Biophysical Characterisation of the Pathological Z Variant of α_1-Antitrypsin

Anja Sylvia Knaupp

Diplom-Industriebiologin (FH)

Submitted for the Degree of Doctor of Philosophy

Department of Biochemistry and Molecular Biology,
Monash University, Melbourne, Australia

July 2012

Notice 1

Under the Copyright Act 1968, this thesis must be used only under the normal conditions of scholarly fair dealing. In particular no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this thesis.

Notice 2

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Table of Contents

Table of Contents iii

Summary vii

General Declaration ix

Acknowledgements xi

Abbreviations xiii

Chapter 1 Introduction 1

1.1 Protein folding and misfolding 3
- 1.1.1 Classical models of protein folding ____ 3
- 1.1.2 The energy landscape view of protein folding ____ 4
- 1.1.3 Metastable proteins ____ 6
- 1.1.4 Protein misfolding ____ 7
- 1.1.5 The conformational diseases ____ 8

1.2 The serpins 11
- 1.2.1 Serpin nomenclature ____ 11
- 1.2.2 Serpin structure ____ 13
- 1.2.3 Serpin mechanism of proteinase inhibition ____ 15
- 1.2.4 Serpin metastability and its role in conformational change ____ 18
- 1.2.5 Alternative high stability serpin states ____ 21
  - 1.2.5.1 Uncleaved monomeric high stability states ____ 21
  - 1.2.5.2 Polymeric high stability states ____ 23
- 1.2.6 Mechanisms of serpin folding and misfolding ____ 26
  - 1.2.6.1 Folding into the metastable native state ____ 26
  - 1.2.6.2 Misfolding into polymers ____ 27
- 1.2.7 α_1AT deficiency and the pathological Z variant of α_1AT ____ 29

1.3 Aims of this study 31

Chapter 2 Materials and Methods 33

2.1 Reagents 35

2.2 Microbiological techniques 35
- 2.2.1 *Escherichia coli (E. coli)* ____ 35
  - 2.2.1.1 Media and agar plates ____ 35
  - 2.2.1.2 Strains ____ 36
  - 2.2.1.3 Preparation of competent cells ____ 36
  - 2.2.1.4 Transformation ____ 36
  - 2.2.1.5 Glycerol stock preparation ____ 37

2.2.2 *Pichia pastoris (P. pastoris)* 37
    2.2.2.1 Media and agar plates 37
    2.2.2.2 Strains 38
    2.2.2.3 Preparation of electrocompetent cells 38
    2.2.2.4 Transformation 38
    2.2.2.5 Glycerol stock preparation 38
2.3 Molecular biological techniques 39
2.3.1 Cloning and expression vectors 39
2.3.2 Primers 39
2.3.3 Determination of DNA concentration 40
2.3.4 Site-directed mutagenesis 40
    2.3.4.1 Pfu DNA polymerase site-directed mutagenesis 40
    2.3.4.2 KOD DNA polymerase site-directed mutagenesis 41
2.3.5 Agarose gel electrophoresis 41
    2.3.5.1 Solutions 41
    2.3.5.2 Procedure 42
2.3.6 Restriction enzyme digests 42
    2.3.6.1 DpnI digest 42
    2.3.6.2 EcoRI digest 43
    2.3.6.3 SacI digest 43
2.3.7 DNA purification from agarose gels 44
2.3.8 Dephosphorylation 44
2.3.9 Ligation 44
2.3.10 Plasmid purification 45
2.3.11 DNA sequencing 45
2.3.12 Ethanol precipitation 46
2.4. Preparation of α_1AT....46
2.4.1 α_1AT small scale expression 46
2.4.2 α_1AT large scale expression 47
2.4.3 α_1AT purification 47
2.4.4 Determination of α_1AT concentration 48
2.5 Polyacrylamide gel electrophoresis (PAGE)....49
2.5.1 Continuous Native PAGE 49
    2.5.1.1 Solutions 49
    2.5.1.2 Procedure 49
2.5.2 Discontinuous Native PAGE 50
    2.5.2.1 Solutions 50
    2.5.2.2 Procedure 50
2.5.3 SDS PAGE 51
    2.5.3.1 Solutions 51
    2.5.3.2 Procedure 52

2.6 Western blot 52
2.6.1 Solutions 52
2.6.2 Procedure 53
2.7 Size exclusion chromatography 53
2.8 Preparation of BSA coated plates..54
2.9 Determination of α_1AT inhibitory properties...54
2.9.1 Solutions 54
2.9.2 Stoichiometry of inhibition (SI) 54
2.9.3 Continuous association rate constant determination (k_{ass}) 55
2.10 Circular Dichroism (CD) techniques 56
2.10.1 Spectral measurements 56
2.10.2 Thermal denaturation 56
2.11 Fluorescence spectroscopy 57
2.11.1 Tryptophan fluorescence 57
2.11.2 bis-ANS fluorescence 57
2.11.3 Quenching 58
2.11.4 Anisotropy 58
2.12 Equilibrium unfolding studies 58
2.13 Kinetic unfolding studies..59
2.14 Polymerisation studies...59

Chapter 3 Kinetic Instability of the Serpin Z α_1-Antitrypsin Promotes Aggregation 61

Chapter 4 Structural Change in β-Sheet A of Z α_1-Antitrypsin is Responsible for Accelerated Polymerisation and Disease 75

Chapter 5 Conformational Properties of the Disease-Causing Z Variant of α_1-Antitrypsin Revealed by Theory and Experiment 91

Chapter 6 General Discussion...105

References...115

Appendix...135

Summary

Protein misfolding is associated with a range of diseases and occurs when a protein meanders from its normal folding pathway resulting in the formation of a non-native state that can self-associate. One protein superfamily commonly associated with misfolding and disease are the serpins (serine proteinase inhibitors). Several members of this superfamily are prone to self-association which is linked to a range of diverse disorders including emphysema, liver disease, angioedema, neurodegeneration and thrombosis. The serpins are particularly susceptible to misfolding and aggregation as they are metastable which means their native state is not the thermodynamic minimum for the polypeptide chain. The inherent tension sustained in the native state is necessary for protease inhibition, however, mutations can easily result in conformational rearrangements which lead to polymeric states with significantly increased stability.

α_1-Antitrypsin (α_1AT) is the serpin most commonly associated with misfolding and aggregation as there are several naturally occurring variants linked to disease. Z α_1AT is the most common pathological variant of α_1AT characterised by a glutamate to a lysine substitution at amino acid position 342 (Glu342Lys) which results in the loss of a salt bridge to Lys290. In healthy individuals, α_1AT is expressed in hepatocytes and secreted into circulation to control the proteolytic activity of neutrophil elastase in the lower respiratory tract. The Z mutation leads to an increased propensity of α_1AT to polymerise at its place of synthesis, the endoplasmic reticulum of hepatocytes, which is associated with loss-of-function and gain-of-toxic-function mechanisms.

Why the Z mutation renders α_1AT prone to polymerisation is not known and the data available on Z α_1AT is limited and conflicting. This thesis has therefore focused on the *in vitro* characterisation of Z α_1AT to determine the effect of the Z mutation on the α_1AT molecule. An extensive biophysical analysis of Z α_1AT was conducted and revealed that the molecule adopts an altered but functional conformation. In order to map the conformational change induced by the Z mutation onto the α_1AT molecule single tryptophan mutants were created and subjected to an extensive fluorescence spectroscopic analysis. The single tryptophan data indicate that the Z mutation leads to a conformational rearrangement in the top of β-sheet A whereas the structural integrity of β-sheet B is unaffected. This structural rearrangement of β-sheet A does, however, not result in a significant destabilisation of Z α_1AT in comparison to the wild-type. This observation is further supported by molecular dynamics simulations which suggest that the Z mutation results in the formation of new interactions that compensate for the lost interactions. Analysis of the role of the side chain charge of amino acid 342 revealed that it is not solely the loss of the 290-342 salt bridge but also the positive charge of Lys342 in Z α_1AT that leads to the adoption of an alternative conformation and its increased propensity to polymerise. Together, the data presented in this thesis suggest that the Z mutation results in a decrease in the barrier height to aggregation-prone species and consequently in increased polymer formation.

General Declaration

In accordance with Monash University Doctorate Regulation 17/ Doctor of Philosophy and Master of Philosophy (MPhil) regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes three original papers published in peer reviewed journals and no unpublished publications. The core theme of the thesis is the biophysical characterisation of the pathological Z variant of α_1-antitrypsin. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the Department of Biochemistry and Molecular Biology under the supervision of Prof Stephen Bottomley.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of chapter 3, 4 and 5 my contribution to the work involved the following:

General Declaration

Thesis chapter	Publication title	Publication status*	Nature and extent of candidate's contribution
3	Kinetic Instability of the Serpin Z α_1-Antitrypsin Promotes Aggregation	Published	Experimental design, protein preparation, carried out experiments, data analysis, wrote the first draft of the manuscript and edited the final version.
4	Structural Change in β-Sheet A of Z α_1-Antitrypsin Is Responsible for Accelerated Polymerization and Disease	Published	Experimental design, protein construction and preparation, carried out all experiments, data analysis, wrote the first draft of the manuscript and edited the final version.
5	Conformational Properties of the Disease-Causing Z Variant of α1-Antitrypsin Revealed by Theory and Experiment	Published	Experimental design for *in vitro* experiments, protein construction and preparation, *in vitro* experiments, *in vitro* data analysis and manuscript preparation.

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Signed: [redacted].. **Date:** …26.10.2012…

Acknowledgements

The past few years spent on my thesis have been arduous but rewarding. I have learnt a lot about science and life in general and would like to thank the people without whom that would not have been possible.

Firstly, I would like to thank my supervisor Prof Steve Bottomley for all his help, support, patience and calming influence which were invaluable in completing this thesis. He always encouraged me to overcome my fear of presenting my work at conferences and to believe in my capabilities, which in hindsight is very much appreciated.

I am indebted to Vita Levina who always went beyond the call of duty in helping me understand concepts and equations. My deep thanks to Amy Robertson for our scientific discussions, her advice on English grammar and for calming me down when everything seemed too difficult. I would further like to thank Nik Sotirellis for always going out of his way to find a free FPLC for me. Thanks also to the people in the Bottomley lab who provided a constructive and supportive working environment, in particular, Mary Pearce, Victoria Hughes, Su-Ling Leck and Austin Powers. I probably owe an apology to Mary and Victoria as it must have been hard to share an office with me over the last couple of months.

I would further like to thank Dr Itamar Kass and A/Prof Ashley Buckle for the collaboration and the molecular dynamics simulation. Additionally, thanks to Dr Keith Al-Hasani for introducing me to Steve and his lab.

Thanks also to my friends outside the lab who were always on hand with some words of support and encouragement or a nana-blanket and a cup of tea. These include Luisa Müller, Martina Hurlebaus, Nadine Heidhoff, Sandra Heinrich, Sandra Leffers, Tim Holloway and my sister Claudia Brender. And of course Luke Cossins who probably feels he has been through a PhD himself. I would like to thank him for being so understanding and encouraging and for making me that great cup of coffee that got me going in the morning.

Last but not least I would like to thank my parents for all their love and support and for believing in me. .

Abbreviations

α_1ACT	α_1-Antichymotrypsin
α_1AT	α_1-Antitrypsin
α_2AP	α_2-Antiplasmin
Aβ	Amyloid β
A_{214}	Absorbance at 214 nm
A_{280}	Absorbance at 280 nm
A_{405}	Absorbance at 405 nm
AGT	Angiotensinogen
ALP	Autophagy-lysosome pathway
APP	Amyloid precursor protein
APS	ammonium persulfate
AT	Antithrombin
bis-ANS	4, 4’-Dianilino-1, 1’-binaphthyl-5, 5’-disulfonic acid
BME	β-Mercaptoethanol
BSA	Bovine serum albumin
C1 INH	C1 inhibitor
CAP	Cytoplasmic antiprotease
CBG	Cortisol binding globulin
CBP	Collagen binding protein
CD	Circular dichroism
CrmA	Cytokine response modifier A
CV	Column volumes
ddH_2O	Double-distilled water
D_m	Midpoint of denaturation transition
DMSO	Dimethyl sulfonic acid
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
GDN	Glia derived nexin
GdnHCl	Guanidium hydrochloride
GdnSCN	Guanidinium thiocyanate

h	Hours
HCII	Heparin cofactor II
HNE	Human neutrophil elastase
Hsp47	Heat shock protein 47
htt	Huntingtin
I	Folding intermediate
IAPP	Islet amyloid polypeptide
λ_{em}	Emission wavelength
λ_{ex}	Excitation wavelength
λ_{max}	Wavelength at emission maximum
k_{ass}	Association rate constant
k_{assapp}	Apparent association rate constant
LEI	Leukocyte elastase inhibitor
MENT	Myeloid and Erythroid Nuclear Termination stage-specific protein
min	Minutes
MNEI	Monocyte neutrophil elastase inhibitor
N	Native state
OD_{260}	Optical density at 260 nm
OD_{600}	Optical density at 600 nm
P	Polymer
PAI-1	Plasminogen activator inhibitor-1
PAI-2	Plasminogen activator inhibitor-2
PCI	Protein C inhibitor
PEDF	Pigment epithelium derived factor
PEG-8000	Polyethylene glycol
PI	Proteinase inhibitor
PMSF	Phenylmethylsulphonyl fluoride
PN1	Protease nexin 1
PS1	Presenilins-1
PS2	Presenilins-2
PTI	Placental thrombin inhibitor
R	“Relaxed” serpin conformation
RCL	Reactive centre loop

s	Seconds
S	“Stressed” serpin conformation
s1/..6A/..C	Strand 1/..6 of β-sheet A/..C
SCCA1	Squamous cell carcinoma antigen-1
SCCA2	Squamous cell carcinoma antigen-2
SDS	Sodium dodecyl sulfate
Serpin	Serine proteinase inhibitor
SI	Stoichiometry of inhibition
θ_{222}	CD signal at 222 nm
TBG	Thyroxine binding globulin
TEMED	N,N,N’,N’-Tetramethylethylenediamine
T_m	Midpoint of thermal denaturation transition
tPA	Tissue-type plasminogen activator
TTR	Transthyretin
TUG	Transverse urea gradient
U	Unfolded state
uPA	Urokinase-type plasminogen activator
UPR	Unfolded protein response
UPS	Ubiquitin-proteasome system
v	Volume
w	Weight
ZPI	Protein Z-dependent proteinase inhibitor

Chapter 1

Introduction

1.1 Protein folding and misfolding

Protein folding is the process by which a polypeptide chain adopts its characteristic three-dimensional structure. The question of how the functional native state of a protein is obtained with such high fidelity considering the vast amount of possible conformations the polypeptide chain could fold into has intrigued scientists for many decades. That protein folding cannot be completely random was pointed out by Levinthal in 1969 who noted that if a folding polypeptide chain was to search randomly through all possible conformations the correct native fold could not be attained in a realistic time frame [1]. That it is the amino acid sequence that dictates the folding process became evident from pioneering work of Anfinsen. In 1973, he showed that ribonuclease can refold spontaneously into its functional native conformation after denaturation and cleavage of disulphide bonds [2]. A corollary to Anfinsen's work is that protein folding is under thermodynamic control and research has since progressed to determine the rules that govern protein folding.

