

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600

UMI<sup>®</sup>

PREVIEW

CHARACTERIZATION OF FOLDING INTERMEDIATES OF TIM BARREL  
PROTEINS BY HYDROGEN EXCHANGE AND MASS SPECTROMETRY

by

Hai Pan

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree Doctor of Philosophy

Major: Chemistry

Under the Supervision of Professor David L. Smith

Lincoln, Nebraska

December, 2002

UMI Number: 3070132

UMI<sup>®</sup>

---

UMI Microform 3070132

Copyright 2003 by ProQuest Information and Learning Company.

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

---

ProQuest Information and Learning Company  
300 North Zeeb Road  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

DISSERTATION TITLE  
CHARACTERIZATION OF FOLDING INTERMEDIATES OF TIM BARREL  
PROTEINS BY HYDROGEN EXCHANGE AND MASS SPECTROMETRY

---

BY

Hai Pan

---

SUPERVISORY COMMITTEE:

Approved

Date

Mark Griep  
Signature Mark Griep

Dec 5, 02

Typed Name

John Stezowski  
Signature John Stezowski

Dec 05 02

Typed Name

David Smith  
Signature David Smith

Dec 05 02

Typed Name

David Hage  
Signature David Hage

Dec. 5, 02

Typed Name

Robert Spreitzer  
Signature Robert Spreitzer

12/5/02

Typed Name

Signature

Typed Name

**Nebraska** UNIVERSITY OF  
GRADUATE COLLEGE

# CHARACTERIZATION OF FOLDING INTERMEDIATES OF TIM BARREL PROTEINS BY HYDROGEN EXCHANGE AND MASS SPECTROMETRY

Hai Pan, Ph.D.

University of Nebraska, 2002

Advisor: David L. Smith

How the amino acid sequence of a protein specifies its three dimensional structure is referred as protein folding problem. It is the major unsolved problem in structural biology. TIM barrel is one of the most frequently encountered and the most versatile structures among the enzymes. Characterization of the folding intermediates of TIM barrel proteins is critical to solve the folding problem for TIM barrel proteins and it potentially offers insights to the folding of other protein architectures as well.

A new method, based on hydrogen/deuterium exchange, protein fragmentation and mass spectrometry was developed to identify and characterize the structures of folding intermediates at high resolution. In this method, the protein can be fragmented by an acid protease after the protein undergoes H/D exchange. The deuterium level and isotope pattern present in each fragment is determined by HPLC and mass spectrometry.

Equilibrium and kinetic folding intermediates of four TIM barrel proteins have been detected and characterized by this method. The results from rabbit muscle aldolase, a tetrameric TIM barrel protein, showed that folding to the native state involves three steps and two intermediates. Approximately 110 residues fold to highly compact forms in each step. These results also showed that each folding domain includes widely

separated regions of the backbone. Interestingly, two intermediates are also populated in the equilibrium and kinetic folding experiments of monomeric *S. aureus* aldolase, which has high sequence homology with rabbit muscle aldolase. The structures of two intermediates are similar but not the same as those observed in the rabbit muscle aldolase, suggesting that quaternary structure affects folding/unfolding pathways. The sequential folding of two domains was observed in the folding of triosephosphate isomerase (TIM), the prototype in TIM barrel family. The C-terminal four ( $\alpha/\beta$ ) units folds much faster than the N-terminal four ( $\alpha/\beta$ ) units. Two intermediates are characterized in the equilibrium unfolding of  $\alpha$  subunit of bacterial luciferase. The domain that unfolds at lower concentrations of denaturant is located at the C-terminus of the protein including the last two ( $\alpha/\beta$ ) units.

All these results suggest that several different models are required to describe the folding of TIM barrel proteins, despite their high level of structural and functional similarity.

***TO MY PARENTS AND BROTHER,  
FOR THEIR SUPPORT,  
TO MY WIFE, YUAN (VIVIAN) WANG,  
FOR HER UNDERSTANDING AND SACRIFICE***



## ACKNOWLEDGEMENTS

I would like to thank Dr. David Smith, my research supervisor, for his excellent research ideas, patient guidance and tremendous support during my Ph.D. study. I would also like to thank Dr. Jean Smith for many helpful discussions.

