 50% from **5**. c) 4N HCl in 1,4-dioxane, rt, 5 h. d) desthiobiotin (1.5 eq.), HBTU (1.5 eq.), DIPEA (6.0 eq.), DMF, rt, 1 h, 84% from **3**. e) MeOH, NaOH aq. (3.0eq.), rt, 15 h, HCl work-up. f) HOSu (1.3 eq.), PL-EDC (2.3 eq.), 4Å molecular sieves, DMF, $CHCl_3$, rt, 15 h, 78% from **4**.

Compound 3. A mixture of ethyl bromoacetate (58 μ l), compound **2** (130 mg, Senn Chemical), potassium carbonate (145 mg) and DMF (1 ml) was stirred at room temperature for 1.5 h. The reaction mixture was worked up in a usual manner (water/ethyl acetate extraction) to give a crude mixture of mono-alkylated and di-alkylated product. A mixture of the crude product, DIPEA (83 μ l), HBTU (65 mg), **5** (37 mg) and DMF (0.8 ml) was stirred at room temperature for 1.0 h. Only the mono-alkylated product was consumed to afford **3**. The reaction mixture was worked up in a usual manner (water/ethyl acetate extraction). The ethyl acetate extraction solution was washed carefully with dilute hydrochloric acid at pH~4 to remove the unreacted di-alkylated product into the aqueous phase, followed by sodium bicarbonate solution and saturated sodium chloride solution, and evaporated. The crude residue was purified by preparative TLC (hexane/ ethyl acetate = 1/1) to afford compound **3** (43.7 mg, 50% from **2**).

Compound 4. Compound **3** (22mg) was treated with 4N hydrogen chloride in 1,4-dioxane (1 ml) at room temperature for 5 h, evaporated, co-evaporated with ethyl alcohol (2 times) and hexanes (1 time) to afford the de-Boc product. A mixture of the crude product, desthiobiotin (13 mg), HBTU (23 mg), DIPEA (42 μ l) and DMF (0.2 ml) was stirred at room temperature for 1.0 h. The reaction mixture was worked up in a usual manner (water/ethyl acetate extraction) and purified by preparative TLC (dichloromethane/ methanol = 10/1) to afford compound **4** (not weighed).

Carboxylic acid form of cross-linker 1. All of the compound **4** obtained above was dissolved in methanol (1.2 ml) and treated with aqueous 1 N sodium hydroxide solution (120 μ l) at room temperature overnight, under dark. 4 N hydrogen chloride in 1,4-dioxane (33 μ l) was added to the mixture and evaporated. The residue was purified by preparative TLC (dichloromethane/ methanol/ acetic acid = 8/1/0.2) and eluted with methanol. The purified batch includes silica-gel from the TLC, that was carefully removed by dissolving the batch with saturated aqueous sodium chloride solution, extracting with chloroform several times, and evaporation. The silica-gel free batch was triturated in hexanes and dried under reduced pressure to afford carboxylic acid form of cross-linker **1** (21 mg).

Cross-linker 1. A mixture of carboxylic acid form of cross-linker **1** (21 mg), HOSU (5.1 mg, 1.3 eq.) polymer-supported EDC (1.45 μ mol/g, 54 mg, 2.3 eq.) molecular sieves 4 Å powder (pre-dried with microwave for 3 min, 18mg), DMF (0.27 ml), and chloroform (0.54 ml) was stirred in a 10 ml glass flask (pretreated with chlorotrimethylsilane) at the ambient temperature overnight. The reaction mixture was diluted with chloroform (5 ml), filtered, and evaporated. The residual DMF was removed by repeated evaporation with isopropyl alcohol (40 ml, 2 times) then with chloroform (40 ml, 1 time) and pumping up to afford 19 mg of **1**, which was dissolved with DMSO for running of protein cross-linking. The product was confirmed via NMR and MALDI spectra.

EDC activation of carboxylic acid form of cross-linker

EDC activation with EDC from NOVABIOCHEM (N-cyclohexylcarbodiimide, N^o-methyl polystyrene ML (200-400 mesh) 2%DVB Loading: 1.3-1.7 mmole/g. To 5 mg of Compound1, 1.21 mg NHS, 4.3 mg dried MS4A (Molecular Sieves), 500 μ l DMF and 17.9 mg EDC was added to maintain 2.5 molar equivalents of EDC and 1.3 molar equivalents of NHS. Also, the reaction was scaled down 5 fold and 35.8 mg EDC was added for 25 molar equivalents. The reaction was stirred in dark for time periods of 1, 4 and 24 hr in glass vials and centrifuged in the end of the incubation period at 1000 rpm. The supernatant was used to label protein.

Insulin labeling

Insulin labeling was done with Zn-free Insulin in three different conditions. At a 35 fold cross-linker to protein ratio, 5.25 μ l of Insulin (3.8 mg/ml in DMF) was mixed with 8 μ l of the cross-linker (EDC activated Compound1 supernatant) and 0.67 μ l TEA/DMF (9.4 μ g/1.2 μ l). At a 21 fold cross-linker to protein ratio, 10.5 μ l of Insulin (0.57 mg/ml in DMF) was mixed with 1.6 μ l of the cross-linker and 0.58 μ l TEA/DMF. At a 68 fold cross-linker to protein ratio, 10.5 μ l of Insulin (0.7 mg/ml in 10 mM NaHCO₃/150 mM NaCl) was mixed with 4.5 μ l of the cross-linker. The reactions were carried out in dark for 30 min. To stop the reactions glacial acid was added to 1 M in DMF reactions and Tris pH 8.5 was added to 100 mM for the aqueous reaction. 5 μ l of each reaction was prepped using ZipTip C18 and spotted on a MALDI plate with FWI matrix buffer and saturated α -cyano cinnamic acid (50% 2-propanol/0.2% Formic acid) (Dai, Whittal et al. 1999).

Photoactivation

For the laser activation a UV filter (300-400 nm, Oriel Corporation 51660) was combined with an IR blocking filter to control heat at the 1 kW setting. The solution was kept in an uncovered a Petri dish the size of the laser light. The hand-held lamp was positioned on top of the sample in a low affinity 96-well plate at 4C. The white light of a deuterium lamp in the HP Spec was directed at the sample in a quartz cuvette at room temperature keeping the shutter open.

V8 digest was performed with 0.1mg/ml Insulin and 10-20 µg/ml V8 in 0.1 M Tris pH 7.8 at 37°C for 24 hrs.

Streptavidin plate purification

The desthiobiotinilation of insulin was optimized with 1mg/ml insulin in 10 mM NaHCO₃/150 mM NaCl, pH 9. An excess of 18X, 9X, 4.4X, 2.2X and 1.1X desthiobiotin-X-succinimidyl ester (DSB-X) was used and 4.4X and 2.2X labeling conditions gave the best results for labeling with about 3 DSB-X per insulin molecule. To eliminate the free DSB-X and facilitate plate binding, the labeled protein was passed through a SMART System MonoQ column in 20 mM Tris pH 8.2 and to mimic the conditions of a photoactivated sample, it was buffer exchanged into 5 mM CAPS pH 10.5 using an S75 size exclusion column. 10X PBS was added 1:10 bringing the pH down to 7.65. The labeled protein was incubated in PBS or TBS (pH7.6) on a SigmaScreen™

High Capacity Streptavidin Coated Plate (Sigma) for 2 hr at room temperature or for 24 hr in the cold room. Around 40 pmol of protein was loaded to each well with a capacity of 500 pmol. The wells were washed with PBS or TBS. The following eluents were tried: 100 μ M biotin/TBS, 100 μ M biotin/70% acetonitrile/5% formic acid, 1mM biotin/TBS, and 1 mM biotin/ 70% acetonitrile/5% formic acid.