1.1.1 Classical models of protein folding

Several models have been proposed to describe the mechanisms by which a folding polypeptide chain adopts its native conformation. According to the nucleation-growth model the folding process is based on a nucleation event; the formation of structural units that form a scaffold for the remaining structure to rapidly fold on [3]. The framework model assumes the formation of secondary structural elements followed by their docking in the rate-limiting step to give rise to tertiary interactions [4, 5]. The hydrophobic collapse model is based on the idea that the driving force of protein folding is the collapse of the polypeptide chain to exclude water in a nonspecific manner, followed by the rearrangement of the compacted state in the rate-limiting step [6, 7]. According to the jigsaw puzzle model, there is not a unique folding mechanism but evolution has produced multiple routes for a folding polypeptide chain to adopt its native conformation [8]. This folding model shares some similarities with the diffusion-collision model [9], a model that was reconsidered nearly two decades after its proposal in order to include more recent experimental data [10]. The diffusion-collision model proposes that several fluctuating portions of incipient secondary structure or hydrophobic clusters, so-called microdomains, move diffusely and collide with each other repeatedly resulting in their coalescence into larger units. Several diffusion-collision encounters therefore result in a structure with parts of the backbone close to the native conformation which, in the final folding step, then folds into the exact tertiary structure.

Some of these classical folding models have been subject to controversy as they cannot accommodate all of the experimental data obtained over time. The nucleation-growth model, for example, has been contradicted by the observation of folding intermediates: partially organised states, which have been identified on the folding pathways of a large number of proteins [11, 12]. On the other hand, the discovery that protein folding can occur via a simple two-state transition [13] and that secondary and tertiary structural elements can form simultaneously as the protein undergoes a general collapse as shown by Φ-value analysis of the transition state [14] led to criticism of some of the other folding models including the hydrophobic collapse, the framework and the diffusion-collision model. As an alternative, the more recent nucleation-condensation model was proposed [15]. This model is generally accepted and assumes that otherwise weak secondary structural units are stabilised through long-range and other native hydrophobic interactions in the transition state. It has been suggested that the hydrophobic collapse and the framework models present two extreme situations of the nucleation-condensation model and that the folding of most proteins corresponds to a mechanism in between the two [16]. While there are several different models for protein folding, each of which is supported by experimental data, they might still all be true as different proteins could fold via different mechanisms and it remains to be elucidated whether a unifying folding mechanism actually exists [16].

1.1.2 The energy landscape view of protein folding

Experimental evidence suggests that the folding pathway of a polypeptide chain may include numerous transition states and intermediates on the way to the native state (Figure 1.1). A two-state folding mechanism has been proposed for many small proteins (less than 100 amino acids) as no stable intermediate species can be detected experimentally (Figure 1.1a) [15]. Larger proteins commonly sample intermediate states that can be productive or obstructive during the folding process (Figure 1.1b) [17].

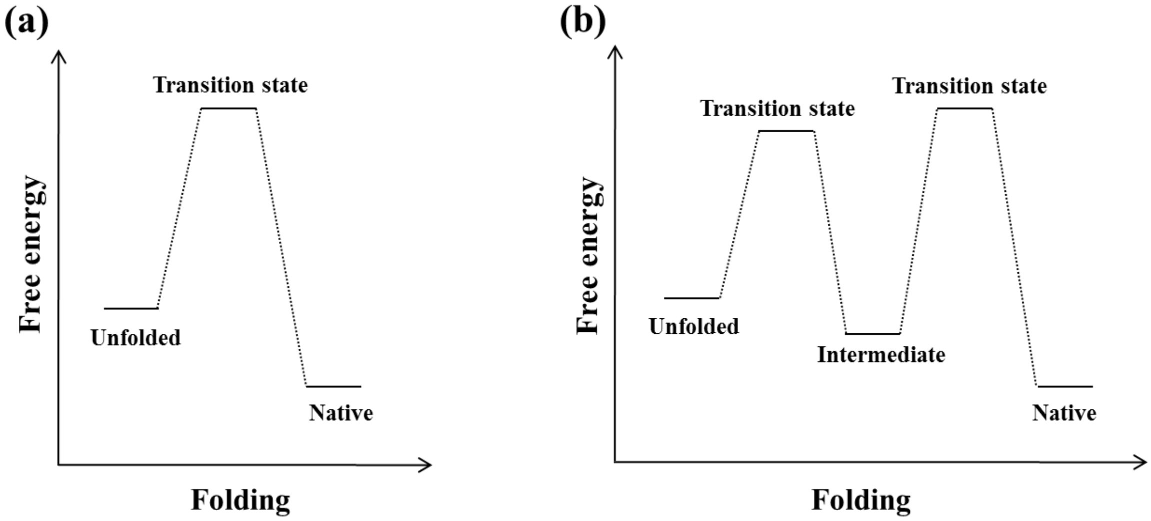


Figure 1.1. Free energy diagrams of protein folding. The higher energy unfolded state proceeds through at least one transition state ensemble, an array of unstable states, before attaining the energetically more favourable native conformation. (a) A two-state folding pathway diagram representing folding via a single transition. (b) A three-state folding pathway diagram representing folding via the formation of an intermediate ensemble, a transiently stable, partially folded state populated on the way to the native conformation.

Folding funnels have been used to describe the theoretical energy landscape of protein folding (Figure 1.2) [18]. This view of protein folding assumes that the folding process is driven by the requirement to minimise the free energy as the number of native contacts increases [19]. The ensemble of unfolded conformations, stabilised by the conformational entropy of the unfolded chain, adopt the native structure possibly via multiple parallel pathways [20]. The formation of native contacts, which are generally energetically more favourable than non-native ones, restrict the conformational space of the folding polypeptide chain and result in a reduction in free energy [21]. Under physiological conditions the difference in free energy between the native and unfolded states of most proteins averages between 5-15 kcal/ mol [22]. An idealised smooth folding funnel describes the energy landscapes of a protein that folds via a two-state transition from the unfolded to the native state (Figure 1.2a). For most proteins energy landscapes of folding are rugged with kinetic traps caused by local energy barriers which can lead to intermediate populations and multistate folding mechanisms (Figure 1.2b) [19].

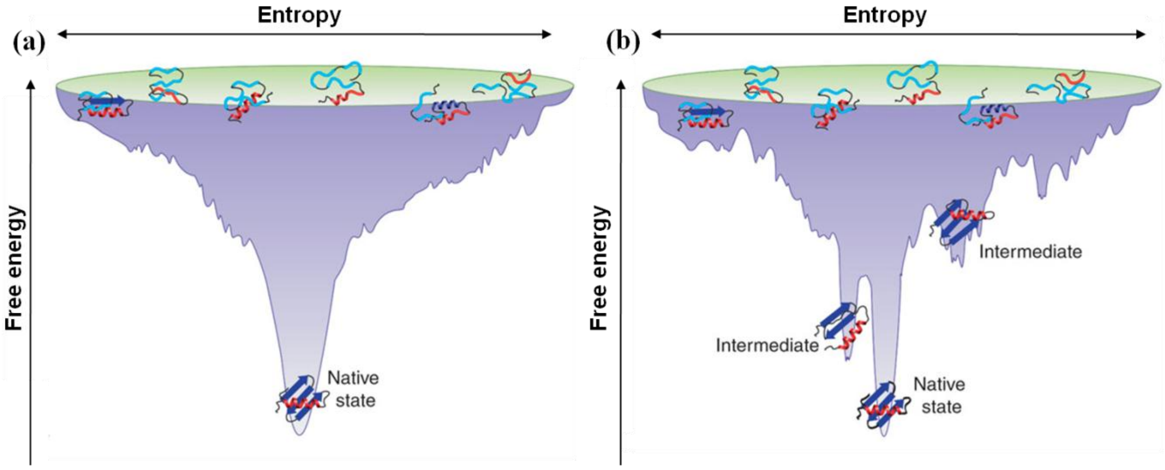


Figure 1.2. Examples of folding funnels. Folding funnels guide the transition from the unfolded state, an ensemble of conformations high in free energy, to the native conformation which is low in free energy. (a) A smooth folding funnel with no deep dips or barriers to interrupt the transition to the native state depicts a two-state folding transition. (b) A rough folding funnel represents a rugged energy landscape which can lead to the formation of transiently stable folding intermediates populated on the way to the native state. In a folding funnel intermediate ensembles would appear as dips and transition state ensembles as peaks. Adapted from [19].

1.1.3 Metastable proteins

Some proteins adopt metastable native states, whereby the native state is essentially a kinetically trapped folding intermediate. The finding that the native conformation is not always the state with the lowest possible free energy suggests that protein folding is under thermodynamic and kinetic control. Metastability occurs when the energy landscape of folding contains minima that are lower in free energy than the native state but the kinetic barrier to attain these states is too high to be overcome. Changes to the system such as a decrease in pH, the binding of a pro region or the cleavage of a peptide bond can allow access to these thermodynamically more favourable states and therefore mestastability often plays a regulatory role in the biological function of those proteins [23, 24]. Metastable protein folding has been observed for several proteins including α-lytic protease [25], influenza haemagglutinin [26], subtilisin [27, 28], gp120 and gp41 [29], luciferase [30, 31] and most serine proteinase inhibitors (serpins) [32].

1.1.4 Protein misfolding

Even though folding into the functional native state takes place with high fidelity, protein folding in the reactive cellular environment, where folding occurs cotranslationally, can be challenged by several factors including macromolecular crowding, temperatures of 37 °C and above and the presence of heavy metals and oxygen free radicals [33, 34]. The competition between intra-molecular and inter-molecular contacts results in a drastic increase in landscape ruggedness and conformational space of a folding polypeptide chain (Figure 1.3) [35]. In order to ensure efficient protein folding under these conditions evolution has produced several protein quality control systems that assist the protein folding process or are involved in the clearance of misfolded proteins.

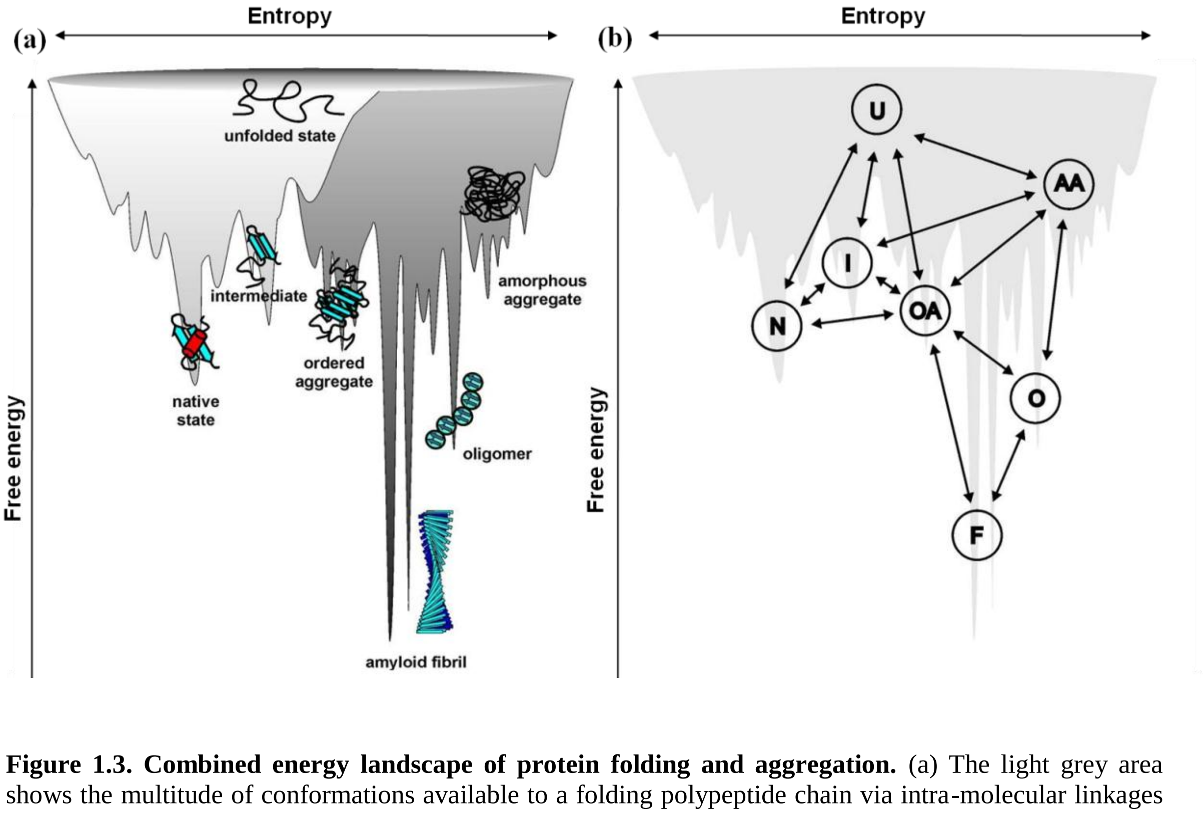

(a)
Entropy
Free energy
unfolded state
intermediate
native state
ordered aggregate
amorphous aggregate
oligomer
amyloid fibril
(b)
Entropy
Free energy
U
AA
I
N
OA
O
F

Figure 1.3. Combined energy landscape of protein folding and aggregation. (a) The light grey area shows the multitude of conformations available to a folding polypeptide chain via intra-molecular linkages on the way to the native state and the dark grey area the multitude of conformations available via inter-molecular linkages during aggregate formation. Cytotoxic oligomers may occur as off-pathway intermediates during the formation of fibrils [17]. (b) Proposed pathways that link the conformational states depicted in (a). Adapted from [35].

Molecular chaperones, for example, are helper molecules which are able to bind to regions of client proteins that are prone to aggregation such as hydrophobic stretches thereby promoting proper folding and preventing misfolding [36]. The unfolded protein response (UPR) is activated in response to increasing amounts of unfolded or misfolded proteins in the endoplasmic reticulum (ER) and involves the upregulation of multiple quality control systems [37]. ER-associated degradation (ERAD) ensures the recognition and targeting of misfolded proteins in the ER and their export and degradation in the cytoplasm [38]. The two major pathways that lead to the degradation of cytosolic and misfolded proteins are the ubiquitin-proteasome system (UPS) and the autophagy-lysosome pathway (ALP) [39]. The UPS is a multi-enzyme degradation system and involves the conjugation of ubiquitin molecules to misfolded or aberrant proteins to target them for degradation [40]. The ALP leads to degradation of cytoplasmic components inside autophagolysosomes. Macroautophagy, microautophagy and chaperone-mediated autophagy are the three distinct pathways via which the substrates are delivered [39]. All of these protein quality control systems play an essential role in ensuring cell viability and their failure can lead to increased protein misfolding and disease. Misfolding diseases ultimately result from the inability of a protein to attain or to remain in its functional three-dimensional state and include conditions where a specific protein is unable to fold correctly, where the native state of a protein is destabilised, where a protein folds into an aberrant native state that is not trafficked correctly or where intermolecular linkages and thus protein aggregates are formed [22].

1.1.5 The conformational diseases

The conformational diseases are a class of disorders caused by protein self-association and subsequent tissue deposition [41]. Various prevalent neurodegenerative disorders such as Alzheimer’s disease and Parkinson’s disease, but also conditions that do not affect the nervous system such as Type II diabetes and α_1-antitrypsin (α_1AT) deficiency have been associated with protein aggregation (Table 1.1). Often, genetic predispositions are linked to these diseases, for example certain mutations in the gene encoding α_1AT have been associated with α_1AT deficiency [42], in the gene encoding α-synuclein with Parkinson’s disease [43-45] and in the genes encoding the amyloid precursor protein

Table 1.1. Examples of conformational diseases and the proteins associated.

Aggregate type	Protein	Predominant native secondary structure	References
Fibrils			
Alzheimer's disease	Amyloid β (Aβ) peptide derived from APP	random coil	[46, 47]
Dialysis-related amyloidosis	β_2-Microglobulin	β-sheet	[48]
Huntington's disease	Huntingtin (htt)	α-helical (full length) or random coil (exon 1)	[49, 50]
Machado-Joseph disease	Ataxin-3	α-helical, β-sheet and random coil	[51, 52]
Parkinson's disease	α-Synuclein	random coil	[53]
Transthyretin Amyloidosis	Transthyretin (TTR)	β-sheet	[54]
Type 2 diabetes mellitus	Islet amyloid polypeptide (IAPP, amylin)	random coil	[55-57]
Polymers			
α_1-Antichymotrypsin (α_1ACT) deficiency	α_1ACT	β-sheet and α-helical	[58, 59]
α_1AT deficiency	α_1AT	β-sheet and α-helical	[42, 60, 61]
Antithrombin deficiency	Antithrombin	β-sheet and α-helical	[62, 63]
Familial encephalopathy with neuroserpin inclusion bodies	Neuroserpin	β-sheet and α-helical	[64, 65]

(APP) or the presenilins-1 and -2 (PS1 and PS2), which have a role in APP processing, with Alzheimer's disease [66, 67].