I must thank Dr. John Stezowski, Dr. Mark Griep, Dr. David Hage, and Dr. Robert Spreitzer for serving on my supervisory committee and taking time in reading my dissertation.

I am grateful to Dr. Kurt Wulser, Dr. Yuzhong Deng, Dr. John Engen and Ashraf Raza for the technical training and support they've provided to me.

Thanks to other Smith group members such as Dr. Lintao Wang, Dr. Veniamin Lapko, Dr. Patrick Wintrode, Jiong Yu, Casey Swaim and Teerapat Rojsajakul for their sincere help and friendship.

I would also like to thank my parents and my brother back in China. They have been a source of endless support and love for me through all these years.

Most important, I am deeply grateful to my wife, Yuan Wang, for her love, support, and sacrifice. Without her, I can not achieve so far.

## TABLE OF CONTENTS

<b>TITLE PAGE</b> .....	i
<b>ABSTRACT</b> .....	ii
<b>ACKNOWLEDEMENTS</b> .....	iv
<b>TABLE OF CONTENTS</b> .....	v
<b>LIST OF FIGURES &amp; TABLES</b> .....	xi
<b>LIST OF ABBREVIATIONS</b> .....	xviii

### CHAPTER 1

#### Introduction and background: Protein folding & TIM barrel proteins

1.1 Protein folding .....	1
1.1.1 Protein folding problem	
1.1.2 The significance of the protein folding	
1.1.3 The classical and new views of protein folding	
1.1.4 Characterization of protein folding intermediates	
1.2 TIM barrel proteins .....	7
1.2.1 Structural characteristics of the TIM barrel protein	
1.2.2 Folding models of TIM barrel proteins	
1.3 References .....	13

### CHAPTER 2

#### Hydrogen exchange and mass spectrometry: A powerful tool to study the protein folding

2.1 Introduction .....	16
2.2 Background on hydrogen exchange .....	17
2.2.1 Backbone amide hydrogen exchange	
2.2.2 Hydrogen exchange in unstructured peptides	
2.2.3 Amide hydrogen exchange in folded proteins	
2.2.4 Combining hydrogen exchange and mass spectrometry	

2.3 Experimental procedures for HX MS to study the protein folding .....	27
2.3.1 <i>Unfold the protein and initiate folding</i>	
2.3.2 <i>Pulse labeling</i>	
2.3.3 <i>Quench the hydrogen exchange reaction</i>	
2.3.4 <i>Intact protein analysis</i>	
2.3.5 <i>peptic fragments analysis</i>	
2.4 HXMS data analysis .....	39
2.4.1 <i>Mass spec data processing</i>	
2.4.2 <i>Determine the deuterium content</i>	
2.4.3 <i>Analysis of isotopic patterns</i>	
2.5 Reference .....	49

## **CHAPTER 3**

### **Improvements to increase sensitivity of HX MS**

3.1 Introduction .....	55
3.2 Methods to improve the sensitivity of HX MS .....	57
3.2.1 <i>Trapping column help online concentration and desalting</i>	
3.2.2 <i>Peptide mapping column improve the signal and the separation</i>	
3.2.3 <i>Preparation of immobilized pepsin column</i>	
3.3 An online system to integrate digestion, concentration, desalting and separation ..	65
3.3.1 <i>High digestion efficiency of the pepsin column</i>	
3.3.2 <i>Temperature dependence of the digestion efficiency</i>	
3.3.3 <i>Evaluation of deuterium recovery</i>	
3.4 Summary .....	75
3.5 Reference .....	77

## **CHAPTER 4**

### **Probing folding intermediates of rabbit muscle aldolase**

4.1 Introduction .....	79
4.2 Materials and methods .....	82
4.2.1 <i>Refolding aldolase by rapid dilution</i>	
4.2.2 <i>Refolding aldolase by dialysis</i>	
4.3 Results .....	87