Alkylation of reduced cysteins

In order to permanently break the disulfide bonds, the digest was reduced with 1000X DTT at 50 °C for 4hrs and carboxymethylated with 2200X iodoacetic acid in 1 M NaOH + 3:4 (v/v) 1M NaOH:1M iodoacetic acid (IAA) to adjust pH to ~8, dark, for 30 min at room temperature (sufficient amount of 3:4 (v/v) 1M NaOH:1M IAA was added to adjust pH to ~8). The carboxymethylation was stopped w/ 4400X DTT.

Chapter 5. The role of disulfide bridges in α LP folding and unfolding

Introduction

Disulfide bridges in α LP were originally studied in the context of the α LP precursor: A single chain of Pro and α LP covalently linked between the C-terminus of Pro and the N-terminus of α LP. In *E. coli* the precursor is transported across the inner membrane to the periplasmic space, where the signal (pre) sequence is removed. Upon proper folding, the precursor is cleaved between the pro region and the protease region. α LP activity is required for cleavage. The two regions have a high affinity for one another and therefore probably remain as a complex while they are conveyed across the outer membrane. If the protein does not fold correctly or cleavage between the pro region and the protease region does not occur, the misfolded/precursor form becomes tightly associated with the outer membrane (Fujishige Boggs and Agard 1996).

In the periplasm, the formation of the disulfides becomes the rate limiting factor for precursor folding and autocatalytic cleavage of the Pro region and subsequent structural rearrangements occur very rapidly in comparison. Partially oxidized forms of the precursor are not competent to yield mature protease (Anderson, Peters et al. 1999). Thus, disulfide formation is necessary for Pro-catalyzed folding.

There are three disulfide bridges in α LP. The N-terminal disulfide bridge 1 (Cys42-Cys58) has a 20-residue sequence separation. A β -hairpin between these residues completes the β -barrel of the N-terminal domain, although the disulfide bridge itself is not located on any of the β -strands. Disulfide bridge 1 is close to Leu41, which extends

towards the active site and has extremely large B-factors in crystal structures of α LP (see Chapter 3).

The C-terminal disulfide bridge 2 (Cys137-Cys159) has a mere 10-residue sequence separation. It anchors two β -strands together.

Disulfide bridge 3 (Cys189-Cys 220A) has the largest sequence separation of 33 residues. There is a β -hairpin, as well as coiled regions before and after the two β -strands. It is located close to the substrate binding pocket. In addition, it is in direct contact with Trp 199 which is involved in introducing the non-planarity to the Phe228 ring (Fuhrmann, Kelch et al. 2004). The two C-terminal disulfide bridges, disulfide bridge 2 and disulfide bridge 3, are both located at the interface where α LP makes contacts with Pro in the α LP•Pro crystal structure (Sauter, Mau et al. 1998).

While the correct disulfide formation is required for Pro-catalyzed folding another important contribution of disulfides to the folding landscape of α LP could be to the unfolding cooperativity. Native α LP unfolds in a perfectly cooperative manner by suppressing local breathing motions as evidenced by hydrogen exchange protection factors of greater than 10^9 throughout a large hydrophobic core that spans both N- and C-terminal domains (Jaswal, Sohl et al. 2002). By closing loops around β -hairpins disulfide bridges could stabilize the structures of individual β -barrel domains of α LP, leaving only few ways for α LP's unfolding and thus making the unfolding transition cooperative.

Here, we study the contributions of the disulfide bridges to the unfolding and uncatalyzed folding of α LP. These preliminary experiments attempt to alter the disulfide contributions by dialing them out in redox buffers of varying red:ox ratios.

Results

α LP unfolding by loss of activity with GSH

To check the effect of reducing conditions on α LP unfolding, a 6 hr time course followed the unfolding of native α LP by loss of activity in a 10mM GSH/GSSG system. To facilitate the unfolding reaction in a realistic time frame, all experiments were carried out with a baseline of 5 M Gdn HCl. The results were fitted with a single exponential. Using with double exponentials did not improve fits.

No significant change in unfolding rate was observed even in the most reducing conditions with the use of 100% GSH (Figure 5.1). A $k_{unf} = 6.98 \cdot 10^{-5} \text{ s}^{-1}$ in fully reducing conditions was compared to a $k_{unf} = 5.58 \cdot 10^{-5} \text{ s}^{-1}$ in the absence of glutathione. While the difference is lightly larger than the error ($3.6 \cdot 10^{-6} \text{ s}^{-1}$) in the reducing unfolding, it only accounts for a mere $\Delta\Delta G = RT \ln(k_{unf,red} / k_{unf,o}) = 0.13 \text{ kcal/mol}$.

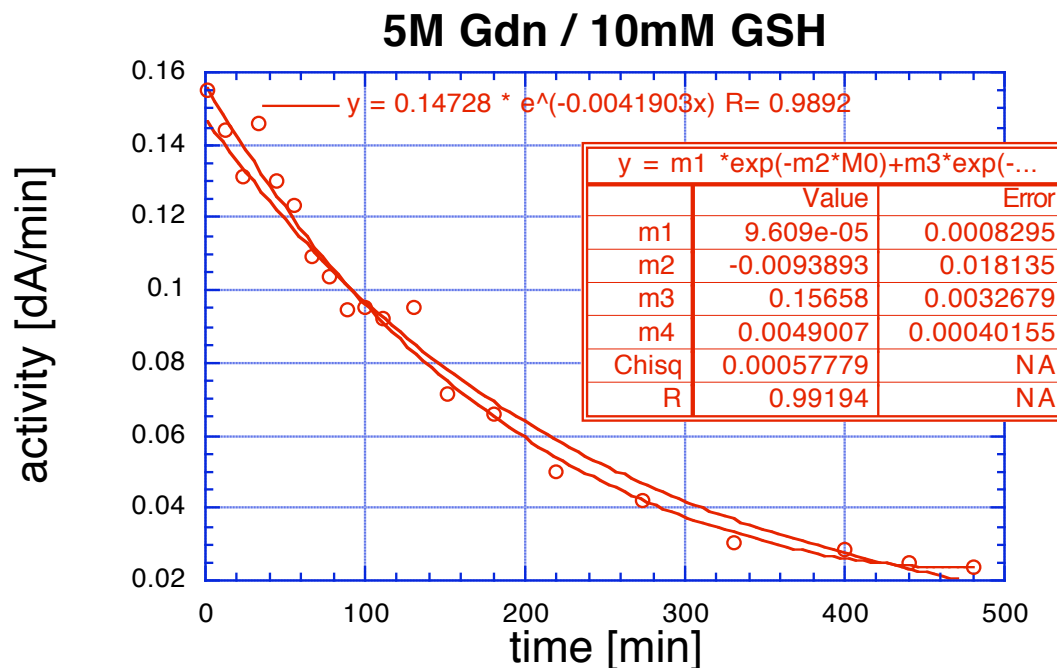


Figure 5.1 α LP unfolding in reducing conditions followed by loss of activity. The unfolding rates with or without reducing agent are indistinguishable.