Many proteins can form amorphous aggregates which lack long-range order and are often observed in the form of inclusion bodies when working with over-expressed recombinant proteins. The disease states discussed above, however, are associated with the formation of ordered aggregates. Extensive *in vitro* studies of aggregates formed from pathological proteins suggest that two classes of ordered aggregates can be formed. These are classified as fibrillar and polymeric aggregates (Figure 1.4) [68]. Disease-associated proteins commonly form fibrils which have a highly stable, unbranched, β-sheet rich structure (Figure 1.4a). In most cases, fibrils are composed of several protofibrils which twist around each other to form the major fibril [68]. Even though fibrillar aggregates are characterised by a significant β-sheet content the proteins that associate to form these aggregates often have no β-sheets in their native conformation (Table 1.1). On the other hand, polymerisation is an aggregation mechanism which does usually not involve a significant conformational rearrangement of self-associating units (Figure 1.4b). Protein polymerisation forms the molecular basis of a group of disorders termed serpinopathies which are associated with several devastating disease states such as emphysema, severe liver disease, dementia, angioedema and thrombosis [41]. The self-associating units that result in the serpinopathies are a group of proteins termed serpins which are the main focus of this thesis.

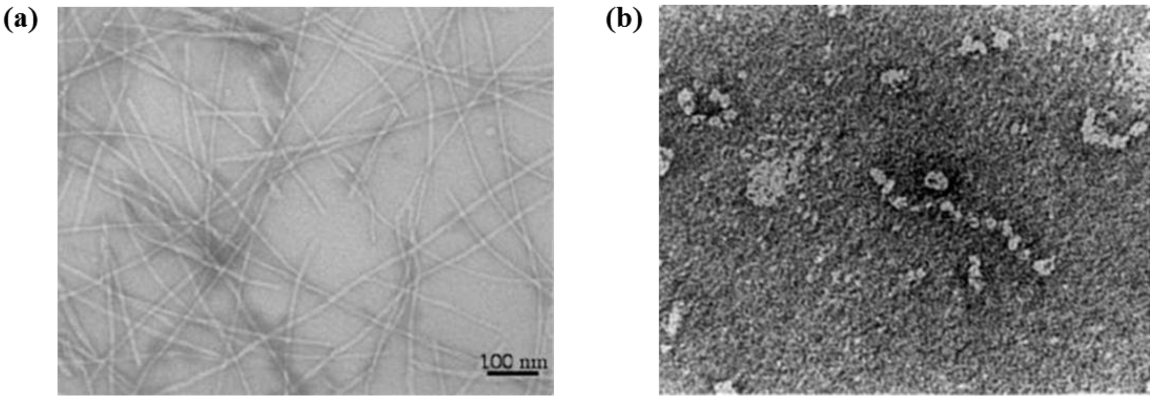


Figure 1.4. Electron micrographs of ordered aggregates. (a) Full-length IAPP fibrils with linear unbranched morphology. Adapted from [69]. (b) α_1AT polymers with “beads on a string” morphology. Adapted from [70].

1.2 The serpins

The serpins are a superfamily of proteins with diverse functions that all share a similar structure. More than 3000 serpins have been identified so far most of which are proteinase inhibitors. Serpins have been shown to play a role in every taxonomic phylum and typically have a size of 350 to 400 amino acid residues and a molecular weight of 40 to 50 kDa [71, 72]. They were first described in 1980 when Hunt and Dayhoff proposed to group ovalbumin, α_1AT and antithrombin into a superfamily based on the homology of their carboxyl-terminal regions [73]. This superfamily was then named the serpins as most members initially identified were serine proteinase inhibitors such as the plasma serpins α_1AT, α_1ACT and antithrombin [74].

The acronym “serpin" is deceiving as this is not a superfamily solely composed of serine proteinase inhibitors. Some serpins are cross-class inhibitors targeting cysteine proteinases as for example the squamous cell carcinoma antigen 1 (SCCA1) [75] and the Myeloid and Erythroid Nuclear Termination stage-specific protein (MENT) [76] and some serpins are dual-class inhibitors which are able to inhibit both serine and cysteine proteinases such as the serpins SRP-2 [77], endopin 2 [78], SQN-5 [79] and the cytokine response modifier A (CrmA) [80]. Additionally, there are serpins without any known inhibitory function including the storage protein ovalbumin [81], the tumour suppressor maspin [82], the collagen binding chaperone heat shock protein 47 (Hsp47) [83] and the hormone transporters thyroxine binding globulin (TBG) and cortisol binding globulin (CBG) [84].

1.2.1 Serpin nomenclature

Based on their phylogenetic relationships the serpins were recently categorised into 16 clades denoted A to P (Figure 1.5). This study renamed the serpins according to their clade and their membership within the clade which, for example, led to the alternative name SERPINA1 for α_1AT as it was denoted the first member of clade A. Ten serpins could not be grouped into any other clade and remained unclassified “orphans” [85].

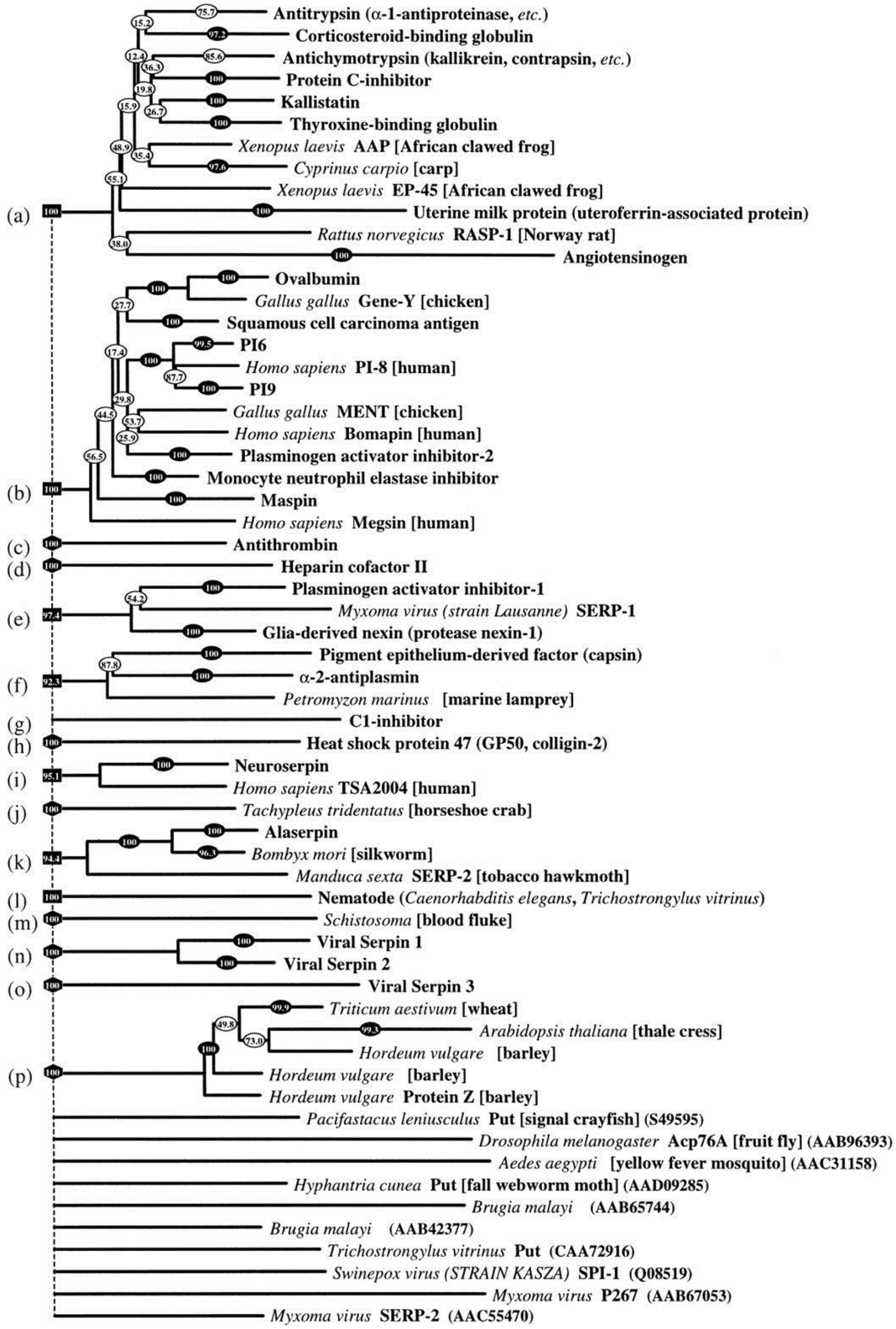

Antitrypsin (α-1-antiproteinase, etc.)
Corticosteroid-binding globulin
Antichymotrypsin (kallikrein, contrapsin, etc.)
Protein C-inhibitor
Kallistatin
Thyroxine-binding globulin
Xenopus laevis AAP [African clawed frog]
Cyprinus carpio [carp]
Xenopus laevis EP-45 [African clawed frog]
(a)
Uterine milk protein (uteroferrin-associated protein)
Rattus norvegicus RASP-1 [Norway rat]
Angiotensinogen
Ovalbumin
Gallus gallus Gene-Y [chicken]
Squamous cell carcinoma antigen
PI6
Homo sapiens PI-8 [human]
PI9
Gallus gallus MENT [chicken]
Homo sapiens Bomapin [human]
Plasminogen activator inhibitor-2
Monocyte neutrophil elastase inhibitor
(b)
Maspin
Homo sapiens Megsin [human]
(c)
Antithrombin
(d)
Heparin cofactor II
Plasminogen activator inhibitor-1
Myxoma virus (strain Lausanne) SERP-1
(e)
Glia-derived nexin (protease nexin-1)
Pigment epithelium-derived factor (capsin)
α-2-antiplasmin
(f)
Petromyzon marinus [marine lamprey]
(g)
C1-inhibitor
(h)
Heat shock protein 47 (GP50, colligin-2)
Neuroserpin
(i)
Homo sapiens TSA2004 [human]
(j)
Tachypleus tridentatus [horseshoe crab]
Alaserpin
Bombyx mori [silkworm]
(k)
Manduca sexta SERP-2 [tobacco hawkmoth]
(l)
Nematode (Caenorhabditis elegans, Trichostrongylus vitrinus)
(m)
Schistosoma [blood fluke]
Viral Serpin 1
(n)
Viral Serpin 2
(o)
Viral Serpin 3
Triticum aestivum [wheat]
Arabidopsis thaliana [thale cress]
Hordeum vulgare [barley]
(p)
Hordeum vulgare [barley]
Hordeum vulgare Protein Z [barley]
Pacifastacus leniusculus Put [signal crayfish] (S49595)
Drosophila melanogaster Acp76A [fruit fly] (AAB96393)
Aedes aegypti [yellow fever mosquito] (AAC31158)
Hyphantria cunea Put [fall webworm moth] (AAD09285)
Brugia malayi (AAB65744)
Brugia malayi (AAB42377)
Trichostrongylus vitrinus Put (CAA72916)
Swinepox virus (STRAIN KASZA) SPI-1 (Q08519)
Myxoma virus P267 (AAB67053)
Myxoma virus SERP-2 (AAC55470)

Figure 1.5. Large-scale phylogenetic tree of the serpin superfamily. The overall relationship between serpins was determined from sequence comparison as described in [85]. The ten orphans are shown at the bottom of the phylogenetic tree. Adapted from [85].

The human serpins are members of clade A to I and most of them are inhibitory (Table 1.2). Like all serpins in higher eukaryotes, the human serpins can broadly be separated into intracellular serpins or so-called ov-serpins and extracellular serpins [85]. The name ov-serpins was coined based on the sequence and structural homology of these serpins to ovalbumin [86]. The ov-serpins are ancestral to the extracellular serpins and constitute clade B [85]. In humans, they play essential roles in several physiological processes including the inhibition of apoptosis (e.g. PAI-2) and the prevention of metastasis (e.g. maspin) [82, 87].

The remaining eight clades of human serpins (A, C-I) comprise extracellular serpins with clade A being the largest [85]. They are involved in the inflammatory response (e.g. α_1AT and α_1ACT), hormone transport (e.g. thyroxine binding globulin and cortisol binding globulin) and the regulation of fibrinolysis (PAI-1), the coagulation cascade (e.g. antithrombin) and the complement system (e.g. C1 inhibitor) [84, 88, 89]. There is also evidence that certain intracellular serpins assist in extracellular roles including apoptosis and cytoprotection (e.g. PI6 and PI9) [90, 91]. It has been proposed, based on their phylogenetic relationship, that neuroserpin or antithrombin may be a linker between intracellular and extracellular serpins [85].

1.2.2 Serpin structure

Even though serpins may have very different physiological roles and amino acid sequence identities as low as 17 % they all share a similar structure consisting of three β-sheets (A-C) and eight to nine α-helices (A-I) (Figure 1.6) [92]. The largest of the three β-sheets is the central β-sheet A which is composed of five strands. β-sheet B and C lie above β-sheet A and consist of six and four strands, respectively. Whereas the strands of β-sheet B and C run anti-parallel to each other, β-sheet A is composed of parallel and anti-parallel strands. The helices are arranged around this β-sheet scaffold with most of them being located at the back of the molecule (relative to β-sheet A). Only one helix is located at the front of the serpin: helix F. Helix F runs across β-sheet A and its role has been implicated in serpin folding, misfolding and proteinase inhibition [93, 94].

Table 1.2. Clade classification of the human serpins.