4.3.1 Aggregation of aldolase by rapid dilution	
4.3.2 Identification of two intermediates during the dialysis	
4.4 Discussion .....	100
4.4.1 Two intermediates are populated during folding of aldolase monomer	
4.4.2 Quaternary structure affects structure of equilibrium intermediates	
4.4.3 Proposed folding model for aldolase	
4.5 Summary .....	114
4.6 References .....	115

## CHAPTER 5

### Unfolding and refolding of *S.aureus* aldolase, a monomeric TIM barrel protein

5.1 Introduction .....	120
5.2 Materials and methods .....	121
5.2.1 Equilibrium unfolding experiments	
5.2.2 Kinetic folding and unfolding experiments	
5.2.3 Analysis of intact protein and peptic fragments	
5.2.4 Data analysis	
5.3 Results .....	125
5.3.1 Identification of two intermediates in the equilibrium unfolding	
5.3.2 Two intermediates populated in the kinetic folding	
5.3.3 Two intermediates populated in the kinetic unfolding	
5.3.4 Sequence analysis of rabbit muscle and <i>S.aureus</i> aldolase	
5.4 Discussion .....	151
5.4.1 <i>S.aureus</i> Aldolase has three unfolding domains	
5.4.2 Comparison with rabbit muscle aldolase	
5.4.3 Comparison with other folding intermediates of TIM barrel proteins	
5.5 Summary .....	157
5.6 References .....	158

## CHAPTER 6

### Kinetic folding studies of triosephosphate isomerase (TIM)

6.1 Introduction .....	161
------------------------	-----

6.2 Materials and methods .....	163
6.2.1 Materials	
6.2.2 CD spectra	
6.2.3 Gel-filtration chromatography	
6.2.4 Kinetic folding experiments	
6.2.5 Double jump experiments	
6.2.6 Analysis of intact protein and peptic fragments	
6.2.7 Data analysis	
6.3 Results .....	167
6.3.1 Reversibility of the folding	
6.3.2 Identification of a kinetic folding intermediate	
6.3.3 Proline isomerization does not change the folding kinetics	
6.3.4 Identification of the folding domain	
6.4 Discussion .....	187
6.4.1 Structure of the Intermediate	
6.4.2 Intermediate is a monomer	
6.4.3 Kinetic model for the TIM folding	
6.4.4 Structural model for the TIM folding	
6.5 Summary .....	197
6.6 References .....	197

## CHAPTER 7

### Equilibrium and kinetic unfolding studies of $\alpha$ subunit of bacterial luciferase

7.1 Introduction .....	203
7.2 Materials and methods .....	206
7.2.1 Materials	
7.2.2 Equilibrium unfolding and fluorescence analysis	
7.2.3 Equilibrium unfolding and hydrogen exchange	
7.2.4 Kinetic unfolding experiments	
7.2.5 Analysis of the intact $\alpha$ subunit and peptic fragments	
7.2.6 Data analysis	
7.3 Results .....	210
7.3.1 Equilibrium unfolding analyzed by fluorescence	
7.3.2 Equilibrium unfolding analyzed by HX MS	
7.3.3 Kinetic unfolding monitored by HX MS	
7.3.4 Peptic fragment analysis of the $I_1$ populated in 0.5 M urea	

7.3.5 Peptic fragment analysis of the $I_2$ populated in 2.0 M urea	
7.4 Discussion .....	230
7.4.1 Structure of $I_1$	
7.4.2 Structure of $I_2$	
7.4.3 Different paths lead to $I_2$	
7.4 $I_2$ is a kinetically trapped species	
7.5 Summary .....	243
7.6 References .....	243
<b>CHAPTER 8</b>	
<b>Preliminary kinetic folding studies of enolase</b>	
8.1 Introduction .....	247
8.2 Materials and methods .....	248
8.3 Results .....	250
8.4 Discussion .....	255
8.5 Summary .....	256
8.6 References .....	256
<b>Conclusion and Perspective.....</b>	<b>259</b>