α LP unfolding by Trp fluorescence

α LP unfolding under reducing conditions was also tested following Trp fluorescence in 5 M GdnHCl to facilitate unfolding in a reasonable time frame. The unfolding reaction proceeded in an apparently single-phase manner as the fits to unfolding traces did not improve with double exponential functions. The exact GdnHCl concentration was checked for each reaction and accounted for most of the variation on reducing conditions, as well as reactions without glutathione (Figure 5.2).

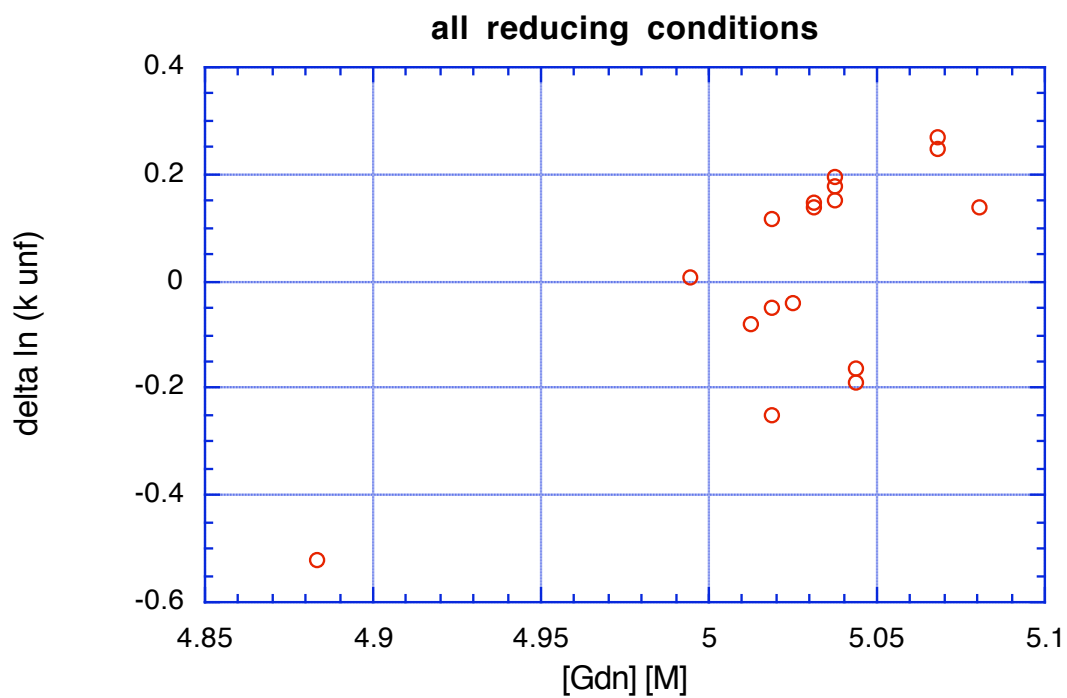


Figure 5.2 α LP unfolding under reducing conditions followed by Trp fluorescence. No significant differences were found between reducing and oxidizing buffers. Instead, the experimental variation in GdnHCl concentration was a stronger determinant of the small differences in unfolding rate.

Uncatalyzed α LP refolding with GSH

In order to study the contribution of disulfides to the uncatalyzed folding reaction, denatured α LP was diluted into Int conditions in the presence of a GSH/GSSG redox system. Even at the least reducing ratio of 1% GSH no protease activity was detected in the end of the incubation period. An aggregate formed immediately after mixing with the redox agent, likely preventing any productive folding. The aggregates were reversed to some degree in 20% SDS gels with beta-mercaptoethanol, but not without (data not shown).

Discussion

It is remarkable that α LP unfolding kinetics is not affected by the reducing conditions tested. All three disulfides in α LP are solvent accessible, although Cys42-Cys58 is slightly protected by the His57 ring and the Ser195 side chain, two of the three residues in the serine protease catalytic triad (Fuhrmann, Kelch et al. 2004).

The other two disulfides, Cys137-Cys159 and Cys189-Cys220A have large exposures to solvent. Thus, it is unlikely that the lack of the glutathione effect on the unfolding is due to chemical inaccessibility.

There are two other explanations for the insensitivity of unfolding rates to reducing conditions: The cooperative nature of α LP might be such that there are no local fluctuations allowing the cysteine side chains to deviate from a well-packed disulfide bond configuration. This is in line with the extremely rigid native state dynamics previously observed and the cooperative nature of the unfolding transition (Jaswal, Sohl et al. 2002). This explanation can be tested by cysteine deletion mutants permanently eliminating the disulfide link.

Alternatively, it is possible that the structural environment around the disulfides does not change in the unfolding transition. This explanation is not unlikely as the TS and N state were found to have extreme similarity at ~80% (Jaswal 2000). A useful TS model with relatively intact N- and C- terminal domains intercalated by water in the domain-

domain interface (Figure 1.2) is supported by an extremely high entropic and low enthalpic barrier to the unfolding transition (Jaswal, Truhlar et al. 2005) and the effect of inter-domain salt bridges on acid sensitivity of unfolding (Kelch, Eagen et al. 2007) and Chapter 2. None of the disulfides span across the two domains and the insensitivity of unfolding to reducing conditions might be due to the fact that the conditions affect the TS and N state similarly leaving the unfolding barrier unchanged.

Even if the disulfides do not undergo structural changes along the unfolding transition, it is still likely that they contribute to the cooperativity. Salt bridges that didn't have a significant effect on unfolding rate did contribute to unfolding cooperativity greatly (Chapter 2). If true, in the absence of cysteines we would observe decreased cooperativity for the unfolding and relaxed protein dynamics for the N state, which can be tested by survival assays and native state H-exchange experiments (Jaswal, Sohl et al. 2002).

The reducing intracellular environments are not amenable to disulfide bonds and therefore these are commonly found only in secreted proteins, like α LP. Previous studies have shown that disulfide bond formation is a rate limiting step for precursor folding in addition to Pro folding and that subsequent structural rearrangements were very rapid compared to disulfide formation (Anderson, Peters et al. 1999). It is unclear if this is due to a stepwise reduction of conformational entropy through formation of individual disulfides or if the disulfide formation marks and locks a concerted structural transformation.

In SA195 α LP precursor, disulfide bonds formed in the presence of reducing agent suggesting there is a prior structural stabilization leading to the formation of native disulfides (Anderson, Peters et al. 1999). The preliminary experiments performed in this project were an attempt to study the role of disulfides in uncatalyzed folding. Aggregation problems prevented the protein from folding at all. In these uncatalyzed reactions aggregation was immediate upon mixing of the glutathione and faster than any such structural stabilization.

Use of 20% glycerol or sucrose prevented aggregation of SA195 precursor into disulfide cross-linked multimers, while protein disulfide isomerase had no effect. It is extremely likely that the aggregation observed in uncatalyzed folding experiments is also due to a cross-linking of chains of similar nature and can be prevented by glycerol or sucrose. Alternatively, a series of cysteine deletion mutants eliminating disulfides can be used to avoid disulfide cross-linking across chains preventing protein aggregation.

Materials and Methods

Protein Purification

α LP was purified as previously described (Kelch, Eagen et al. 2007).