Serpin	Synonyms	Target proteinases or function
SERPINA1	α_1-Antitrypsin (α_1AT), α_1PI	Inhibits neutrophil elastase
SERPINA2	α_1AT-related protein	Probably pseudogene
SERPINA3	α_1-antichymotrypsin (α_1ACT)	Inhibits cathepsin G, chymase and prostate specific antigen
SERPINA4	Kallistatin, kallikrein inhibitor, PI4	Inhibits tissue kallikrein
SERPINA5	Protein C inhibitor (PCI), PAI-3	Inhibits active protein C, uPA, plasma kallikrein and acroin
SERPINA6	Corticosteroid binding globulin (CBG)	Non-inhibitory, cortisol binding
SERPINA7	Thyroxine binding globulin (TBG)	Non-inhibitory, thyroxine binding
SERPINA8	Angiotensinogen (AGT)	Non-inhibitory, N-terminal cleavage by renin releases angiotensin I
SERPINA9	Centerin	Maintenance of naive B cells
SERPINA10	Protein Z-dependent proteinase inhibitor (ZPI)	Inhibits activated factors Z, Xia and Xa
SERPINA11	XP_170754.3	Not characterised
SERPINA12	Vaspin	Insulin-sensitising adipocytokine
SERPINA13	XM_370772	Not characterised
SERPINB1	Monocyte neutrophil elastase inhibitor (MNEI), leukocyte elastase inhibitor (LEI), PI2, EI, ELANH2	Inhibits neutrophil elastase, cathepsin G and proteinase 3
SERPINB2	PAI-2, placental PAI	Inhibits uPA
SERPINB3	Squamous cell carcinoma antigen-1 (SCCA1)	Inhibits cathepsins K, L, S and V
SERPINB4	Squamous cell carcinoma antigen-2 (SCCA2), leupin	Inhibits cathepsins G and chymase
SERPINB5	Maspin, PI5	Non-inhibitory, metastasis suppressor
SERPINB6	Placental thrombin inhibitor (PTI), PI6, cytoplasmic antiprotease (CAP)	Inhibits cathepsin G, thrombin, plasmin and chymotrypsin
SERPINB7	Megsin	Megakaryocyte maturation
SERPINB8	Cytoplasmic antiproteinase 2 (CAP2), PI8	Inhibits furin, thrombin and subtilisin A
SERPINB9	Cytoplasmic antiproteinase 3 (CAP3), PI9	Inhibits granzyme B, caspase 1 and subtilisin A
SERPINB10	Bomapin, PI10	Inhibits thrombin and trypsin
SERPINB11	Epipin	Not characterised
SERPINB12	Yukopin	Inhibits trypsin and plasmin
SERPINB13	Headpin, hurpin, PI13	Inhibits cathepsins L and K
SERPINBP1	PI8-like1	Not characterised, probably pseudogene
SERPINC1	Antithrombin (AT), Antithrombin III (ATIII)	Inhibits thrombin, factors Xa and IXa
SERPIND1	Heparin cofactor II (HCII), leuserpin 2	Inhibits thrombin
SERPINE1	PAI-1, endothelial PAI	Inhibits thrombin, uPA, tPA, plasmin and activated protein C
SERPINE2	Protease nexin 1 (PN1), glia derived nexin (GDN), PI7	Inhibits uPA, tPA, thrombin, acrosin and factor XIa
SERPINE3	Hs.512272	Not characterised
SERPINF1	Pigment epithelium derived factor (PEDF), EPC-1	Non-inhibitory, potent anti-angiogenic factor
SERPINF2	α_2-Antiplasmin (α_2AP), A2AP	Inhibits plasmin
SERPING1	C1 inhibitor (C1 INH)	Inhibits C1r, C1s and plasma kallikrein
SERPINH1	Hsp47, collagen-binding protein 1 (CBP1), colligin 1	Non-inhibitory, chaperone for collagens
SERPINH2	collagen binding protein 2 (CBP2), colligin 2	Rheumatoid arthritis-related antigen
SERPINI1	Neuroserpin, PI12	Inhibits tPA, uPA and plasmin
SERPINI2	Myoepithelium-derived serine proteinase inhibitor, MEPI, pancpin, PI14	Suppression of cancer metastasis

Abbreviations: PI= proteinase inhibitor, uPA = urokinase-type plasminogen activator, tPA = tissue-type plasminogen activator, PAI = plasminogen activator inhibitor. Adapted from [81, 95].

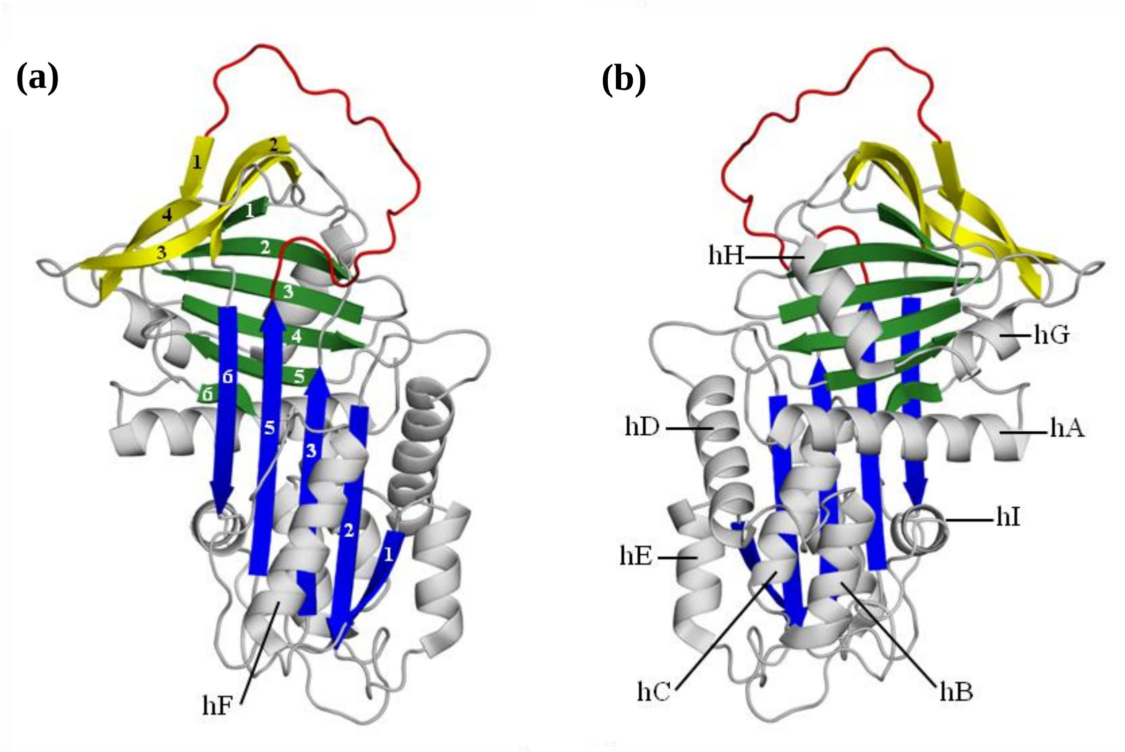


Figure 1.6. Comparison of the front view (a) and back view (b) of the serpin structure. Ribbon diagram of the archetypal serpin α_1AT (Protein Data Bank ID: 1QLP) with β-sheet A in blue, β-sheet B in green, β-sheet C in yellow, the RCL in red and helices (h) in grey.

In the native state β-sheet A supports an exposed reactive centre loop (RCL) that extends from strand five of β-sheet A (s5A) to strand one of β-sheet C (s1C). That the RCL can adopt a helical conformation became evident from the crystal structure of ovalbumin, the first native serpin structure to be solved [96]. A helical RCL conformation has also been observed for an engineered variant of α_1ACT [97], however, this is not a general feature of the serpins. The RCL of α_1AT, for example, is in a canonical β-strand conformation [98, 99] and the crystal structures of heterodimeric antithrombin suggest that the RCL exists in an extended conformation with its N-terminal part inserted between strand three and five of β-sheet A [100-102].

1.2.3 Serpin mechanism of proteinase inhibition

In inhibitory serpins the RCL contains the cleavage site for the target proteinase and it became already evident from the first serpin structure to be crystallised, cleaved α_1AT, that RCL cleavage must be accompanied by a significant conformational rearrangement [103]. This crystal structure revealed that upon cleavage the N-terminal part of the RCL inserts into β-sheet A separating the cleaved residues by approximately 67 Å and turning β-sheet A anti-parallel (Figure 1.7a). Accordingly, the strands of β-sheet A are numbered one to six with the RCL becoming strand four of β-sheet A (s4A) upon insertion (Figure 1.6a).

The crystal structure of cleaved α_1AT together with subsequent extensive biophysical and biochemical studies led to the conclusion that a serpin's inhibitory mechanism involves insertion of the RCL into β-sheet A and translocation of the covalently attached proteinase to the opposite pole of the serpin [104-111]. This was confirmed by Huntington *et al.* in 2000 when the first crystal structure of a serpin-proteinase complex was published (Figure 1.7b) [112]. In this complex between α_1AT and trypsin the serpin showed full insertion of the RCL N-terminal to the cleavage site into β-sheet A between strand 3 and strand 5 similar to the structure of cleaved α_1AT. The proteinase, which remained attached to the RCL during loop insertion, had been moved approximately 71 Å and ended up crushed against the base of the serpin. This translocation process resulted in a loss of approximately 37 % of the native structure of the proteinase therefore suggesting inhibition occurs via deformation.

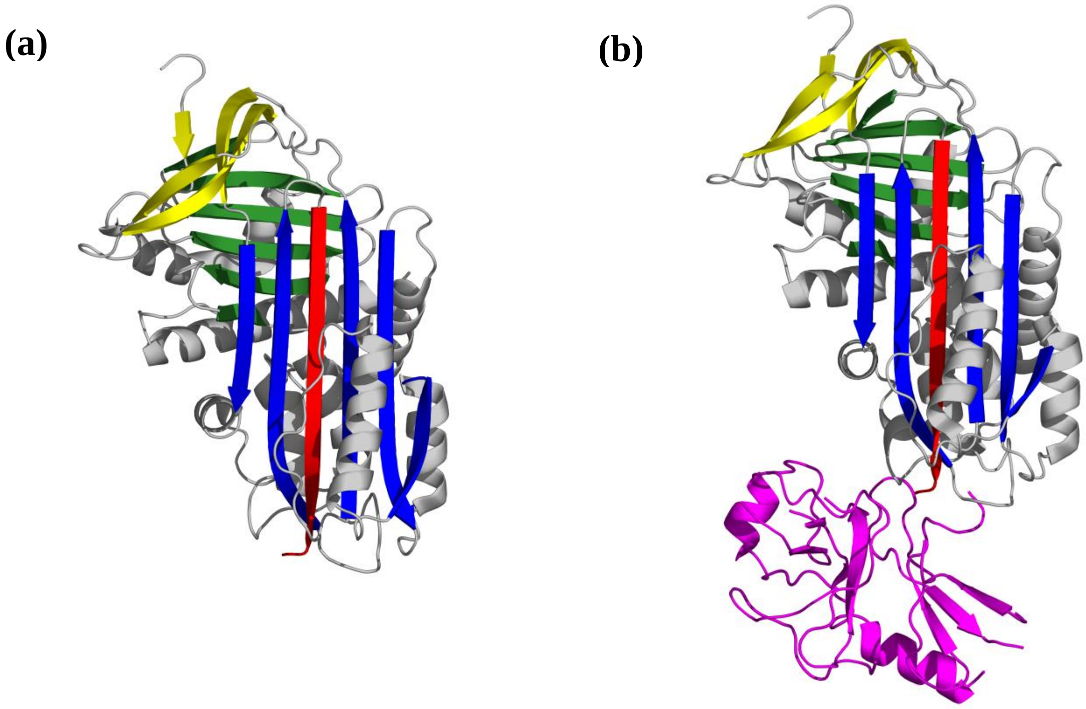


Figure 1.7. Ribbon diagrams of RCL cleaved α_1AT. (a) Crystal structure of cleaved α_1AT (Protein Data Bank ID: 7API). (b) Crystal structure of the covalent complex of α_1AT and trypsin (Protein Data Bank ID: 1EZX)**.** Both structures show full insertion of the cleaved RCL (red) into β-sheet A (blue) of α_1AT. Trypsin is shown in magenta.

The product of a serpin’s inhibitory process is hence a covalent complex between the cleaved serpin and the catalytically inactivated proteinase. Dissociation of the serpin-proteinase complex occurs over time as the trap is kinetic [113, 114], however, the complex is usually cleared from the system prior to breakdown [115-117]. A serpin’s mechanism of proteinase inhibition therefore differs significantly from the so-called “standard mechanism” in two points: Firstly, the serpin does not regain its inhibitory function after completion of the inhibitory process and secondly, the covalent complex formed between the serpin and the proteinase is essentially irreversible [118]. Based on their extraordinary mechanism of inhibition, serpins have therefore also been described as “molecular mousetraps” [119] and “suicide inhibitors” [120].

The initial steps of proteinase inhibition by a serpin are equivalent to the initial steps of substrate hydrolysis by the target proteinase (Figure 1.8). The RCL, which protrudes from the body of the serpin, essentially acts as bait. It contains the scissile bond denoted P1-P1’ according to the nomenclature of Schechter and Berger [121] which is recognised as substrate by the target proteinase. There is, however, evidence that not only the P1-P1’residues mediate this specificity but also other RCL residues are involved [122-125]. The inhibitory process is initiated by the formation of a non-covalent Michaelis complex between serpin and proteinase followed by the formation of a covalent serpin-acyl-proteinase complex which involves an ester bond formation between the P1 residue and the active site of the proteinase and cleavage of the scissile bond.

From this point of the reaction onwards there are two competing pathways: the inhibition pathway and the substrate pathway. If the inhibitory process is successful, cleavage of the scissile bond leads to RCL insertion into β-sheet A and proteinase translocation to the opposite pole of the serpin [81, 126]. It has been proposed that catalytic deacylation is prevented by active site distortion of the proteinase as a result of the translocation process: the acyl intermediate is trapped [112]. The acyl linkage is, however, susceptible to hydrolytic deacylation during the translocation process and completion of the proteolysis reaction can occur leading to the release of the active protease and leaving behind the cleaved serpin. Hence, the outcome of the inhibitory process is determined by the competing rates of RCL insertion and deacylation [81, 126] and mutations that impede with efficient RCL insertion usually result in increased substrate behaviour of the serpin [127, 128].

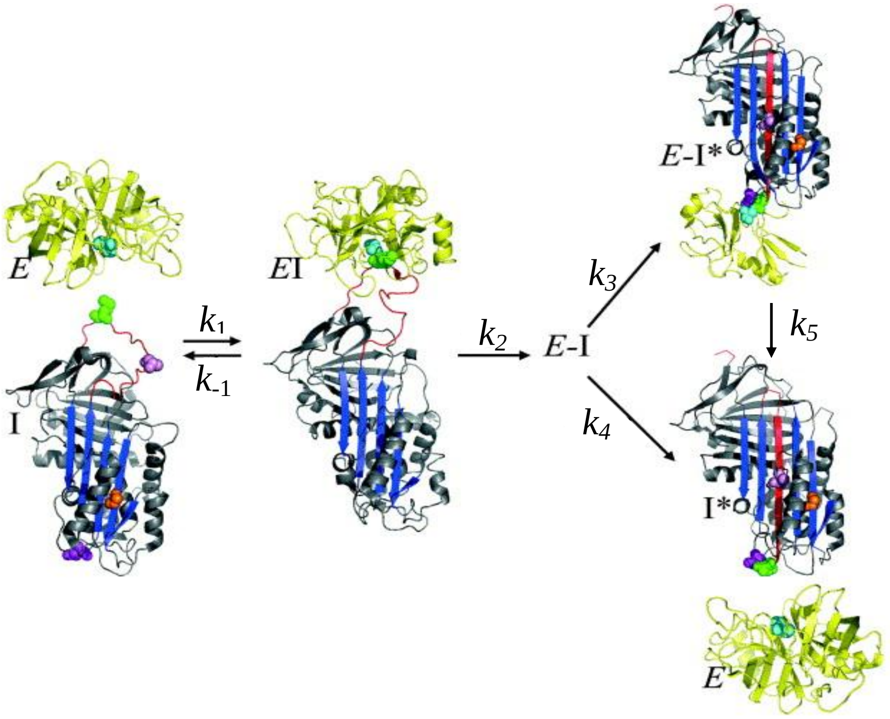


Figure 1.8. The branched pathway mechanisms of serpin proteinase inhibition. Only the essential intermediates are shown with the structure of the proteinase depicted in yellow (*E*) and the serpin in grey with β-sheet A in blue and the RCL in red (I). The formation of the non-covalent Michaelis complex (*E*I) is described by the forward rate constant k_1 and the back rate constant k_{-1}. The *E*I complex progresses to the serpin-acyl-proteinase intermediate (*E*-I) which is thought to represent the native serpin conformation but with the P1-P1' bond cleaved and the active site serine (cyan) of the proteinase linked with the P1 residue (green) of the serpin via an acyl linkage. The overall rate constant for this reaction is k_2. Insertion of the RCL into β-sheet A and translocation of the proteinase leads to kinetic trapping of the serpin-acyl-proteinase intermediate and hence the stable covalent complex between the proteinase and the cleaved serpin (I*) denoted *E*-I*. k_3 is the rate constant for this translocation reaction. The competing substrate-like deacylation reaction, which results in cleaved serpin and free enzymatically active proteinase, is described by the rate constant k_4. The rate constant for the slow breakdown of the *E*-I* complex into cleaved serpin and free proteinase is k_5. Adapted from [129].