## LIST OF FIGURES & TABLES

Description	Page
<b>CHAPTER 1</b>	
Figure 1.1 Schematic representation of a simple folding reaction with one intermediate	2
Figure 1.2 Energy landscape for protein folding showing two-state folding kinetics	5
Figure 1.3 Energy landscape view of multiple-states model	6
Figure 1.4 Topological and schematic diagrams of triosephosphate isomerase (TIM)	11
Table 1.1 Proteins containing TIM barrel fold	9
<b>CHAPTER 2</b>	
Figure 2.1 Three kinds of hydrogen found in proteins	18
Figure 2.2 Effect of pH on the amide hydrogen exchange rate constant	20
Figure 2.3 Protein hydrogen exchange models	23
Figure 2.4 General procedures used for HX MS to study protein folding	28
Figure 2.5 LC/MS setup for the analysis of deuterated protein	34
Figure 2.6 Triple play scheme	36
Figure 2.7 LC/MS setup for the analysis of deuterated peptides	38
Figure 2.8 Representative chromatogram and mass spectra of intact protein in HPLC/MS analysis	40
Figure 2.9 Representative chromatogram and mass spectra of peptic fragments in HPLC/MS analysis	42
Figure 2.10 Mass spectra of a peptic fragment from cytochrome <i>c</i> that is (A) non-deuterated, (B) partially deuterated and (C) fully deuterated	44
Figure 2.11 Mass spectra of a peptic fragment from <i>S.aureus aldolase</i> that was incubated	

in (A) 0 M urea , (B) 1 M urea and (C) 3 M urea for 24 h and then was pulse-labeled for 5 s .....	47
---	----

Table 2.1 Hydrogen/deuterium exchange rate constants, half-lives and pH minima calculated for amide hydrogens located at peptide amide linkages in angiotensin calculated for 0°C and pH 2.5 .....	32
--	----

### CHAPTER 3

Figure 3.1 LC/MS setup with a trapping column to analyze peptic fragments .....	58
Figure 3.2 MS spectra of 0.1 $\mu$ M peptide angiotension I (200 $\mu$ l) using a loop (A) and a trapping column (B) .....	59
Figure 3.3 MS spectra of 0.1 $\mu$ M peptide angiotension I (200 $\mu$ l in 3 M GdHCl and 25 mM DTT) using a loop (A) and a trapping column (B) .....	61
Figure 3.4 Chromatogram of peptic digest of aldolase using a C18 column and a peptide mapping column .....	63
Figure 3.5 Mass spectra and single ion plots of a peptic fragment from aldolase using a C18 column and a peptide mapping column .....	64
Figure 3.6 An Online system to digest labeled proteins and to analyze the peptic fragments by LC/MS .....	68
Figure 3.7 HPLC chromatograms of cytochrome <i>c</i> .....	70
Figure 3.8 Total ion chromatograms of peptic digests of triosephosphate isomerase ...	72
Figure 3.9 Total ion current chromatograms of cytochrome <i>c</i> digested using immobilized pepsin and the on-line system .....	74
Scheme 3.1 Protocol to immobilize pepsin on poros-20AL Media .....	66
Scheme 3.2 Protocol to pack immobilized pepsin column .....	67
Table 3.1 Deuterium recoveries for several peptic fragments from peptic digestion of cytochrome <i>c</i> in solution or in a column packed with immobilized pepsin .....	76