α LP unfolding by loss of activity with GSH

13 μ M native α LP was incubated in 5M GdnHCl, 10 mM HEPES pH7.2 and a 10 mM glutathione (GSH, L-gamma-glutamyl-L-cysteinylglycine)/glutathione disulfide (GSSG) system. The GSH/GSSG ratio was varied 0-1. The incubations were carried out for 6 hrs, taking aliquots at several time points. To test activity, the sample was diluted 50-fold in 0.1 M Tris pH 8.0. The activity in the aliquots was measured following the conversion of a pNA substrate at 410nm in a Perkin Elmer UV spectrometer. The activity along the time course was fitted with a single exponential to determine k_{unf} . The fit was repeated with a double exponential to check for double-phased unfolding.

The GdnHCl concentrations were checked before and after the experiment with a refractrometer.

α LP unfolding by Trp fluorescence

AEBSF inhibited α LP was prepared using 10mM AEBSF in 10 mM NaHPO₄ (pH 7.2) at room temperature and was used for the fluorimeter unfolding experiments. 1.75 μ M native α LP was unfolded in 5M GdnHCl, 10 mM HEPES pH 7.2 and a 10 mM

glutathione (GSH, L-gamma-glutamyl-L-cysteinylglycine)/glutathione disulfide (GSSG) system. 5 control reactions were performed omitting the GSH/GSSG. Unfolding was followed at 322 nm as previously described (Sohl, Jaswal et al. 1998). The GdnHCl concentrations were checked before and after the experiment with a refractrometer. The unfolding trace was fitted with a single exponential. The fits were repeated with double exponentials to probe for double-phased unfolding.

Uncatalyzed α LP refolding with GSH

Native α LP was denatured in 6M GdnHCl, pH 2.0, overnight at room temperature. Intermediate α LP (Int) was obtained by diluting denatured α LP down to 5 μ M in 10 mM HEPES, pH 7.2. 10 mM GSH/GSSG system was used with 1-25% GSH.

Refolding was stopped at 45 min intervals by the addition of pepsin (70 mg/ml) in 35 mM Glycine, pH 3 for 30 min on ice. The activity was checked with a succinyl-Ala-Ala-Pro-Ala-thiobenzyl ester as described previously (Sohl, Jaswal et al. 1998) at room temperature, following the enzyme activity at 324 nm. The thio-benzyl substrate activity was calibrated prior to use with known concentrations of α LP.

To check self digestion of α LP, aliquots were taken at 45 min intervals and refolded with 16 molar equivalents excess wt Pro region (obtained from Erin Cunningham) on ice, overnight. The excess Pro was digested by the addition of trypsin (6.7 mg/ml, pH 8) at room temperature, for 10-15 min. The enzyme activity was checked with the pNA substrate as described above.

Conclusions

Role of the domain-domain interface in TS

aLP is one of the most extreme examples of kinetically stable proteins. In addition to an extremely high unfolding barrier (Sohl, Jaswal et al. 1998) and extreme cooperativity (Jaswal, Sohl et al. 2002), the entropic contribution to the unfolding barrier lead to a model involving the N- and C- terminal domains staying relatively intact, but separating from each other to expose the domain-domain interface to solvent. Additional studies focusing on the inter-domain salt bridges (Kelch, Eagen et al. 2007) and the domain bridge (Kelch and Agard 2007) provided supporting evidence for this model.

A more detailed study on the role of salt bridges in aLP domain-domain interface further supports the evidence for the importance of the domain-domain interface in aLP unfolding Chapter 2. Of the three salt bridges across the two domains, salt bridge 3 contributions to the height of the unfolding barrier are dominant implying the involvement of structural rearrangements in the vicinity of the salt bridge 3 in the unfolding transition.

In addition to effects on the height of the unfolding barrier, the salt bridges in concert contribute to the cooperativity of aLP unfolding. This finding proposes the domain-domain interface to be important not only for the structural rearrangements that must take place going from the native state to the TS, but also for the concerted fashion aLP has evolved to do so.

The inter-domain salt bridge mutants provide a means to quantify cooperativity in protein unfolding studies. The underlying principal of this analysis is that dampened protein dynamics at a salt bridge site will reduce the apparent pK_a of the acidic residue, since the protein will spend a lot less time in an open conformation allowing for protonation. The coupling effect observed with pair wise combinations of salt bridges allows for an energetic dissection of the cooperativity contributions. α LP combines the cooperative effects from the inter-domain salt bridges extremely efficiently. In contrast, its acidophilic homolog NAPase avoids acid sensitivity by relocating salt bridges from the domain-domain interface that is involved in the unfolding transition at the expense of achieving a cooperativity level as high as that seen in α LP (Kelch, Eagen et al. 2007).

For the energetic dissection of cooperativity, it is, however, not a requirement that the coupled structural elements be salt bridges. Any number of experimental setups that can register open/closed equilibria in protein dynamics in response to changes in conditions such as red:ox potential, ligand concentration etc. can be exploited to carry out such analysis in combinations with mutants that are thought to contribute to cooperativity.

Extremely low native state pK_a 's

Extremely low native state pK_a 's are predicted for the acidic groups in inter-domain salt bridges in α LP. It is possible that some of the other acidic residues have similarly low values due to dampened native state dynamics of α LP (Jaswal, Sohl et al.

2002). Measuring the native state pK_a 's would be informative, but challenging due to much faster unfolding of α LP at lower pH values. However, because the salt bridges are solvent exposed and such low pK_a values are only observed in buried salt bridges with polar interactions, a more accurate measure of the native state pK_a 's of α LP would be instrumental in demonstrating how big of a role the protein dynamics can play in tuning the apparent strength of salt bridges.

Further characterization unfolding barrier using salt bridge mutants

The salt bridge mutants studied in Chapter 2 were instrumental in deciphering how cooperativity is achieved on a structural basis. Moreover, the pK_a 's of the mutants offered a way to quantitatively assess cooperativity to the unfolding transitions of wt α LP relative to the salt bridge mutants. According to this interpretation, the double salt-bridge mutants were less cooperative than wt α LP and the single salt-bridge mutants were even further less. In addition, the crystal structure of the SB1+2 mutant, lacking salt-bridge 3 included a peptide at its active site, most likely as a result of autolysis. While this can be a result of the overall decrease in the unfolding barrier for this mutant, it could also be a result of compromised cooperativity.

Survival assays provided a useful method to assess cooperativity in α LP and its homologs (Jaswal, Sohl et al. 2002), (Kelch, Eagen et al. 2007), (Kelch 2007). The salt-

bridge mutants are prime candidates for a cooperativity analysis by survival assays, as cooperativity levels are quantitatively predicted for these from the pH sensitivity studies. Comparing the survival rates of the mutants against a common protease can easily rank the magnitude of their cooperativity and assess the strength of the quantitative cooperativity predictions that emerged from the pH sensitivity studies.

The pH sensitivity studies with the salt-bridge mutants interpreted the energetic coupling between structurally remote salt-bridges to be of entropic origin adding order to the TS. As the native state pK_a 's for the salt-bridges were found to be well below the accessible pH range in these experiments, the salt-bridge contributions to the native state free energy and entropy is unknown. Eyring analysis of the wt α LP unfolding resulted in a largely entropic component for the unfolding barrier, suggesting ordering in the TS structure and predicting water intercalating between the N- and C-terminal domains that are intact but separated (Jaswal, Truhlar et al. 2005),(Kelch, Eagen et al. 2007), (Kelch and Agard 2007). Eyring analysis of the salt-bridge mutants can determine the entropic and enthalpic components to the unfolding barrier for these mutants. In combination with the quantitative cooperativity predictions from the pH sensitivity experiments and survival assays, the energetic components can decipher in greater detail the factors that establish the extremely high, cooperative unfolding barrier of α LP.