1.2.4 Serpin metastability and its role in conformational change

The conformational change of a serpin that occurs during proteinase inhibition is thermodynamically driven as several studies have shown that RCL insertion into β-sheet A results in a significant increase in stability of the serpin [103, 106, 130-132]. This increase in stability upon RCL cleavage was first shown for the serpins α_1AT and antithrombin by Carrell and Owen in 1985 [133]. Whereas both serpins precipitated out of solution when incubated at 60 °C no precipitation was observed upon RCL cleavage even if the serpins were incubated at 80 °C for two hours. The authors therefore referred to the metatstable native as the "stressed" (S) state which undergoes a structural rearrangement to a more ordered "relaxed" (R) state upon RCL cleavage. Even the non-inhibitory serpin ovalbumin has been shown to be able to undergo the S to R transition with concomitant increase in stability if the critical P14 hinge residue is changed from the bulkier arginine to a threonine residue [134]. Hence, together with a high degree of structural flexibility, metastability of the native state is crucial for a serpin's inhibitory activity but renders the molecule particularly sensitive to mutations which lead to the adoption of dysfunctional folds with increased stability.

Extensive work by Yu and colleagues was able to link metastability of the native state to energetically unfavourable interactions distributed throughout the whole serpin molecule such as over-packed side chains [135], buried polar groups [136], surface exposed hydrophobic pockets [137] and cavities in the hydrophobic core [138] which impose strain. This strain is released upon RCL cleavage or the transition into the alternative high stability states discussed in section 1.2.5. Naturally occurring mutations, mutagenic studies, phylogenetic analysis and molecular modelling have led to the identification of regions within the serpin molecule that modulate serpin conformational change (Figure 1.9) [71, 85, 139, 140]. Several of these regions play an important role

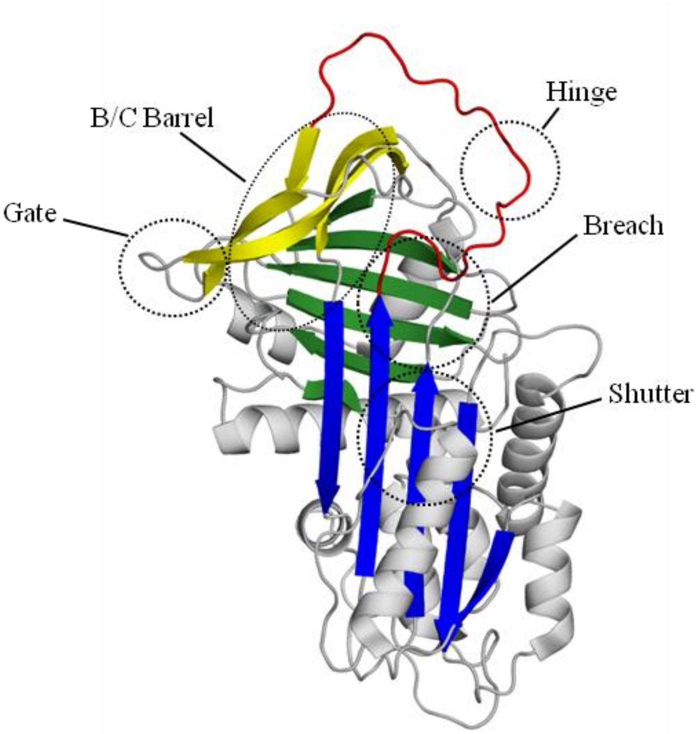


Figure 1.9. Regions of the serpin molecule that modulate conformational change. Ribbon diagram of α_1AT (Protein Data Bank ID: 1QLP) indicating the regions that are involved in the molecular switches with β-sheet A shown in blue, β-sheet B in green, β-sheet C in yellow and the RCL in red.

during the inhibitory process and hence need to display considerable structural flexibility which makes them vulnerable to mutations that result in misfolding. Such mobile structural elements include the proximal hinge region, the breach region, the shutter region and the gate region whereas the so-called B/C barrel remains rather static during proteinase inhibition but its role has been implicated in the transition to latency and serpin polymers [71, 85, 140].

The proximal hinge region includes a portion of the RCL which is required to display considerable mobility in order to allow efficient RCL insertion during the inhibitory process. It is comprised of the residues P15 to P9 which are located N-terminal to the scissile bond [141]. Inhibitory serpins are usually characterised by a consensus sequence of residues with short side chains in the hinge region to permit the flexibility required for the conformational change [141]. Mutations in the hinge region often lead to increased substrate behaviour during the inhibitory process [85].

The breach region is the area at the top of β-sheet A where the RCL first inserts after cleavage by the proteinase. It consists of s3A, s5A, s2B, s3B and s4B [140]. The breach region is characterised by a high degree of sequence conservation across the serpin superfamily which suggests that the hydrogen bond network in this area is particularly important for the inhibitory process [85]. Glu342, which is mutated to a lysine in the most common pathological variant of α_1AT, Z α_1AT, is one of these highly conserved residues. Glu342 forms a hydrogen bond to Thr203 and a salt bridge to Lys290 which are also both conserved residues [85]. Disruption of these interactions by the Z mutation leads to increased polymerisation of α_1AT [142, 143] and impaired proteinase inhibition [144, 145]. Compromised inhibitory activity was also observed for PAI-1 upon mutation of Trp175, which is equivalent to Trp194 in α_1AT, another highly conserved residue in the breach region [146].

The shutter region is the area in the centre of the molecule which, together with the breach region, facilitates β-sheet A opening and RCL insertion during the inhibitory process. It includes parts of s3A, s5A, s6B and the top of helix B [140]. Like the breach region, the shutter region is characterised by many highly conserved residues [85]. Mutations in this region have been associated with misfolding and disease of several naturally occurring serpin variants including Mmalton α_1AT (Phe52 deleted) [147], Siiyama α_1AT (Ser53Phe) [70], the Leu55Pro mutant of α_1ACT [148] and the Pro80Thr mutant of antithrombin [149].

The gate region, which is also characterised by a large number of highly conserved residues, includes s3C and s4C [85]. This region is involved in the transition to latency [150, 151] and it has been proposed that the high number of conserved residues might play a role in maintaining the native state [85].

Also the B/C barrel, the area between β-sheet B and C including s1B, s3B, s1C and s2C, has been implicated in serpin misfolding as mutation of residues in this region in α_1AT, α_1ACT and C1 inhibitor result in the formation of latent or polymeric species [71, 152].

1.2.5 Alternative high stability serpin states

Besides the cleaved serpin conformation, there is a range of alternative high stability states that can be adopted by certain serpins when exposed to stressors such as elevated temperatures, low denaturant concentrations and acidic pH values or through the introduction of specific mutations [139, 142, 153-155]. Most of these alternative high stability states involve at least partial insertion of the RCL into the serpin either intra-molecular forming the latent or the delta conformations or inter-molecular leading to serpin polymers.

1.2.5.1 Uncleaved monomeric high stability states

Latency represents the most stable uncleaved monomeric serpin conformation and has first been described for PAI-1 [156]. Incubation of PAI-1 in the absence of its cofactor vitronectin leads to spontaneous conversion to the inactive latent state [156] and latency might therefore play a physiological role in controlling the inhibitory activity of this serpin [157]. The structural basis of latency became evident in 1992 when a crystal structure of latent PAI-1 was solved [150]. This structure showed full insertion of the intact RCL into β-sheet A between s3A and s5A and dislodgement of s1C in order to accommodate the relocation of the RCL (Figure 1.10a). The self-inserted RCL forms new interactions in β-sheet A resulting in an increase in stability in comparison to the

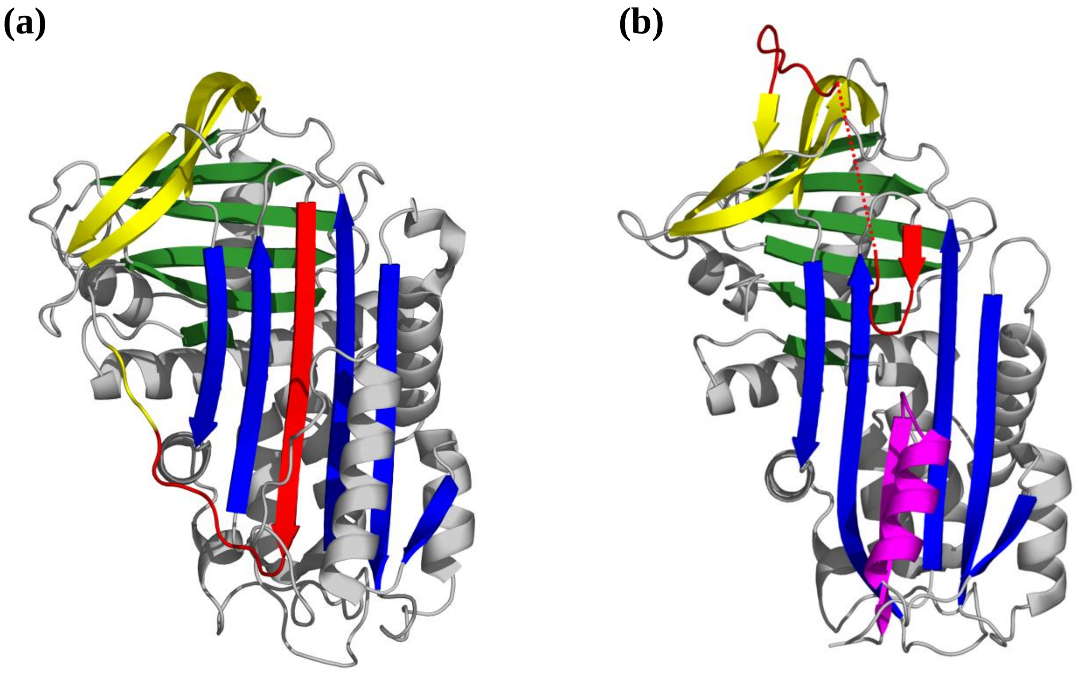


Figure 1.10. Alternative high stability monomeric serpin conformations. (a) Crystal structure of PAI-1 in the latent conformation (Protein Data Bank ID: 1DVN). The RCL (red) is inserted into β-sheet A (blue) and the first strand of β-sheet C (yellow) is dislodged. (b) Crystal structure of α_1ACT in the delta conformation (Protein Data Bank ID: 1QMN). The RCL (red) is partially inserted into the top of β-sheet A (blue) and parts of helix F (magenta) have rearranged to insert into the bottom of β-sheet A. β-Sheet B is shown in green in both structures.

native state. The latent conformation is less stable than the cleaved conformation which might be due to destabilisation of β-sheet C [155, 158]. Besides PAI-1 other serpins have also been proposed to be capable of undergoing the transition to latency including α_1ACT [159], α_1AT [152, 155] and antithrombin [100, 160, 161].

An alternative high stability inactive monomeric serpin conformation is the delta (δ) conformation which was first described for the naturally occurring Leu55Pro variant of α_1ACT. [59]. Its crystal structure showed partial insertion of the RCL into the top of β-sheet A between s3A and s5A (Figure 1.10b). Additionally, the last turn of helix F and the loop linking it to s3A had rearranged to insert into the lower part of β-sheet A. It has been proposed that the delta conformation might represent a structure similar to the intermediates adopted during complex formation and polymerisation [59]. A more recent study, however, which showed that wild-type α_1ACT can also adopt the delta conformation under physiological conditions, suggests that the delta conformation is rather an end product in the misfolding pathway of α_1ACT [162]. This study further proposed that in previous work stable monomeric alternative states might have been mistaken for latent whereas they may have actually been in the delta conformation.

1.2.5.2 Polymeric high stability states

Serpin polymerisation is the most clinically relevant adoption of high stability states as it forms the basis of most serpinopathies (serpin related disorders). Tissue deposition of serpin polymers can lead to gain-of-toxic-function diseases and the decrease of circulating serpin to loss-of-function diseases [61, 68]. Electron micrograph images of serpin polymers show a "beads on a string" morphology (Figure 1.4b) [163, 164], however, the exact nature of the inter-molecular linkages formed between the monomeric units ("beads") to give rise to the polymer chain remains controversial. Additionally, there is evidence that the polymerisation mechanism can vary depending on the conditions provided and the serpin involved [106, 165-167].

Several polymerisation models have been proposed, many of which are based on crystallographic data. Except for the disulphide polymerisation model, which suggests that polymerisation occurs via the formation of disulphide-linked dimers which further propagate into linear polymers via surface interactions involving β-sheet A [168], all of these models propose that the RCL is involved in the formation of the inter-molecular linkages.

Cleaved α_1AT polymers were the first serpin polymers to be crystallised [169, 170]. These structures showed polymerisation can arise when residues N-terminal to the cleavage site insert into β-sheet A of the same molecule and the C-terminal portion inserts into the partially occupied β-sheet A of another molecule (Figure 1.11a).

A crystal structure of PAI-1 revealed serpin polymerisation can also occur via annealing of the RCL of a donor molecule as an edge strand to s6A of an acceptor molecule (Figure 1.11b) [171]. The physiological relevance of strand 7A polymers, however, remains questionable as the interactions formed are reversible [171] but several studies indicate *in vivo* formed serpin polymers are stable [172-174].

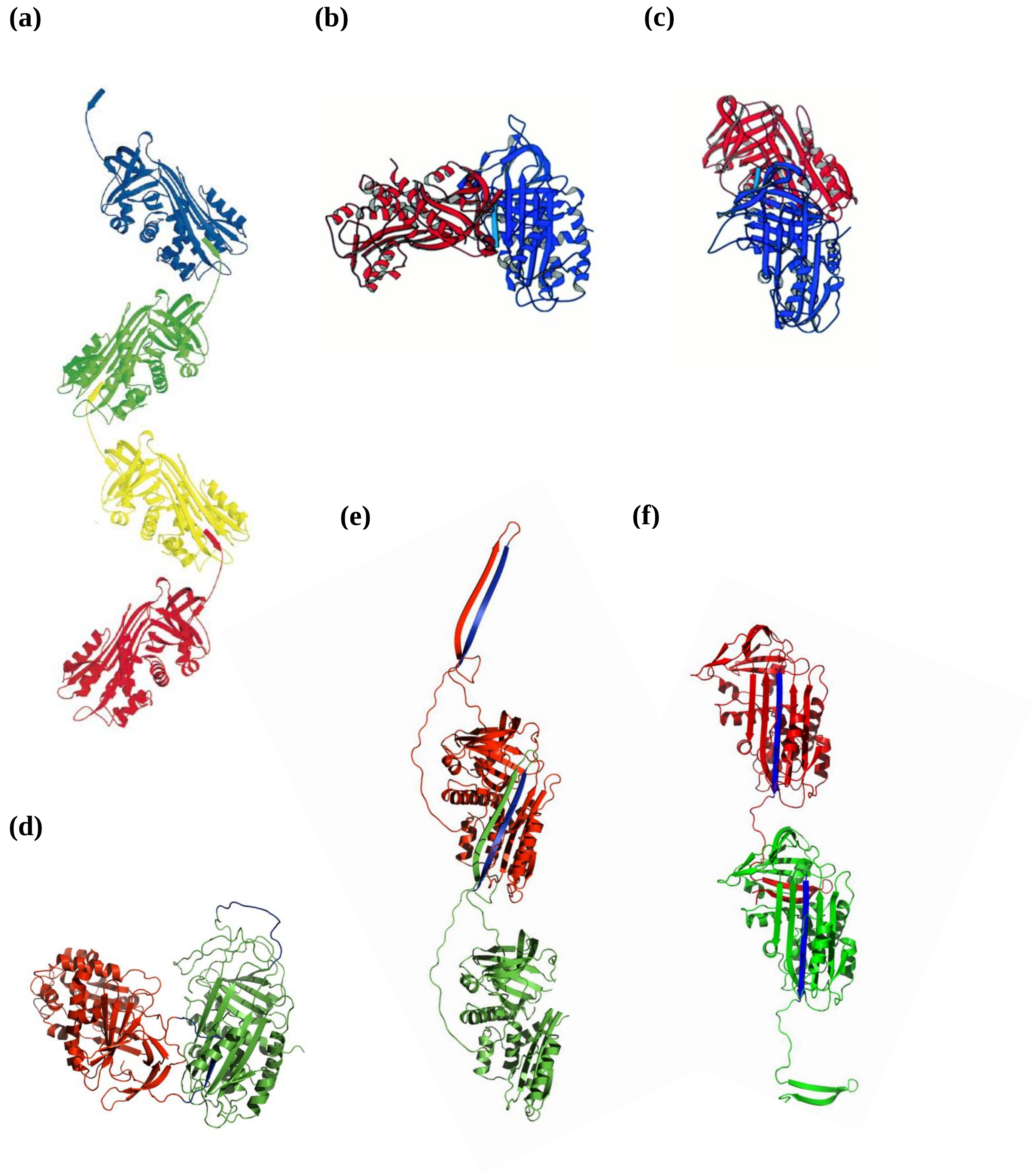


Figure 1.11. Serpin polymerisation models. (a) Cleaved polymer (Protein Data Bank ID: 18mb). Adapted from [170]. (b) Strand 7A polymer. Adapted from [175]. (c) Loop C-sheet polymer. Adapted from [175]. (d) Loop A-sheet polymer. Adapted from [176]. (e) s4A/s5A swap polymer. Adapted from [176]. (f) C-terminal swap polymer. Adapted from [176].