## CHAPTER 4

Figure 4.1 Crystal structure of rabbit muscle aldolase tetramer .....	81
Figure 4.2 Size exclusion chromatography of native aldolase and the refolded aldolase .....	88
Figure 4.3 Electrospray ionization mass spectra of intact aldolase (-38 charge state) folded in 0.8 M urea .....	90
Figure 4.4 Mass spectra of three peptic fragments of aldolase following dialysis of the intact protein for various times in GdHCl/H <sub>2</sub> O .....	93
Figure 4.5 Percentage of aldolase molecules folded in regions represented by peptic fragments as a function of the dialysis time .....	97
Figure 4.6 Percentage of aldolase molecules folded in regions represented by b-series ions produced by collision-induced-dissociation of a peptic fragment (residues 188-203) cleaved from aldolase dialyzed for various times .....	99
Figure 4.7 Aldolase folding progress with dialysis time determined by CD (222 nm) and enzymatic activity .....	101
Figure 4.8 Three dimensional views of aldolase monomers illustrating the folding and unfolding domains .....	104
Figure 4.9 Primary structure of aldolase illustrating locations of secondary structural units, peptic fragments used in this study and the three folding domains .....	105
Figure 4.10 Contact map for backbone segments of aldolase based on its crystal structure PDB 1ADO .....	113
Scheme 4.1 Protocols to measure the activity of the rabbit muscle aldolase .....	84
Table 4.1 Folding of the aldolase backbone represented by 27 peptic fragments .....	96
Table 4.2 Segments of the aldolase backbone that are in one folding domain, but in a different unfolding domain .....	110

## CHAPTER 5

Figure 5.1 Electrospray ionization mass spectra of intact aldolase (+36 charge state) equilibrated in various concentrations of urea .....	126
--	-----

Figure 5.2 Populations of folded (F), two intermediates ( $I_1$ and $I_2$ ) and Unfolded (U) <i>S. aureus</i> aldolase equilibrated in different concentrations of urea .....	130
Figure 5.3. Free energy change of each unfolding step of aldolase versus urea concentration .....	132
Figure 5.4 Electrospray ionization mass spectra of three peptic fragments derived from intact aldolase equilibrated in various concentrations of urea .....	134
Figure 5.5 Populations of aldolase molecules unfolded in three unfolding domains as a function of the concentration of urea .....	140
Figure 5.6 Mass spectra of the peptic fragment 56-73 derived from intact aldolase equilibrated in various concentrations of urea .....	142
Figure 5.7 Electrospray ionization mass spectra of intact aldolase (+36 charge state) folded in 0.3 M urea .....	145
Figure 5.8. Electrospray ionization mass spectra of intact aldolase (+36 charge state) unfolded in 3 M and 2.5 M urea .....	148
Figure 5.9 Primary structures of <i>S. aureus</i> and rabbit muscle aldolase illustrating their homology, secondary structure and unfolding domains .....	150
Figure 5.10. Three unfolding domains in the <i>S. aureus</i> and rabbit muscle aldolase ...	155
Table 5.1 Peptides from the peptic digest of <i>S. aureus</i> aldolase identified by the MS/MS .....	124
Table 5.2 Deuterium levels found in native state (F), the two equilibrium unfolding intermediates ( $I_1$ , $I_2$ ) and the unfolded state (U) of <i>S. aureus</i> aldolase equilibrated in different concentrations of urea .....	128
Table 5.3 Deuterium levels of folded ref, unfolded ref, low-mass envelope and high-mass envelope for 28 peptic fragments from <i>S. aureus</i> aldolase .....	136
Table 5.4 The $C_m$ value for peptic fragments of aldolase from <i>S. aureus</i> .....	139
Table 5.5 Deuterium levels of native state (F), two intermediates ( $I_1$ , $I_2$ ) and unfolded state (U) of <i>S. aureus</i> aldolase during folding in 0.3 M urea .....	146