Recapturing stability at the salt bridge 3 site

Out of the three salt-bridges at the domain-domain interface of α LP, salt-bridge 3 seemed to be involved in structural rearrangements in the unfolding transition and contributed to the acid sensitivity of α LP the most. In contrast, NAPase achieves acid resistance without lowering the unfolding barrier to levels observed in the absence of salt bridge 3 from α LP. While NAPase likely does so by stabilizing the protein elsewhere, I was interested to see if the lost stability provided by this salt-bridge could be recouped by picking substitutions different than the corresponding residues in NAPase.

Dan Mandell performed a fixed-backbone repacking calculation with this salt-bridge and determined the R64M/E182T variant more favorable over the NAPase equivalents. The design is interesting, since the threonine satisfies a buried backbone hydrogen bond with D63 in the active site, which may affect the catalytic activity. Buried hydrogen bonds were found to have greatly lowered pK_a values suggesting significant free energy contributions (Pace, Huyghues-Despointes et al. 2002), (Bartik, Redfield et al. 1994), thus these computational predictions may indeed result in a high unfolding barrier that is also acid resistant.

Further structural elements are thought to contribute to cooperativity

Neema Salimi performed molecular dynamics unfolding simulations of α LP and compared it to Trypsin (Salimi, Ho et al.). Analyzing the pairs of contacts throughout the simulations, he was able to cluster contacts according to their cooperativity as measured by their correlated behavior (unpublished data), i.e. if two contacts were formed and broken in synchrony, they were considered to have high cooperativity. In α LP, 340 analyzed contacts resulted in 1238 cooperative pairs above a threshold, while in trypsin only 174 contacts made up the 318 cooperative pairs. While α LP's major networks representing semi-local structural regions were interconnected by "hubs", only one non-local network was found in trypsin and the individual networks were mostly isolated. A more careful look at the conserved networks only accentuated the behavior in each protein.

The hubs provided several sites that could be the backbone of a wide spread cooperativity network in α LP spanning a large fraction of the structure. Some of these sites are interestingly at structural components that were previously thought to be important for the kinetic stability of α LP. Among those, four sites where the contact is maintained by a side chain are amenable for mutational studies: 1) V119:L131 is located at the base of the beta-hairpin that forms a five strand beta-sheet with Pro in the Pro•N complex and its role in catalyzed and uncatalyzed folding was studied by Stephanie Truhlar (Truhlar and Agard 2005). 2) V79:V86 and V79:V88 are located on the domain

bridge, which is thought to be involved in the unfolding transition (Kelch and Agard 2007). 3) Y9:I46 is a contact at the chore of the N-terminal beta-barrel. A double mutant involving both residues might have too drastic consequences on the kinetic and/or thermodynamic stability of the protein and thus, single mutants might be more adequate for study. I have cloned and expressed Y9A, however lost most of the protein when I tried to purify it under the same conditions of the more resilient wt α LP. 4) S18:G6 contact looks especially interesting because of the packing differences seen around this contact in α LP and Trypsin. In addition, it is in close proximity to salt-bridge 1, which provides significant cooperativity to the unfolding transition. Mutating G6 to Trypsin counterpart seems deleterious due to tight packing around it. I have cloned the S18G mutant, however wasn't able to observe much expression at 12 °C. It may behave differently under different expression conditions or the mutation might have altered the folding landscape so significantly, it may not be possible to yield any native protein.

A method for structural characterization of nonnative protein species motivated by α LP Int

Structural information of nonnative states such as folding intermediates is crucial in our understanding of the folding landscapes. The nonnative states are tough to study structurally: They are often short-lived, prone to aggregations, especially at high concentrations that are often required for structural determination, and structurally

diverse rendering each species rare and making it hard to deconvolute their combined signal.

α LP Int is an especially interesting examples, as its study would provide insights into the nature of the extremely high folding barrier of α LP. The importance of this question motivated the extremely challenging task of the development of a method for the characterization of nonnative structures like α LP Int.

The method development discussed in Chapter 4 summarize this attempt to find out more about the structural content in α LP Int. The method involves intramolecular cross-linking of the protein followed by proteolytic digest and the analysis of cross-linked peptides via mass spectrometry. The tri-functional cross-linker designed and synthesized for this project possesses crucial properties for such a method. The cross-linker is able to react efficiently with both the chemical and photoactive functionalities under the conditions optimized for the method. The desthiobiotin affinity tag provides a means for selective purification of the cross-linked peptides. The same tag fragments in a reproducibly characteristic fashion in CID experiments allowing precursor ion scan experiments to be used in order to enhance the signal of the cross-linked peptides over the rest of the peptide species, which are more abundant in any sample.

Challenges related to the low abundance of each molecular species due to unspecific cross-linking chemistry renders the detection of each cross-linked peptide beyond the limits of current instrumentation. For the successful use of this method, a

more specific cross-linking chemistry might be unavoidable in order to decrease the number of cross-linked peptide species and increase the abundance of each. In the future, improvements in the signal to noise levels and detection limits of mass spectrometers may allow for the use of this method as it is for structural assessment of native and nonnative proteins.

Uncatalyzed folding of α LP under reducing conditions

Uncatalyzed folding experiments of α LP in redox buffers faced challenges with protein aggregation. Protein aggregation is immediate after mixing α LP Int with reducing agents. The aggregates appear to be intermolecular cross-links that can be reversed in reducing SDS gels. α LP precursor also forms cross-linked multimers. To avoid the intermolecular cross-links the use of protein disulfide isomerase has no effect, in contrast to glycerol and sucrose (Anderson, Peters et al. 1999). Thus, approaches that reduce the local concentration of protein could help remedy the intermolecular cross-linking and avoid protein aggregates.

Even in the absence of aggregation challenges, however, it may be very difficult to observe uncatalyzed folding with eliminated disulfide bridges. Catalyzed and uncatalyzed folding reactions most likely share a very similar TS (Derman and Agard 2000) and thus proceed through similar transitions. Uncatalyzed folding is extremely slow with $t_{1/2} = 1,800$ years. In the Pro-catalyzed folding, correct disulfide formation was

required and the rate-limiting step for folding (Anderson, Peters et al. 1999). Eliminating the disulfide bridges would thus slow the rate limiting step of an already very slow reaction.

α LP unfolding rates are unaffected by reducing conditions

Remarkably, fully reducing conditions do not have an affect on the unfolding rate of α LP suggesting that i) the native state dynamics in α LP are so suppressed that the cysteins in the disulfide bridges never access conformations where the bond can be reduced or ii) the disulfide bridges are not altered in the TS compared to the N state. Much like with the inter-domain salt bridges, it is still possible that an interaction is contributing to the cooperativity of the unfolding transition, instead of the barrier height. Alanine substitutions of the cysteins in individual salt bridges could help decipher both the disulfide bridge contribution to the barrier height by following unfolding kinetics, and their contribution to unfolding cooperativity via survival assays or native state hydrogen exchange studies (Jaswal, Sohl et al. 2002).