Loop C-sheet polymers, on the other hand, might be of physiological relevance as it has been proposed that the pathological Mmalton α_1AT variant (Phe52 deleted) polymerises via this mechanism [147]. Evidence for loop C-sheet polymerisation came from structural data of dimeric antithrombin which showed one molecule in the latent conformation with s1C displaced from the molecule to allow full RCL insertion [100, 102, 161]. Inter-molecular linkages occur via the RCL of the second molecule which anneals to s2C and fills the position of s1C left vacant (Figure 1.11c). Displacement of s1C as an initial event during polymerisation has further been supported by hydrogen/deuterium exchange [177] and disulphide-trapping [178] experiments.

Until recently, one of the most accepted models for α_1AT polymerisation has been the loop A-sheet model, which is also known as the s4A swap model. This model explains the formation of intermolecular linkages via the RCL of a donor molecule inserting between s3A and s5A of an acceptor molecule [142] (Figure 1.11d). Evidence for loop A-sheet polymerisation came from peptide annealing studies [106, 142, 179-182], fluorescence resonance energy transfer (FRET) analyses [183] and hydrogen/deuterium exchange experiments [177]. Two common pathological α_1AT variants, Z α_1AT (Glu342Lys) and Siiyama α_1AT (Ser53Phe), have been proposed to polymerise via loop A-sheet linkages [70, 142].

The loop A-sheet polymerisation model was challenged in 2008 when the crystal structure of a self-terminating antithrombin dimer was published [184]. This crystal structure showed the two molecules had not just exchanged the RCL but also the adjacent s5A. Based on this crystal structure the s4A/s5A swap polymerisation model was proposed which explains polymerisation via the dual-strand insertion of s4A and s5A into β-sheet A of the acceptor molecule (Figure 1.11e). The authors additionally conducted disulphide-trapping experiments which suggest that guanidine hydrochloride induced α_1AT polymerisation also follows this polymerisation pathway [184].

Recently, a crystal structure of a self-terminating trimer of a disulphide variant of α_1AT was published which suggests an even more excessive inter-molecular exchange during polymerisation [167]. The three α_1AT molecules were linked via a C-terminal swap including s1C, s4B and s5B (Figure 1.11f). Based on this crystal structure the C-terminal swap mechanism of serpin polymerisation was proposed. Notably, C-terminal swap polymers are recognised by the monoclonal 2C1 antibody [167] which is specific for the pathogenic polymers in hepatocellular inclusions [185], suggesting that these polymers are of physiological relevance.

1.2.6 Mechanisms of serpin folding and misfolding

The question of how serpins attain their metastable functional conformation and avoid folding into the thermodynamically more favourable states discussed in section 1.2.5 has intrigued scientists for decades.

1.2.6.1 Folding into the metastable native state

In vitro analysis of the folding pathway of a number of serpins suggests the presence of at least one intermediate ensemble on the way to the native conformation which can be populated in low denaturant concentrations [186-190]. Conclusively, serpin folding can be described by the following equation:

$$U \leftrightarrow I \leftrightarrow N$$

where U represents the unfolded ensemble, I the partially folded, transiently stable intermediate ensemble and N the metastable native fold.

Biophysical studies suggest that the intermediate ensemble contains about 80 % of the secondary structural elements of the native state [186] and is characterised by a relatively unperturbed β-sheet B, partially folded β-sheet A and β-sheet C [186-188, 191], a folded helix B [191] and a grossly disrupted helix F [93, 94, 191]. Based on some of these findings it has been proposed that the first three strands of β-sheet B together with helix G and helix H might serve as a nucleus for folding [188]. Also, hydrogen/deuterium exchange studies identified s2B and s3B as areas of local persistent structure in the intermediate ensemble [192]. Additionally, this study suggested that the folding intermediate has many of the properties of a molten globule including highly dynamic and transient secondary structural elements [192].

Folding intermediates often retard the folding process as they form non-native contacts, the disruption of which presents a kinetic barrier to the native fold and controversy exists as to whether they exist as off-pathway or on-pathway products [20, 193]. In regards to serpin folding it has been proposed that the formation of such non-native contacts is productive as they restrict the conformational space available to the folding molecule and hence prevent misfolding into more stable alternative structures [68]. Particularly non-native interactions formed by helix F could play an important role in attaining the native serpin conformation and preventing folding into higher stability states. Biophysical studies [94] together with crystallographic data [59] suggest that the top of helix F inserts between s3A and s5A in the folding intermediate which would consequently prevent the intra- or inter-molecular insertion of the RCL [71].

1.2.6.2 Misfolding into polymers

Incubation of the intermediate ensemble populated in low denaturant concentrations often leads to polymer formation which suggests that the intermediate species is prone to polymerisation [142, 143, 167, 194-196]. In addition to low denaturant concentrations, elevated temperatures [93, 94, 154, 167, 197] and altered pH values [153, 175] can lead to serpin polymerisation *in vitro*. Together with the observation that destabilising mutations in the serpin molecule often result in polymer formation [139], this suggests that disruption of the native state plays a significant role in the polymerisation process.

Spectroscopic analysis of the heat-induced polymerisation reaction of α_1AT revealed the formation of a non-native species in a fast, concentration-independent reaction. This non-native species was termed M* and is considered to be the polymeric monomeric species which is capable of the formation of inter-molecular linkages. Subsequent polymerisation was observed in a slower, concentration dependent second reaction [198]. These data suggest the following kinetic scheme of serpin polymerisation:

$$\mathrm{M} + \mathrm{M} \overset{k_1}{\leftrightarrow} \mathrm{M}^* + \mathrm{M}^* \xrightarrow{k_2} \mathrm{P}$$

where M represents the native monomeric conformation, M* the polymerogenic non-native monomeric species and P the polymeric species. The transition of M to M* (represented by k_1) is approximately tenfold faster than the transition of M* to P (represented by k_2) [198].

According to this kinetic scheme, serpin polymerisation is dependent on the concentration of M* and therefore the rate of M* formation. Mutations that either lead to destabilisation of the native fold and hence favour the formation of M* or that decrease the kinetic barrier to attain M* will result in an increased polymerisation rate.

Similar to the heat-induced polymerisation reaction of α_1AT discussed above also pH-induced polymerisation of α_1AT [153], pH-induced polymerisation of PAI-1 [175] and heat induced polymerisation of α_1ACT [197] revealed the formation of a non-native species early on in the polymerisation reaction. However, whether the different denaturation methods lead to the formation of the same polymerogenic species and consequently result in the same inter-molecular linkages is unclear.

The folding intermediate I populated in low denaturant concentrations and M* populated under heat, for example, share several similarities besides their polymerogenic nature. They both bind the fluorescent dye 4, 4'-dianilino-1, 1'-binaphthyl-5, 5'-disulfonic acid (bis-ANS) [154, 187] and the formation of both, which is accompanied by an increase in intrinsic tryptophan fluorescence [187, 198], has been associated with the disruption of β-sheet A [154, 188, 198]. Whilst α_1AT polymers formed under heat, however, bind the 2C1 antibody [167, 199] which also recognises the pathogenic polymers in hepatocellular inclusions [185], polymers formed by incubation in low denaturant concentrations do not [167, 199]. This suggests that depending on the *in vitro* polymerisation conditions α_1AT forms different types of polymers. Whether that is also the case *in vivo*, however, is not clear at this stage. Yet, only 8-10 % of individuals homozygote for the pathological Z mutation of α_1AT develop liver disease by a gain-of-toxic-function mechanism with remarkable variability in its phenotypic expression [200, 201]. Whether these differences in disease manifestation are due to different polymerisation pathways of Z α_1AT, which factors influence the choice of pathway and how the potential heterogeneity of polymer formation affects their clearance are important questions to be addressed. Answers to these questions play an essential role in the development of therapeutic strategies to prevent polymerisation of α_1AT and other serpins *in vivo* and hence disease.

1.2.7 α_1AT deficiency and the pathological Z variant of α_1AT

The most common serpin-related misfolding disease or serpinopathy is α_1AT deficiency which affects approximately 1 in 1800 live births in Northern Europe and North America [202]. α_1AT is the archetypal member of the serpin superfamily and the most abundant proteinase inhibitor found in human plasma. It is expressed in hepatocytes and secreted into circulation to control the proteolytic activity of neutrophil elastase in the lower respiratory tract. Neutrophil elastase is a serine proteinase involved in the inflammatory response and its uncontrolled activity leads to parenchymal lung destruction as it degrades several connective tissue constituents. Since α_1AT is the major inhibitor of neutrophil elastase a significant decrease of α_1AT plasma levels can lead to emphysema due to a protease-antiprotease imbalance [203-205].

Several deficiency variants with varying severities have been identified including null variants, S α_1AT (Glu264Val), I α_1AT (Arg39Cys), Siiyama α_1AT (Ser53Phe), Mmalton α_1AT (ΔF52) and Z α_1AT (Glu342Lys) [164]. Z α_1AT is the most common pathological α_1AT variant and is responsible for approximately 95 % of severe α_1AT deficiency cases [206]. The Z mutation results in an increased propensity of α_1AT to polymerise in the ER of hepatocytes [142, 207, 208] which leads to a reduction in α_1AT plasma levels to approximately 10 % of normal [164]. The significant decrease in α_1AT plasma levels manifests itself in early-onset emphysema and the presence of Z α_1AT polymers in the ER of hepatocytes can culminate in severe liver disease [164, 205].

The mechanism by which the Z mutation results in misfolding and polymerisation is not well defined. Considering that Z α_1AT mainly polymerises in the ER suggests that polymerisation occurs during folding, potentially from the folding intermediate I. That the folding intermediate of α_1AT has polymerogenic properties is supported by *in vitro* studies which showed that incubation of α_1AT under conditions that populate I (e.g. low denaturant concentration) also results in the formation of α_1AT polymers [142, 143, 167, 195]. Yet, there is evidence that Z α_1AT can also form polymers once it has attained its native state under *in vitro* conditions under which the wild-type does not polymerise [142-144]. *In vivo*, this is supported by the detection of Z α_1AT polymers in circulation [207, 209] and in the lung [172]. Whether Z α_1AT polymerisation occurs from I, from N or from both and whether potentially two different polymer types are formed remains to be elucidated.

The Z mutation is a glutamate to a lysine substitution at amino acid position 342. Glu342 lies at the top of s5A and the base of the RCL and forms a hydrogen bond to Thr203 and a salt bridge to Lys290 [85]. It has been proposed that it is the loss of the 290-342 salt bridge upon introduction of the positive charge that leads to the pathological phenotype of Z α_1AT [103, 210]. Reestablishment of this salt bridge in Z α_1AT by introducing a second mutation (Lys290Glu) resulted in restoration of wild-type-like secretion levels in transfected COS I cells [210]. A second study utilising a HEPA 1a cell system, however, suggests that the introduction of the second mutation (Lys290Glu) increases secretion levels from 17 % to only 38 % of normal [211]. Additionally, this work and another study [212] suggest that disruption of the 290-342 salt bridge from the other side by replacing Lys290 with a glutamate residue in wild-type α_1AT does not decrease secretion levels to the same extent as the Z mutation. Based on these results and additional mutational analyses the authors concluded that is mainly the introduction of a positive charge at amino acid position 342 that causes misfolding of Z α_1AT and potentially results in an altered native conformation [211]. Unfortunately, there is currently no crystal structure of Z α_1AT available and experimental data on its conformation are very limited. Near-UV circular dichroism (CD) spectra suggest structural differences between Z α_1AT and the wild-type [143] and it has been proposed that this conformational change involves opening of β-sheet A in Z α_1AT [180]. Yet, the authors of the latter study based their conclusion on the interactions of Z α_1AT with peptides whose exact binding sites are unknown at this stage [180].

Thermal denaturation experiments suggest that Z α_1AT is less stable than the wild-type and it has been proposed that the decrease in thermodynamic stability of Z α_1AT results in an increase in the equilibrium concentration of the polymer precursor M* and hence increased polymer formation [198]. This is in direct contradiction with a second study which showed that wild-type and Z α_1AT have similar transition midpoints in transverse urea gradient (TUG) gels and hence comparable stabilities [213]. This study suggests that the Z mutation leads to retardation of the folding transition from the polymerogenic intermediate ensemble I to the native state and it is this folding defect that causes increased polymer formation [213].

1.3 Aims of this study

Considering the limited and yet conflicting amount of information available on Z α_1AT, several intriguing question remain to be answered: What is the effect of the Z mutation on the structure of α_1AT? How does the Z mutation result in increased polymerisation of α_1AT and hence disease? Does it alter the kinetic or thermodynamic stability of the molecule or both? Is it mainly the loss of the 290-342 salt bridge or the positive charge at amino acid position 342 that leads to increased α_1AT polymerisation upon Z mutation?

A major impediment to studying the effect of the Z mutation on α_1AT in detail has been the lack of recombinant Z α_1AT and until recently research was limited to Z α_1AT purified from plasma of patients. In 2009, however, we succeeded in developing a protocol for expressing and purifying recombinant Z α_1AT from *Pichia pastoris* (for details see Appendix/ reference 144). This ensured a supply of Z α_1AT for analysis and allowed the molecular characterisation of the pathological mutant. Based on this work, this thesis has examined some of the contradictions in the field by:

1.) Conducting an extensive biophysical characterisation of the effect of the Z mutation on structure, thermodynamic and kinetic stability of α_1AT.

2.) Mapping the conformational change induced by the Z mutation onto the tertiary scaffold of α_1AT utilising single tryptophan mutants of Z α_1AT.

3.) Determining the effect of the side chain charge of amino acid 342 on the structure and rate of polymerisation of α_1AT.