## CHAPTER 6

Figure 6.1 Far-UV CD spectra of native TIM and refolded TIM .....	168
Figure 6.2 Gel-filtration analysis of native TIM and refolded TIM .....	169
Figure 6.3 Mass spectra of intact TIM folding in 0.3 M GdHCl .....	171
Figure 6.4 Fractional populations of the unfolded, folded and intermediate as a function of the folding time .....	174
Figure 6.5 Mass spectra of intact TIM folding in 0.8 M urea .....	175
Figure 6.6 Mass spectra of intact TIM folding in 0.4 M GdHCl .....	177
Figure 6.7 Mass spectra of intact TIM folded in 0.5 M GdHCl .....	178
Figure 6.8 Fractional populations of the unfolded, folded and intermediate as a function of the folding time. (A) Folding in 0.4 M GdHCl (B) Folding in 0.5 M GdHCl .....	179
Figure 6.9 Fractional population of the intermediate as a function of the folding time in the double experiment .....	182
Figure 6.10 Mass spectra of two peptic fragments derived from intact TIM folding in 0.3 M GdHCl .....	183
Figure 6.11 Fractional populations of peptic fragments including residues 191-204 and 25-38 as a function of folding time .....	185
Figure 6.12 Primary structures of TIM from human and rabbit muscle illustrating the locations of secondary structural units .....	189
Figure 6.13 Three-dimensional views of a TIM monomer. Slow folding and Fast folding domains are indicated by the green and red, respectively .....	191
Table 6.1. The deuterium levels in folded dimer ( $F_D$ ), partially folded intermediate ( $I$ ), unfolded monomer ( $U_M$ ) .....	173
Table 6.2 The folding rate constants and their amplitudes for 21 peptic fragments from TIM .....	186

## Chapter 7

Figure 7.1 Topographic diagram of the secondary structure of the $\alpha$ and $\beta$ subunit of luciferase .....	204
Figure 7.2 Equilibrium denaturation curve analyzed by fluorescence .....	211
Figure 7.3 Mass spectra of intact $\alpha$ subunit incubated in various concentrations of urea .....	212
Figure 7.4 The fractional populations of various structural forms as a function of urea .....	216
Figure 7.5 Mass spectra of intact $\alpha$ subunit unfolding in 1.0 M urea .....	218
Figure 7.6 Mass spectra of intact $\alpha$ subunit unfolding in 1.5 M urea .....	220
Figure 7.7 Mass spectra of intact $\alpha$ subunit unfolding in 2.5 M urea .....	222
Figure 7.8 Mass spectra of the intact $\alpha$ subunit unfolding in 3.5 M urea .....	223
Figure 7.9 Mass spectra of three peptic fragments including residues 310-322, 175-192 and 280-296 taken from the intact $\alpha$ subunit that incubated in 0.5 M urea and labeled with D <sub>2</sub> O for 5 s .....	224
Figure 7.10 Mass spectra four peptic fragments including residues 37-49, 50-75, 175-192 and 310-322 taken from the intact $\alpha$ subunit that incubated in 2.0 M urea and labeled with D <sub>2</sub> O for 5 s .....	229
Figure 7.11. The 3-D structure of the $\alpha$ subunit .....	232
Figure 7.12 The percentage of the unfolding populations of fragments in the $\alpha$ subunit incubated in 2 M urea .....	236
Figure 7.13 The 3-D representation of luciferase heterodimer .....	237
Figure 7.14 Schematic diagram of free energy surface of the refolding and unfolding from the $\alpha$ subunit equilibrated in 2 M urea .....	241

Table 7.1. Deuterium levels of native state (F), two partially unfolded forms ( $I_1$ , $I_2$ ) and unfolded state (U) of the $\alpha$ subunit in different concentrations of urea .....	215
--	-----

Table 7.2. Deuterium levels of native state (F), three partially unfolded forms ( $P_1$ , $P_2$ , $I_2$ ) of the $\alpha$ subunit unfolded in 1.0 M urea .....	219
--	-----

Table 7.3. The percentage of the unfolded populations of different peptic fragments in the $\alpha$ subunit .....	228
---	-----

## Chapter 8

Figure 8.1. The crystal structure of enolase monomer from baker's yeast .....	249
---	-----

Figure 8.2 The time course study of enolase folding from 2 s to 600 s followed by pulse-labeling hydrogen exchange and mass spectrometry .....	251
--	-----

Figure 8.3 The fractional populations of F, $I_1$ and $I_2$ as a function of the folding time .....	254
---	-----

Table 8.1. Deuterium levels of native state (F), two intermediates ( $I_1$ , $I_2$ ) and the unfolded state (U) of enolase from baker's yeast found during folding .....	253
--	-----