References

- Adams, P. D., P. V. Afonine, et al. "PHENIX: a comprehensive Python-based system for macromolecular structure solution." Acta Crystallogr D Biol Crystallogr **66**(Pt 2): 213-21.
- Anderson, D. E., R. J. Peters, et al. (1999). "alpha-lytic protease precursor: characterization of a structured folding intermediate." Biochemistry **38**(15): 4728-35.
- Baker, D., J. L. Silen, et al. (1992). "Protease pro region required for folding is a potent inhibitor of the mature enzyme." Proteins **12**(4): 339-44.
- Baker, D., J. L. Sohl, et al. (1992). "A protein-folding reaction under kinetic control." Nature **356**(6366): 263-5.
- Balch, W. E., R. I. Morimoto, et al. (2008). "Adapting proteostasis for disease intervention." Science **319**(5865): 916-9.
- Barrick, D. and R. L. Baldwin (1993). "Three-state analysis of sperm whale apomyoglobin folding." Biochemistry **32**(14): 3790-6.
- Bartik, K., C. Redfield, et al. (1994). "Measurement of the individual pKa values of acidic residues of hen and turkey lysozymes by two-dimensional ¹H NMR." Biophys J **66**(4): 1180-4.
- Baskakov, I. V., G. Legname, et al. (2001). "Folding of prion protein to its native alpha-helical conformation is under kinetic control." J Biol Chem **276**(23): 19687-90.
- Bauer, C. A., R. C. Thompson, et al. (1976). "The active centers of Streptomyces griseus protease 3, alpha-chymotrypsin, and elastase: enzyme-substrate interactions close to the scissile bond." Biochemistry **15**(6): 1296-9.
- Bharadwaj, P. R., A. K. Dubey, et al. (2009). "Abeta aggregation and possible implications in Alzheimer's disease pathogenesis." J Cell Mol Med **13**(3): 412-21.
- Bocharova, O. V., N. Makarava, et al. (2006). "Annealing prion protein amyloid fibrils at high temperature results in extension of a proteinase K-resistant core." J Biol Chem **281**(4): 2373-9.
- Bone, R., D. Frank, et al. (1989). "Structural analysis of specificity: alpha-lytic protease complexes with analogues of reaction intermediates." Biochemistry **28**(19): 7600-9.
- Bone, R., A. Fujishige, et al. (1991). "Structural basis for broad specificity in alpha-lytic protease mutants." Biochemistry **30**(43): 10388-98.
- Bone, R., N. S. Sampson, et al. (1991). "Crystal structures of alpha-lytic protease complexes with irreversibly bound phosphonate esters." Biochemistry **30**(8): 2263-72.
- Bone, R., A. B. Shenvi, et al. (1987). "Serine protease mechanism: structure of an inhibitory complex of alpha-lytic protease and a tightly bound peptide boronic acid." Biochemistry **26**(24): 7609-14.
- Bone, R., J. L. Silen, et al. (1989). "Structural plasticity broadens the specificity of an engineered protease." Nature **339**(6221): 191-5.

- Bonisch, H., C. L. Schmidt, et al. (2002). "The structure of the soluble domain of an archaeal Rieske iron-sulfur protein at 1.1 Å resolution." *J Mol Biol* **319**(3): 791-805.
- Borisov, O. V., M. B. Goshe, et al. (2002). "Low-energy collision-induced dissociation fragmentation analysis of cysteinyl-modified peptides." *Anal Chem* **74**(10): 2284-92.
- Butler, J. S. and S. N. Loh (2005). "Kinetic partitioning during folding of the p53 DNA binding domain." *J Mol Biol* **350**(5): 906-18.
- Butler, J. S. and S. N. Loh (2006). "Folding and misfolding mechanisms of the p53 DNA binding domain at physiological temperature." *Protein Sci* **15**(11): 2457-65.
- Canet, D., M. Sunde, et al. (1999). "Mechanistic studies of the folding of human lysozyme and the origin of amyloidogenic behavior in its disease-related variants." *Biochemistry* **38**(20): 6419-27.
- Chen, H. M., J. L. You, et al. (1991). "Kinetic analysis of the acid and the alkaline unfolded states of staphylococcal nuclease." *J Mol Biol* **220**(3): 771-8.
- Cho, S. S., P. Weinkam, et al. (2008). "Origins of barriers and barrierless folding in BBL." *Proc Natl Acad Sci U S A* **105**(1): 118-23.
- Cooper, J. B., G. Khan, et al. (1990). "X-ray analyses of aspartic proteinases. II. Three-dimensional structure of the hexagonal crystal form of porcine pepsin at 2.3 Å resolution." *J Mol Biol* **214**(1): 199-222.
- Costas, M., D. Rodriguez-Larrea, et al. (2009). "Between-species variation in the kinetic stability of TIM proteins linked to solvation-barrier free energies." *J Mol Biol* **385**(3): 924-37.
- Cunningham, E. L. (2003). Function and destruction of protein folding catalysts : the pro region-mediated folding of alpha-lytic protease and S. Griseus protease B / by Erin Lynn Cunningham. *Biophysics*. San Francisco, University of California, San Francisco. **Ph.D.**
- Cunningham, E. L. and D. A. Agard (2003). "Interdependent folding of the N- and C-terminal domains defines the cooperative folding of alpha-lytic protease." *Biochemistry* **42**(45): 13212-9.
- Cunningham, E. L. and D. A. Agard (2004). "Disabling the folding catalyst is the last critical step in alpha-lytic protease folding." *Protein Sci* **13**(2): 325-31.
- Cunningham, E. L., S. S. Jaswal, et al. (1999). "Kinetic stability as a mechanism for protease longevity." *Proc Natl Acad Sci U S A* **96**(20): 11008-14.
- D'Aquino, J. A., J. Gomez, et al. (1996). "The magnitude of the backbone conformational entropy change in protein folding." *Proteins* **25**(2): 143-56.
- Dai, Y., R. M. Whittal, et al. (1999). "Two-layer sample preparation: a method for MALDI-MS analysis of complex peptide and protein mixtures." *Anal Chem* **71**(5): 1087-91.
- Dao-pin, S., U. Sauer, et al. (1991). "Contributions of engineered surface salt bridges to the stability of T4 lysozyme determined by directed mutagenesis." *Biochemistry* **30**: 7142-7153.
- Davis, I. W., A. Leaver-Fay, et al. (2007). "MolProbity: all-atom contacts and structure validation for proteins and nucleic acids." *Nucleic Acids Res* **35**(Web Server issue): W375-83.