Chapter 2

Materials and Methods

2.1 Reagents

Enzymes were obtained from either Promega or Novagen. Dimethyl sulfonic acid (DMSO) was purchased from Merck and Triton®X-100 from MP Biomedicals. Tween® 20 and sodium dodecyl sulfate (SDS) were obtained from AMRESCO. 4, 4’-dianilino-1, 1’-binaphthyl-5, 5’–disulfonic acid (bis-ANS) was purchased from Invitrogen. Phenylmethylsulphonyl fluoride (PMSF), β-mercaptoethanol (BME), bovine serum albumin (BSA), ethylenediaminetetraacetic acid (EDTA), N,N,N’,N’-Tetramethylethylenediamine (TEMED), ammonium persulfate (APS), polyethylene glycol (PEG-8000), ampicillin, kanamycin, guanidium hydrochloride (GdnHCl), guanidinium thiocyanate (GdnSCN) and urea were obtained from Sigma Aldrich. The bis-ANS, GdnSCN, GndHCl and urea solutions were prepared analytically by weight and passed through 0.22-μm membranes. The concentrations of GdnHCl and urea were determined from the refractive index of the prepared solutions [214]. Bovine α-chymotrypsin (chymotrypsin) was purchased from Sigma Aldrich, prepared analytically by weight and stored in 1 mM HCl at -80 °C. Human neutrophil elastase (HNE) was obtained from Calbiochem, prepared analytically by weight in 50 mM sodium acetate, 200 mM NaCl, pH 5.5 and stored at -80 °C. Chromogenic substrates for chymotrypsin (N-Succinyl-Ala-Ala-Pro-Phe p-nitroanilide) and HNE (N-Methoxysuccinyl-Ala-Ala-Pro-Val p-nitronailide) were purchased from Sigma Aldrich, prepared analytically by weight in DMSO and stored at -20 °C.

2.2 Microbiological techniques

2.2.1 *Escherichia coli (E. coli)*

2.2.1.1 Media and agar plates

The media used for *E. coli* growth was 2YT media (1.6 % (w/v) tryptone, 1 % (w/v) yeast extract, 0.5 % (w/v) NaCl) and SOC media (0.5 % (w/v) yeast extract, 2 % (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM dextrose (added after autoclaving)) was used for recovery after transformation. 2YT plates were prepared by supplementing 2YT media with 1.5 % (w/v) agar prior to autoclaving. Media were autoclaved at 120 °C for 20 min and dextrose was autoclaved at 120 °C for 10 min. For 2YT media or agar plates containing ampicillin (100 µg/mL) or kanamycin (25 µg/mL) the antibiotic was added after autoclaving once the media had cooled below 50 °C.

2.2.1.2 Strains

The *E. coli* strain JM107 was used for plasmid maintenance and the transformation of ligation and Pfu DNA polymerase site-directed mutagenesis products. NovaBlue Competent Cells (Novagen) are a K-12 derived strain and were used for the transformation of KOD DNA polymerase site-directed mutagenesis products.

2.2.1.3 Preparation of competent cells

Competent JM107 cells were prepared using the rubidium chloride method of Hanahan [215] with slight modifications. 10 mL of 2YT media were inoculated with a scraping of cells from a glycerol stock and grown with shaking overnight at 37 °C. 100 mL of 2YT were then inoculated with 1 mL of the overnight culture and incubated with shaking at 37 °C until the optical density at 600 nm (OD_{600}) was approximately 0.5. The cells were incubated on ice for 5 min and then collected by centrifugation (1,882 x g, 10 min, 4 °C). The supernatant was discarded and the cell pellet resuspended in 30 mL of ice-cold, sterile Transformation Buffer 1 (100 mM RbCl, 50 mM $MnCl_2$, 10 mM $CaCl_2$, 30 mM potassium acetate, 15 % (w/v) glycerol, pH 5.8). The cells were incubated on ice for 90 min and then collected by centrifugation (2,562 x g, 10 min, 4 °C). The supernatant was discarded and the cell pellet resuspended in 4 mL of ice-cold, sterile Transformation Buffer 2 (10 mM RbCl, 75 mM $CaCl_2$, 10 mM MOPS, 15 % (w/v) glycerol, pH 6.8). The cells were aliquoted into 100 µL volumes, snap frozen in a dry ice/ ethanol bath and immediately transferred to -80 °C for storage.

2.2.1.4 Transformation

1-10 µL of plasmid DNA were added to competent *E. coli* cells (100 µL of competent JM107 or 50 µL of NovaBlue Competent Cells) and incubated on ice (competent JM107 for 30 min, NovaBlue Competent Cells for 5 min). The cells were heat-shocked at 42 °C (competent JM107 for 90 s, NovaBlue Competent Cells for 30 s) and then incubated on ice for approximately 2 min. 2.5 volumes of SOC media were added and cells were incubated at 37 °C with shaking for 60-90 min. Cells were then plated onto 2YT agar plates containing the according antibiotic and incubated at 37 °C overnight.

2.2.1.5 Glycerol stock preparation

5 mL of 2YT media containing the according antibiotic were inoculated with *E. coli* cells containing the vector of interest. The cells were incubated with shaking overnight at 37 °C. 850 µL of the overnight culture and 150 µL of sterile glycerol were then mixed by inversion to produce the glycerol stock that was stored frozen at -80 °C.

2.2.2 *Pichia pastoris (P. pastoris)*

2.2.2.1 Media and agar plates

The media used for *P. pastoris* growth was YPD media (1 % (w/v) yeast extract, 2 % (w/v) peptone, 2 % (w/v) dextrose (added after autoclaving)) and the media used for α_1AT expression in *P. pastoris* was YPM media (1 % (w/v) yeast extract, 2 % (w/v) peptone, 0.06 % (v/v) methanol (added after autoclaving)). Test expression media (1 % (w/v) yeast extract, 2 % (w/v) peptone, 0.4 % (w/v) dextrose (added after autoclaving), 3 % (v/v) methanol (added after autoclaving)) was used to screen for α_1AT expressing *P. pastoris* transformants. Media were autoclaved at 120 °C for 20 min and dextrose was autoclaved at 120 °C for 10 min. Kanamycin was added to all media after autoclaving once the media had cooled below 50 °C to a final concentration of 25 µg/mL. RDB plates (1 M sorbitol, 2 % (w/v) dextrose, 1.34 % (w/v) yeast nitrogen base, 4 x 10^{-5} % (w/v) biotin, 0.005 % (w/v) amino acids (L-glutamic acid, L-methionine, L-lysine, L-leucine, and L-isoleucine), 2 % (w/v) agar) were prepared as described in the *Pichia* Expression Kit (Invitrogen).

2.2.2.2 Strains

The *P. pastoris* strain SMD1163 (His4$^-$, pep4, PRB1) was used for protein expression.

2.2.2.3 Preparation of electrocompetent cells

Electrocompetent *P. pastoris* were prepared according to the method of Wu [216] with slight modifications. 6 mL of YPD were inoculated with a scraping of cells from a glycerol stock and grown with shaking overnight at 30 °C. 100 mL of YPD were inoculated with 2 mL of the overnight culture and incubated shaking at 30 °C until the OD_{600} was between 1.0 and 1.5. For each transformation reaction 8 x 10^8 cells were collected by centrifugation (3,346 x g, 10 min, 4 °C) with the number of cells calculated according to the formula 1 OD_{600} = 5 x 10^7 cells/mL. The cell pellet was resuspended in 8 mL of ice-cold, sterile LiAc buffer (100 mM LiAc, 10 mM DTT, 0.6 M sorbitol, 10 mM Tris, pH 7.5) and incubated for 30 min at room temperature. The cells were collected by centrifugation and the cell pellet was washed once with 30 mL of ice-cold, sterile ddH_2O followed by a wash with 30 mL of ice-cold, sterile 1 M sorbitol (3,346 x g, 10 min, 4 °C). The cell pellet was then resuspended in 80 µL of ice-cold, sterile 1 M sorbitol.

2.2.2.4 Transformation

For each pHIL-D2 construct 2 µL of the SacI-linearised DNA (section 2.3.6.3) and 80 µL of electrocompetent *P. pastoris* (section 2.2.2.3) were added to a 0.2-cm long electrode electroporation cuvette (Cell Projects) and pulsed once on the “Pic” programme (2.00 kV, 1 pulse) on a MicroPulser^{TM} electroporator (Bio-Rad). The electroporated cells were immediately diluted in 1 mL of ice-cold, sterile 1 M sorbitol and 100 µL aliquots were spread on RDB plates. The plates were incubated at 30 °C for 3 days or until colonies appeared.

2.2.2.5 Glycerol stock preparation

5 mL of YPD media were inoculated with *P. pastoris* cells containing the construct of interest. The cells were incubated with shaking overnight at 30 °C. The overnight culture was washed once with 1 mL of sterile Tris buffer (100 mM NaCl, 50 mM Tris, pH 8.0)

(16,000 x g, 1 min) and the cell pellet was resuspended in sterile Tris buffer to a total volume of 850 µL. 150 µL of sterile glycerol were added, the sample was mixed by inversion and frozen at -80 °C.

2.3 Molecular biological techniques

2.3.1 Cloning and expression vectors

The vectors used in this study were the 3.5 kb pCR®-Blunt cloning vector (Invitrogen) and the 8.2 kb pHIL-D2 expression vector (Invitrogen). pCR®-Blunt was initially used as sub-cloning vector due to difficulties with Pfu DNA polymerase site-directed mutagenesis of α_1AT DNA in the larger pHIL-D2 vector. KOD DNA polymerase site-directed mutagenesis of α_1AT DNA was later then directly performed in pHIL-D2.

2.3.2 Primers

All primers were purchased from Geneworks (Australia). 40 nmol of freeze-dried sequencing/ PCR grade oligonucleotides were reconstituted in sterile ddH_2O to a final concentration of 100 µM and stored at -20 °C.

Primers (5'-3')		**TM (°C)**
Single tryptophan mutations		
5'W194F	ATCTTCTTTAAAGGCAAA**TTT**GAGAGACCCTTTGAAGTC	62
3'W194F	GACTTCAAAGGGTCTCTC**AAA**TTTGCCTTTAAAGAAGAT	62
5'W238F	TGTAAGAAGCTGTCCAGC**TTT**GTGCTGCTGATGAAATAC	65
3'W238F	GTATTTCATCAGCAGCAC**AAA**GCTGGACAGCTTCTTACA	65
E342 mutations		
5'E342Q	CTGACCATCGAC**CAG**AAAGGTACTGAA	60
3'E342Q	TTCAGTACCTTT**CTG**GTCGATGGTCAG	60
5'E342R	CTGACCATCGAC**CGG**AAAGGTACTGAA	61
3'E342R	TTCAGTACCTTT**CCG**GTCGATGGTCAG	61
Sequencing		
5'M13	GTAAAACGACGGCCAG	46
3'M13	CAGGAAACAGCTATGAC	44
5'*AOX1*	GACTGGTTCCAATTGACAAGC	52
3'*AOX1*	GCAAATGGCATTCTGACATCC	52

2.3.3 Determination of DNA concentration

Where appropriate, sample dilutions were performed in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) and the DNA concentration of the sample was calculated according to the formula 1 optical density at 260 nm (OD_{260}) = 50 ng/µL taking the according dilution factor into consideration.

2.3.4 Site-directed mutagenesis

Complementary primers containing the mutation of interest were designed (section 2.3.2) and mutations incorporated into vectors utilising the QuikChange® method (Stratagene). The mutated vector was subsequently transformed into *E. coli*. All mutagenesis was confirmed by sequencing (section 2.3.11).

2.3.4.1 Pfu DNA polymerase site-directed mutagenesis

Single tryptophan mutations were introduced into α_1AT DNA in pCR®-Blunt [144] utilising Pfu DNA polymerase. The mutated α_1AT DNA was subsequently cut out of pCR®-Blunt using EcoRI (section 2.3.6.2) and ligated into pHIL-D2 (section 2.3.9).

Reaction setup (25 µL):

Sample	Final concentration
10 x Pfu DNA polymerase buffer (Promega)	1 x
dNTPs (Promega)	200 µM
5’ primer (section 2.3.2)	500 nM
3’ primer (section 2.3.2)	500 nM
DNA template	30 ng
Pfu DNA polymerase (Promega)	1.5 U

Temperature cycling:

Step	Temperature	Time	Cycles
Initial Denaturation	95 °C	1 min	1
Denaturation	95 °C	1 min	20
Annealing	X °C	1 min	
Extension	72 °C	11 min	
Final Extension	72 °C	11 min	1

X = 5 °C below the TM of the primers

5 µL of the PCR product were separated on a 1 % agarose gel (section 2.3.5) and the remaining sample digested with DpnI (section 2.3.6.1).

2.3.4.2 KOD DNA polymerase site-directed mutagenesis

E342 mutations were introduced into α_1AT DNA in pHIL-D2 [144] utilising KOD DNA polymerase.

Reaction setup (25 µL):

Sample	Final concentration
10 x KOD DNA polymerase buffer # 2 (Novagen)	1 x
dNTPs (Promega)	400 µM
5’ primer (section 2.3.2)	1.2 µM
3’ primer (section 2.3.2)	1.2 µM
DNA template	60 ng
$MgCl_2$ (Novagen)	1.5 mM
DMSO	4 % (v/v)
KOD DNA polymerase (Novagen)	0.5 U

Temperature cycling:

Step	Temperature	Time	Cycles
Initial Denaturation	94 °C	3 min	1
Denaturation	94 °C	25 s	20
Annealing	50 °C	25 s	
Extension	72 °C	3 min	
Final Extension	72 °C	5 min	1

5 µL of the PCR product were separated on a 1 % agarose gel (section 2.3.5) and the remaining sample digested with DpnI (section 2.3.6.1).

2.3.5 Agarose gel electrophoresis

2.3.5.1 Solutions

50 x TAE buffer: 242 g/L Tris, 57.1 mL/L glacial acetic acid, 100 mL/L 0.5 M EDTA, pH 8.0

Blue/Orange 6 x loading dye (Promega)

1kb DNA ladder (Promega)

SYBR Green (Invitrogen)

2.3.5.2 Procedure

A final concentration of 0.8 % (w/v), 1 % (w/v) or 2 % (w/v) of agarose (Promega) was added to 1 x TAE buffer and dissolved by heating. Upon cooling to approximately 60 °C the solution was poured into a DNA gel frame with a comb inserted and allowed to set. The gel frame was then placed into the running tank (Amersham Biosciences) containing 1 x TAE buffer and the comb was removed. 1kb DNA ladder and DNA samples were prepared with Blue/Orange 6 x loading dye and SYBR Green according to the manufacturers' specifications and loaded into the wells. Electrophoresis was performed at a constant voltage of 100 V until the dye front had reached the bottom of the gel. The gel was then exposed to UV light to visualise DNA bands.

2.3.6 Restriction enzyme digests

2.3.6.1 DpnI digest

DpnI digestion was performed after site-directed mutagenesis (section 2.3.4) to remove parental DNA.

Reaction setup (25 µL):

Sample	Volume (µL)
PCR product	20
10 x Buffer B (Promega)	2.5
10 mg/mL BSA (Promega)	0.25
10 U/µL DpnI (Promega)	1
ddH$_2$O	1.25

Samples were incubated at 37 °C for 2 h. The digests containing Pfu DNA polymerase site-directed mutagenesis products were then frozen at -20 °C until transformed into competent JM107 (section 2.2.1.4). The digests containing KOD DNA polymerase site-directed mutagenesis products were ethanol precipitated (section 2.3.12) and the DNA pellet was resuspended in 10 µL of sterile ddH_2O. Samples were then frozen at -20 °C until transformed into NovaBlue Competent Cells (section 2.2.1.4).

2.3.6.2 EcoRI digest

EcoRI was used to excise template DNA from pCR®-Blunt for ligatation into EcoRI-linearised pHIL-D2 (section 2.3.9).

Reaction setup (50 µL):

Sample	Final concentration
10 x Buffer H (Promega)	1 x
DNA (pCR®-Blunt + insert/ pHIL-D2)	3 - 6 µg
BSA (Promega)	5 µg
EcoRI (Promega)	24 U

Samples were incubated for 3.5 h at 37 °C and then stored at -20 °C until gel purified (section 2.3.7).

2.3.6.3 SacI digest

pHIL-D2 constructs were linearised using SacI for *P. pastoris* transformation.

Reaction setup (300 µL):

Sample	Final concentration
pHIL-D2 construct	ca. 100 µg
10 x Buffer J (Promega)	1 x
BSA (Promega)	30 µg
SacI (Promega)	400 U

Samples were incubated for 5 h at 37 °C and then ethanol precipitated (section 2.3.12). DNA pellets were resuspended in 20 µL of sterile ddH_2O and stored at -20 °C until *P. pastoris* transformation (section 2.2.2.4).