## LIST OF ABBREVIATIONS

ACN	acetonitrile
CD	circular dichroism
CID	collision induced dissociation
Cyt C	cytochrome c
E/S	enzyme substrate ratio
ESI	electrospray ionization
GdHCl	guanidine hydrogen chloride
HisF	imidazoleglycerol phosphate synthase
HPLC	high performance liquid chromatography
HX	hydrogen exchange
HX MS	hydrogen exchange and mass spectrometry
LCMS	liquid chromatography and mass spectrometry
LEM	linear extrapolation method
MW	molecular weight
MS	mass spectrometry
M/z	mass to charge ratio
NMR	nuclear magnetic resonance
PDB	protein data bank
S/N	signal to noise ratio
TIM	triosephosphate isomerase
TFA	trifluoroacetic acid
w/w	weight to weight ratio

## Chapter 1

### Introduction and Background:

### Protein folding & TIM barrel proteins

#### 1.1 Protein Folding

##### *1.1.1 Protein folding problem*

Proteins are the molecules of life, responsible for a wide range of fundamental biological processes such as catalyzing the numerous chemical reactions, transport, signaling and energy conversion. Proteins are synthesized on the ribosomes as linear chains of amino acids. To become fully functional, all proteins must fold to unique three-dimensional structures. The question of how proteins efficiently and reliably achieve their native states following synthesis is referred as protein folding problem, which is one of the most intriguing questions in structural biology.

In 1973, Anfinsen established that the sequence of the protein contains all the information necessary to direct the unfolded protein to reach the native state (Anfinsen, 1973). However, the practice of this rule is not simple, which is illustrated by a very famous question known as the Levinthal paradox (Dill and Chan, 1997). One simple way to explain this problem follows. If it is assumed that each of the backbone  $\phi$  and  $\psi$  angles in a polypeptide chain can adopt only three different values, there will be  $3^{100}$  possible conformations for the protein with 100 residues. If only  $10^{-11}$  s is required to change from one conformation to the other, a random search of all conformations will take about  $10^{29}$  years, which is far more than the time since the earth was born. However, proteins are frequently observed to fold on the seconds to hours scale.

The classical view to explain the Levinthal paradox is that the protein folding is not the random search. Instead, the protein passes through a limited number of intermediates (Fig 1.1). These intermediates guide the protein to reach the native state.

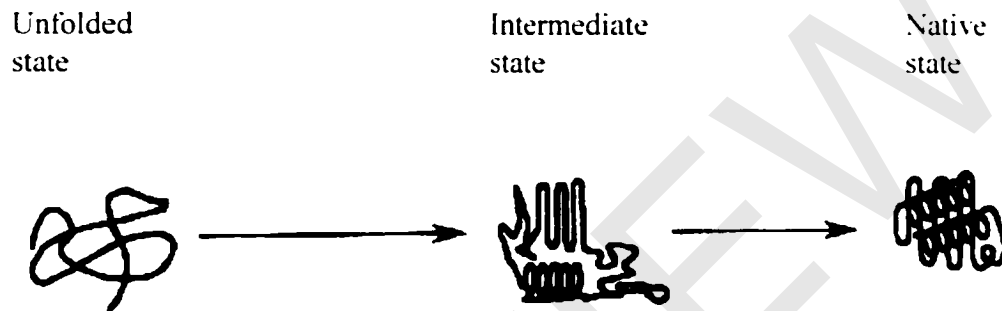


Figure 1.1 Schematic representation of a simple folding reaction with one intermediate

### 1.1.2 The significance of the protein folding

The protein folding is not just a fascinating challenge from an intellectual point of view, it has important practical implications. With the completion of the human genome project, solving protein folding problem is even more urgent. Tremendous efforts have been made to predict three dimensional structures from amino acid sequences. In addition, protein misfolding is found to relate to a number of diseases, such as Parkinson and Alzheimer's diseases (Dobson, 1999). For example, misfolded  $\beta$ -amyloid peptides form insoluble fibrils in the human brain, which is likely the cause of the Alzheimer's disease (Walsh, Lomakin et al., 1997; Walsh, Hartley et al., 1999). There is also