- de Prat Gay, G., C. M. Johnson, et al. (1994). "Contribution of a proline residue and a salt bridge to the stability of a type I reverse turn in chymotrypsin inhibitor-2." Protein Eng **7**(1): 103-8.
- Dee, D. R. and R. Y. Yada "The prosegment catalyzes pepsin folding to a kinetically trapped native state." Biochemistry **49**(2): 365-71.
- Derman, A. I. and D. A. Agard (2000). "Two energetically disparate folding pathways of alpha-lytic protease share a single transition state." Nat Struct Biol **7**(5): 394-7.
- Ecroyd, H. and J. A. Carver (2009). "Crystallin proteins and amyloid fibrils." Cell Mol Life Sci **66**(1): 62-81.
- Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-32.
- Eyring, H. (1935). "The activated complex in chemical reactions." J Chem Phys **3**.
- Fersht, A. R. (1999). Structure and mechanism in protein science : a guide to enzyme catalysis and protein folding. New York, W.H. Freeman.
- Flanagan, M. A., B. Garcia-Moreno, et al. (1983). "Contributions of individual amino acid residues to the structural stability of cetacean myoglobins." Biochemistry **22**(25): 6027-37.
- Friedler, A., D. B. Veprintsev, et al. (2003). "Kinetic instability of p53 core domain mutants: implications for rescue by small molecules." J Biol Chem **278**(26): 24108-12.
- Fuhrmann, C. N. (2005). Structural studies of alpha-Lytic Protease at sub-angstrom resolution reveal insights into the mechanisms of serine protease catalysis and kinetic stability. Biochemistry. San Francisco, University of California, San Francisco. **Ph.D.**
- Fuhrmann, C. N., M. D. Daugherty, et al. (2006). "Subangstrom crystallography reveals that short ionic hydrogen bonds, and not a His-Asp low-barrier hydrogen bond, stabilize the transition state in serine protease catalysis." J Am Chem Soc **128**(28): 9086-102.
- Fuhrmann, C. N., B. A. Kelch, et al. (2004). "The 0.83 Å resolution crystal structure of alpha-lytic protease reveals the detailed structure of the active site and identifies a source of conformational strain." J Mol Biol **338**(5): 999-1013.
- Fujishige Boggs, A. and D. A. Agard (1996). Bacterial extracellular secretion: Transport of alpha-Lytic Protease across the outer membrane of Escherichia Coli. Membrane protein transport : a multi-volume treatise. S. S. Rothman. London, Greenwich, Conn. **3**: 165-179.
- Fushinobu, S., K. Ito, et al. (1998). "Crystallographic and mutational analyses of an extremely acidophilic and acid-stable xylanase: biased distribution of acidic residues and importance of Asp37 for catalysis at low pH." Protein Eng **11**(12): 1121-8.
- Garcia-Arribas, O., R. Mateo, et al. (2007). "Thermodynamic stability of a cold-adapted protein, type III antifreeze protein, and energetic contribution of salt bridges." Protein Sci **16**(2): 227-38.
- Garcia-Mira, M. M., M. Sadqi, et al. (2002). "Experimental identification of downhill protein folding." Science **298**(5601): 2191-5.

- Geierstanger, B., M. Jamin, et al. (1998). "Protonation behavior of histidine 24 and histidine 119 in forming the pH 4 folding intermediate of apomyoglobin." Biochemistry **37**(12): 4254-65.
- Gillespie, J. R. and V. N. Uversky (2000). "Structure and function of alpha-fetoprotein: a biophysical overview." Biochim Biophys Acta **1480**(1-2): 41-56.
- Godoy-Ruiz, R., F. Ariza, et al. (2006). "Natural selection for kinetic stability is a likely origin of correlations between mutational effects on protein energetics and frequencies of amino acid occurrences in sequence alignments." J Mol Biol **362**(5): 966-78.
- Goto, Y. and S. Nishikiori (1991). "Role of electrostatic repulsion in the acidic molten globule of cytochrome c." J Mol Biol **222**(3): 679-86.
- Gotz, J., L. M. Ittner, et al. (2009). "Common features between diabetes mellitus and Alzheimer's disease." Cell Mol Life Sci **66**(8): 1321-5.
- Hermanson, G. (1996). Bioconjugate techniques. San Diego, Calif., Academic Press.
- Higaki, J. N. and A. Light (1986). "Independent refolding of domains in the pancreatic serine proteinases." J Biol Chem **261**(23): 10606-9.
- Hill, C. R. and G. Tomalin (1981). "The kinetics of hydrolysis of some extended N-aminoacyl-L-phenylalanine methyl esters by bovine chymotrypsin A-alpha. Evidence for enzyme subsite S5." Biochim Biophys Acta **660**(1): 65-72.
- Horovitz, A., L. Serrano, et al. (1990). "Strength and co-operativity of contributions of surface salt bridges to protein stability." J Mol Biol **216**(4): 1031-44.
- Hughson, F. M. and R. L. Baldwin (1989). "Use of site-directed mutagenesis to destabilize native apomyoglobin relative to folding intermediates." Biochemistry **28**(10): 4415-22.
- Huyghues-Despointes, B. M., J. M. Scholtz, et al. (1999). "Protein conformational stabilities can be determined from hydrogen exchange rates." Nat Struct Biol **6**(10): 910-2.
- Ionescu, R. M. and M. R. Eftink (1997). "Global analysis of the acid-induced and urea-induced unfolding of staphylococcal nuclease and two of its variants." Biochemistry **36**(5): 1129-40.
- Jaswal, S. S. (2000). Thermodynamics, Kinetics and Landscapes in alpha-Lytic Protease: A Role for Pro Regions in Kinetic Stability., University of California, San Francisco.
- Jaswal, S. S., J. L. Sohl, et al. (2002). "Energetic landscape of alpha-lytic protease optimizes longevity through kinetic stability." Nature **415**(6869): 343-6.
- Jaswal, S. S., S. M. Truhlar, et al. (2005). "Comprehensive analysis of protein folding activation thermodynamics reveals a universal behavior violated by kinetically stable proteases." J Mol Biol **347**(2): 355-66.
- Johnson, S. M., R. L. Wiseman, et al. (2005). "Native state kinetic stabilization as a strategy to ameliorate protein misfolding diseases: a focus on the transthyretin amyloidoses." Acc Chem Res **38**(12): 911-21.
- Kadima, W., L. Ogendal, et al. (1993). "The influence of ionic strength and pH on the aggregation properties of zinc-free insulin studied by static and dynamic laser light scattering." Biopolymers **33**(11): 1643-57.
- Kaganovich, D., R. Kopito, et al. (2008). "Misfolded proteins partition between two distinct quality control compartments." Nature **454**(7208): 1088-95.

- Kashiwagi, T., N. Kunishima, et al. (1997). "The novel acidophilic structure of the killer toxin from halotolerant yeast demonstrates remarkable folding similarity with a fungal killer toxin." Structure **5**(1): 81-94.
- Kelch, B. A. (2007). Identifying the structural determinants of extreme folding and unfolding barriers. Biochemistry. San Francisco, University of California, San Francisco. **Ph.D.**
- Kelch, B. A. and D. A. Agard (2007). "Mesophile versus thermophile: insights into the structural mechanisms of kinetic stability." J Mol Biol **370**(4): 784-95.
- Kelch, B. A., K. P. Eagen, et al. (2007). "Structural and mechanistic exploration of acid resistance: kinetic stability facilitates evolution of extremophilic behavior." J Mol Biol **368**(3): 870-83.
- Kloss, E., N. Courtemanche, et al. (2008). "Repeat-protein folding: new insights into origins of cooperativity, stability, and topology." Arch Biochem Biophys **469**(1): 83-99.
- Li, H., A. D. Robertson, et al. (2005). "Very fast empirical prediction and rationalization of protein pKa values." Proteins **61**(4): 704-21.
- Li, R. and C. Woodward (1999). "The hydrogen exchange core and protein folding." Protein Sci **8**(8): 1571-90.
- Light, A. and A. M. al-Obeidi (1991). "Further evidence for independent folding of domains in serine proteases." J Biol Chem **266**(12): 7694-8.
- Lyu, P. C., P. J. Gans, et al. (1992). "Energetic contribution of solvent-exposed ion pairs to alpha-helix structure." J Mol Biol **223**(1): 343-50.
- Mace, J. E. and D. A. Agard (1995). "Kinetic and structural characterization of mutations of glycine 216 in alpha-lytic protease: a new target for engineering substrate specificity." J Mol Biol **254**(4): 720-36.
- Mace, J. E., B. J. Wilk, et al. (1995). "Functional linkage between the active site of alpha-lytic protease and distant regions of structure: scanning alanine mutagenesis of a surface loop affects activity and substrate specificity." J Mol Biol **251**(1): 116-34.
- Marti, D. N. and H. R. Bosshard (2003). "Electrostatic interactions in leucine zippers: thermodynamic analysis of the contributions of Glu and His residues and the effect of mutating salt bridges." J Mol Biol **330**(3): 621-37.
- McCoy, A. J., R. W. Grosse-Kunstleve, et al. (2007). "Phaser crystallographic software." J Appl Crystallogr **40**(Pt 4): 658-674.
- Meehan, S., Y. Berry, et al. (2004). "Amyloid fibril formation by lens crystallin proteins and its implications for cataract formation." J Biol Chem **279**(5): 3413-9.
- Merritt, E. A. (1999). "Expanding the model: anisotropic displacement parameters in protein structure refinement." Acta Crystallogr D Biol Crystallogr **55**(Pt 6): 1109-17.
- Miller, D. W. and D. A. Agard (1999). "Enzyme specificity under dynamic control: a normal mode analysis of alpha-lytic protease." J Mol Biol **286**(1): 267-78.
- Myers, J. K., C. N. Pace, et al. (1995). "Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding." Protein Sci **4**(10): 2138-48.
- Nozaki, Y. and C. Tanford (1967). "Examination of titration behavior." Methods in Enzymology **11**: 715-734.

- Ota, N. and D. A. Agard (2001). "Enzyme specificity under dynamic control II: Principal component analysis of alpha-lytic protease using global and local solvent boundary conditions." Protein Sci **10**(7): 1403-14.
- Otwinowski, Z. and W. Minor (1997). "Processing of X-ray diffraction data collected in oscillation mode." Methods in Enzymology **276**: 307-326.
- Pace, C. N., B. M. Huyghues-Despointes, et al. (2002). "Charge-charge interactions are the primary determinants of the pK values of the ionizable groups in Ribonuclease T1." Biophys Chem **101-102**: 211-9.
- Park, C., S. Zhou, et al. (2007). "Energetics-based protein profiling on a proteomic scale: identification of proteins resistant to proteolysis." J Mol Biol **368**(5): 1426-37.
- Pearson, K. M., L. K. Pannell, et al. (2002). "Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method." Rapid Commun Mass Spectrom **16**(3): 149-59.
- Peters, R. J., A. K. Shiau, et al. (1998). "Pro region C-terminus:protease active site interactions are critical in catalyzing the folding of alpha-lytic protease." Biochemistry **37**(35): 12058-67.
- Rodriguez-Almazan, C., R. Arreola, et al. (2008). "Structural basis of human triosephosphate isomerase deficiency: mutation E104D is related to alterations of a conserved water network at the dimer interface." J Biol Chem **283**(34): 23254-63.
- Sali, D., M. Bycroft, et al. (1991). "Surface electrostatic interactions contribute little of stability of barnase." J Mol Biol **220**(3): 779-88.
- Salimi, N. L., B. Ho, et al. "Unfolding simulations reveal the mechanism of extreme unfolding cooperativity in the kinetically stable alpha-lytic protease." PLoS Comput Biol **6**(2): e1000689.
- Sauter, N. K., T. Mau, et al. (1998). "Structure of alpha-lytic protease complexed with its pro region." Nat Struct Biol **5**(11): 945-50.
- Schafer, K., U. Magnusson, et al. (2004). "X-ray structures of the maltose-maltodextrin-binding protein of the thermoacidophilic bacterium *Alicyclobacillus acidocaldarius* provide insight into acid stability of proteins." J Mol Biol **335**(1): 261-74.
- Serrano, L., A. Horovitz, et al. (1990). "Estimating the contribution of engineered surface electrostatic interactions to protein stability by using double-mutant cycles." Biochemistry **29**(40): 9343-52.
- Settembre, E. C., J. R. Chittuluru, et al. (2004). "Acidophilic adaptations in the structure of *Acetobacter acetii* N5-carboxyaminoimidazole ribonucleotide mutase (PurE)." Acta Crystallogr D Biol Crystallogr **60**(Pt 10): 1753-60.
- Silen, J. L., D. Frank, et al. (1989). "Analysis of prepro-alpha-lytic protease expression in *Escherichia coli* reveals that the pro region is required for activity." J Bacteriol **171**(3): 1320-5.
- Sohl, J. L. (1997). The folding and inhibition of alpha-lytic protease by its pro region. Biophysics. San Francisco, University of California, San Francisco. **Ph.D.**
- Sohl, J. L., S. S. Jaswal, et al. (1998). "Unfolded conformations of alpha-lytic protease are more stable than its native state." Nature **395**(6704): 817-9.
- Tanford, C. (1970). "Protein denaturation. C. Theoretical models for the mechanism of denaturation." Adv Protein Chem **24**: 1-95.

- Thibodeau, P. H., C. A. Brautigam, et al. (2005). "Side chain and backbone contributions of Phe508 to CFTR folding." Nat Struct Mol Biol **12**(1): 10-6.
- Thirumalai, D., D. K. Klimov, et al. (2003). "Emerging ideas on the molecular basis of protein and peptide aggregation." Curr Opin Struct Biol **13**(2): 146-59.
- Thurkill, R. L., G. R. Grimsley, et al. (2006). "pK values of the ionizable groups of proteins." Protein Sci **15**(5): 1214-8.
- Truhlar, S. M. (2004). Decoupled Folding and Functional Landscapes Allow alpha-Lytic Protease and Streptomyces griseus Protease B to Develop Novel Native-State Properties, University of California, San Francisco.
- Truhlar, S. M. and D. A. Agard (2005). "The folding landscape of an alpha-lytic protease variant reveals the role of a conserved beta-hairpin in the development of kinetic stability." Proteins **61**(1): 105-14.
- Truhlar, S. M., E. L. Cunningham, et al. (2004). "The folding landscape of Streptomyces griseus protease B reveals the energetic costs and benefits associated with evolving kinetic stability." Protein Sci **13**(2): 381-90.
- Vasquez, J. R., L. B. Evin, et al. (1989). "An expression system for trypsin." J Cell Biochem **39**(3): 265-76.
- Watters, A. L., P. Deka, et al. (2007). "The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection." Cell **128**(3): 613-24.
- Wisiz, M. S. and H. W. Hellinga (2003). "An empirical model for electrostatic interactions in proteins incorporating multiple geometry-dependent dielectric constants." Proteins **51**(3): 360-77.
- Xia, K., M. Manning, et al. (2007). "Identifying the subproteome of kinetically stable proteins via diagonal 2D SDS/PAGE." Proc Natl Acad Sci U S A **104**(44): 17329-34.
- Yang, A. S. and B. Honig (1992). "Electrostatic effects on protein stability." Curr Opin Struct Biol **2**: 40-45.
- Young, M. M., N. Tang, et al. (2000). "High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry." Proc Natl Acad Sci U S A **97**(11): 5802-6.
- Yu, B. and D. L. Caspar (1998). "Structure of cubic insulin crystals in glucose solutions." Biophys J **74**(1): 616-22.
- Zhang, Y., J. L. Whittingham, et al. (2002). "Crystallization and preliminary crystallographic investigation of a low-pH native insulin monomer with flexible behaviour." Acta Crystallogr D Biol Crystallogr **58**(Pt 1): 186-7.
- Zubillaga, R. A., E. Garcia-Hernandez, et al. (2006). "Effect of a new ionic pair on the unfolding activation barrier of beta-glucosidase B." Protein Pept Lett **13**(2): 113-8.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Robert Bailey
Author Signature

04/02/2010
Date