2.3.7 DNA purification from agarose gels

EcoRI digests containing pHIL-D2 were separated on 0.8 % (w/v) agarose gels and EcoRI digests containing pCR®-Blunt including the insert of interest were separated on 2 % (w/v) agarose gels (section 2.3.5). To obtain purified DNA fragments prior to ligation, the fragments of interest were excised from the agarose gels and purified using the Wizard® SV Gel and PCR Clean-Up System (Promega) according to the manufacturer's specifications. In the final purification step, samples were eluted in 50 µL of sterile ddH_2O and frozen at -20 °C until dephosphorylation (section 2.3.8) or ligation (section 2.3.9).

2.3.8 Dephosphorylation

EcoRI-linearised and gel purified pHIL-D2 was dephosphorylated with TSAP to prevent re-circularisation.

Reaction setup (20 µL):

Sample	Volume (µL)
pHIL-D2	16
10 x Ligation Buffer (Promega)	2
1 U/µL TSAP (Promega)	2

The sample was incubated at 37 °C for 15 min. TSAP was then heat-inactivated at 74 °C for 15 min and sample was used for ligation (section 2.3.9).

2.3.9 Ligation

T4 DNA ligase (Promega) and 10 x Ligation buffer (Promega) were used to set up the ligation reactions. The reaction volume was 20 µL containing 3 U of T4 ligase and approximately 200 ng of dephosphorylated pHIL-D2. The molar ratio of gel purified insert was varied. 1:1, 1:3 and 3:1 (insert:vector) ratios were used and the amount of insert required for each molar ratio was determined using Equation 2.1.

$$\frac{\text{ng of vector x insert size (kb)}}{\text{vector size (kb)}} \text{ x molar ratio of insert to vector = ng of insert}$$

Equation 2.1

All ligation reactions were carried out at room temperature for 3 h. Samples were then heat inactivated at 70 °C for 15 min and frozen at -20 °C until transformation into competent JM107 (section 2.2.1.4).

2.3.10 Plasmid purification

For minipreps, 5 mL of 2YT media containing the according antibiotic were inoculated with *E. coli* cells containing the plasmid of interest. The cells were incubated with shaking overnight at 37 °C. Plasmid DNA was purified using the Wizard® Plus SV Minipreps DNA Purification System (Promega) according to the manufacturer's specifications. For midipreps, 100 mL of 2YT containing ampicillin were inoculated with *E. coli* cells containing the plasmid of interest. The cells were incubated with shaking overnight at 37 °C. Plasmid DNA was purified using the Wizard® Plus SV Midipreps DNA Purification System (Promega) according to the manufacturer's specifications.

2.3.11 DNA sequencing

To confirm the integrity of clones and site-directed mutagenesis DNA sequencing was performed. For sequencing reactions with pCR®-Blunt the M13 primers (section 2.3.2) and for sequencing reactions with pHIL-D2 the *AOX1* primers (section 2.3.2) were used.

Sample setup (20 µL):

Sample	Volume (µL)
DNA (ca. 100 ng/µL)	10
10 x PCR reaction buffer + Mg (Roche)	2
Primer (10 µM)	0.5
Big Dye Terminator v3.1	1
DMSO (100 %)	1
ddH_2O	5.5

Temperature cycling:

Step	Temperature	Time	Cycles
Initial Denaturation	90 °C	1 min	1
Denaturation	90 °C	1 min	25
Annealing	50 °C	1 min	
Extension	60 °C	4 min	
Final Extension	60 °C	4 min	1

Following the PCR, the samples were ethanol precipitated (section 2.3.12). Sample analysis was performed by the DNA Sequencing Facility at Micromon (Monash University, Australia).

2.3.12 Ethanol precipitation

0.1 volume of 3 M Na Acetate (pH 5.2) and 2 volumes of 100 % (v/v) ice-cold ethanol were added to the sample and incubated at -80 °C for 20 min. The sample was centrifuged for 10 min at 16,000 x g and the supernatant removed. 5 times the initial sample volume was added in ice-cold 70 % (v/v) ethanol and the sample was vortexed for approximately 3 s. The sample was centrifuged for 10 min at 16,000 x g and the ethanol was carefully removed. This wash step was repeated once. Any remaining ethanol in the sample was then evaporated by incubation of the sample on a dry heating block set at 90 ºC for 5 min.

2.4. Preparation of α_1AT

2.4.1 α_1AT small scale expression

A minimum of six colonies were picked from each RDB plate after *P. pastoris* transformation (section 2.2.2.4) and re-streaked on fresh RDB plates to obtain single colonies. The plates were incubated at 30 °C for 3 days or until colonies appeared which were then analysed for α_1AT expression. At least 3 colonies per plate were screened. Each colony was used to inoculate 2 mL of test expression media which was then incubated shaking for 3 days at 30 °C. The cells were collected by centrifugation (3,346 x g, 10 min, 4 °C) and the cell pellet was resuspended in 1 mL of ddH_2O and transferred into a new tube. The cells were collected by centrifugation (16,000 x g, 2 min) and the cell pellet was resuspended in 300 µL of 10 % (w/v) SDS. The sample was boiled for 10 min and then centrifuged at 16,000 x g for 5 min. 5 µL of the supernatant were analysed by 10 % SDS-PAGE and Western blotting (section 2.6).

2.4.2 α_1AT large scale expression

α_1AT variants were expressed in *P. pastoris*. The transformant with the highest expression level of the α_1AT variant of interest was used to inoculate 5 mL of YPD media and incubated with shaking overnight at 30 °C. Two baffled flasks each containing 500 mL of YPD media were inoculated with 500 µL of overnight culture and incubated shaking at 30 °C for approximately 30 h. The flasks were then left on the bench overnight to collect the cells by gravity. The media was poured off under sterile conditions and 500 mL of YPM were added to each flask to resuspend the cells. The cells were incubated shaking at 30 °C for approximately 72 h and then collected by centrifugation (3,346 x g, 10 min, 4 °C). The cell pellets were pooled and either used straight away for protein purification or frozen in liquid nitrogen and stored at -80 °C.

2.4.3 α_1AT purification

The cell pellets (section 2.4.2) were resuspended in 50 mL of lysis buffer (500 mM NaCl, 25 mM imidazole, 2.5 mM EDTA, 1 mM BME, 20 mM Tris, pH 8.0) containing 0.5-1 Complete EDTA free tablet (Roche) and 2 mM PMSF. Approximately 15 mL aliquots of the cell suspension were distributed to 50 mL Falcon™ tubes (BD Biosciences) and approximately 24 g of 0.5-mm glass beads (Daintree Scientific) were added to the cell suspension. The samples were vortexed for 30 s and then incubated on ice for at least 30 s. This was repeated 7 times for each sample. The samples were then centrifuged (3,346 x g, 15 min, 4 °C) and the supernatants were poured off and pooled. Triton®X-100 was added to a final concentration of 0.05 % (v/v) and sample mixed by inversion. The sample was centrifuged (58,545 x g, 1 h, 6 °C) and the lipid layer was carefully removed by pipetting and discarded. The supernatant was then poured off and passed through 0.22-µm membranes. A minimum of 1.5 volumes of dilution buffer (500 mM NaCl, 25 mM imidazole, 1 mM BME, 20 mM Tris, pH 8.0) were added to the supernatant and the protein was purified on an ÄKTAxpress™ system (GE Healthcare).

The sample was loaded at 1 mL/min onto a 1 mL HisTrap™ HP column (GE Healthcare) pre-equilibrated with dilution buffer. After loading, the column was washed with wash buffer (500 mM NaCl, 25 mM imidazole, 0.5 % (v/v) Triton®X-100, 1 mM BME, 20 mM Tris, pH 8.0) at 0.5 mL/min until stabilisation of the absorbance at 280 nm (A_{280}) with a maximum of 20 column volumes (CV). The protein was eluted with elution buffer (500 mM NaCl, 500 mM imidazole, 0.5 % (v/v) Triton®X-100, 1 mM BME, 20 mM Tris, pH 8.0) at 0.8 mL/min over 5 CV with collection of the A_{280} peak into a storage loop. The sample was then buffer exchanged at a flow rate of 8 mL/min over 1.3 CV on a HiPrep™ 26/10 desalting column (GE Healthcare) pre-equilibrated with low salt buffer (75 mM NaCl, 50 mM Tris, pH 8.0). The A_{280} peak was collected into a storage loop and loaded at 0.5 mL/min onto a 1 mL HiTrap™ Q column (GE Healthcare) pre-equilibrated with low salt buffer. The protein was eluted from the column via NaCl gradient elution: 0 % B to 60 % B in 30 CV followed by 60 % B to 100 % B in 2 CV and then 5 CV at 100 % B where 0 % B = low salt buffer and 100 % B = high salt buffer (300 mM NaCl, 50 mM Tris, pH 8.0). The flow rate was 0.5 mL/min and 1 mL fractions were collected. The Q column elution allowed the separation of monomeric and polymeric α_1AT [144]. The fractions containing monomeric α_1AT were analysed by SDS gel electrophoresis (section 2.5.3) and fractions containing pure α_1AT were pooled and concentrated to approximately 2 mg/mL. The protein was subsequently either used in experiments or aliquoted and frozen at -80 °C. The monomeric state of the purified α_1AT variants was confirmed by discontinuous Native PAGE (section 2.5.2) or size exclusion chromatography (section 2.7).

2.4.4 Determination of α_1AT concentration

The protein concentration was determined by measurement of A_{280} and where appropriate verified by measuring the intrinsic fluorescence of the proteins unfolded in GdnSCN. The extinction coefficients of all α_1AT variants were calculated using the ProtParam tool of the ExPASy Proteomics Server (http://au.expasy.org).

2.5 Polyacrylamide gel electrophoresis (PAGE)

2.5.1 Continuous Native PAGE

Continuous Native PAGE was performed as described in the Mini-PROTEAN® 3 Cell Instruction Manual (Bio-Rad).

2.5.1.1 Solutions

Acrylamide/ bis-Acrylamide 30 % (w/v) mix ratio 29:1 (Sigma Aldrich)

5 x Continuous buffer: 14.64 g/L Imidazole, 41.7 g/L HEPES, pH 7.4

25 % (w/v) APS

TEMED

4 x Sample buffer: 10 % (v/v) electrophoresis buffer, 40 % (v/v) glycerol, 0.01 % (w/v) Bromphenol Blue (Sigma Aldrich)

Staining solution: 40 % (v/v) methanol, 7 % (v/v) acetic acid, 0.025 % (w/v) Coomassie brilliant blue (R250) (Sigma Aldrich)

Destaining solution: 40 % (v/v) methanol, 7 % (v/v) acetic acid

2.5.1.2 Procedure

Gels were prepared and run using a Mini-PROTEAN® system (Bio-Rad). A 6 % gel was prepared as follows.

Setup:

Solution	Volume (mL)
Acrylamide/ bis-Acrylamide 30 %	1.0
5 x Continuous buffer	1.0
ddH$_2$O	3.04
25 % APS	0.01
TEMED	0.005

Samples were diluted with sample buffer and loaded into the wells. The samples were electrophoresed in 1 x continuous buffer for 1 h at 4 °C at a constant current of 20 mA for two gels. Gels were stained in staining solution for approximately 1 h and subsequently destained in destaining solution until protein bands were visible with little background stain. The gels were scanned on an AlphaImager gel documentation system (Alpha Innotech Corporation) and densitometry analyses were performed using the ImageQuant software.

2.5.2 Discontinuous Native PAGE

2.5.2.1 Solutions

Acrylamide/ bis-Acrylamide 30 % (w/v) mix ratio 29:1 (Sigma Aldrich)

4 x Running gel buffer: 1.5 M Tris, pH 8.8

4 x Stacking gel buffer: 0.5 M Tris, pH 6.8

25 % (w/v) APS

TEMED

2 x Loading buffer: 0.125 M Tris, 20 % (v/v) glycerol, 0.2 M DTT, 0.02 % (w/v) Bromphenol Blue (Sigma Aldrich), pH 8.3

Cathode tank buffer: 5.3 mM Tris, 6.8 M glycine, pH 8.9

Anode tank buffer: 0.1 M Tris, pH 7.8

Staining solution: 40 % (v/v) methanol, 7 % (v/v) acetic acid, 0.025 % (w/v) Coomassie brilliant blue (R250) (Sigma Aldrich)

Destaining solution: 40 % (v/v) methanol, 7 % (v/v) acetic acid

2.5.2.2 Procedure

Gels were prepared and run using a Mini-PROTEAN® system (Bio-Rad). A 10 % running gel and a 4 % stacking gel were prepared as follows.

Setup:

Solution	Running gel	Stacking gel
Acrylamide/ bis-Acrylamide 30 % (mL)	1.65	0.22
4 x Running gel buffer (mL)	1.25	-
4 x Stacking gel buffer (mL)	-	0.42
ddH$_2$O (mL)	2.05	1.18
25 % APS (µL)	30	10
TEMED (µL)	4	1.5

Samples were diluted with loading buffer and loaded into the wells. The samples were electrophoresed in tank buffer at 4 °C at a constant current of 25 mA for one gel or 50 mA for two gels until the dye front reached the bottom of the gel. Gels were then stained in staining solution for approximately 1 h and subsequently destained in destaining solution until protein bands were visible with little background stain. The gels were scanned with the Typhoon Trio Variable Mode Laser Imager (GE Healthcare) using the 633 nm laser and the 670 BP 30 emission filter and densitometry analyses were performed using the ImageQuant software.

2.5.3 SDS PAGE

2.5.3.1 Solutions

Acrylamide/ bis-Acrylamide 30 % (w/v) mix ratio 29:1 (Sigma Aldrich)

4 x Running gel buffer: 1.5 M Tris, pH 8.8

4 x Stacking gel buffer: 0.5 M Tris, pH 6.8

10 % (w/v) SDS

25 % (w/v) APS

TEMED

10 x SDS Tank buffer: 0.025 M Tris, 0.192 M glycine, 0.1 % (w/v) SDS, pH 8.3

4 x reducing SDS loading dye: 0.125 M Tris, 40 % (v/v) glycerol, 4 % (w/v) SDS, 0.4 M DTT, 0.04 % Bromphenol Blue (Sigma Aldrich), pH 8.3

Mark 12 ^{TM} (Invitrogen), used for SDS gel setup

Prestained SDS PAGE standards, broad range (Bio-Rad), used for SDS gel setup and subsequent Western blot analysis

Staining solution: 40 % (v/v) methanol, 7 % (v/v) acetic acid, 0.025 % (w/v) Coomassie brilliant blue (R250) (Sigma Aldrich)

Destaining solution: 40 % (v/v) methanol, 7 % (v/v) acetic acid

2.5.3.2 Procedure

Gels were prepared and run using a Mini-PROTEAN® system (Bio-Rad). A 10 % running gel and a 4 % stacking gel were prepared as follows.

Setup:

Solution	Running gel	Stacking gel
Acrylamide/ bis-Acrylamide 30 % (mL)	1.65	0.22
4 x Running gel buffer (mL)	1.25	-
4 x Stacking gel buffer (mL)	-	0.42
10 % SDS (µL)	50	17
ddH$_2$O (mL)	2	1.02
25 % APS (µL)	30	10
TEMED (µL)	4	1.5

Samples were diluted with SDS loading dye and boiled for 2 min. Protein standards and samples were then loaded into the wells and electrophoresed in 1 x SDS tank buffer at a constant current of 35 mA for one gel or 70 mA for two gels until the dye front reached the bottom of the gel. Gels were then either used for Western blotting (section 2.6) or stained in staining solution for approximately 1 h and subsequently destained in destaining solution until protein bands were visible with little background stain.

2.6 Western blot

2.6.1 Solutions

10 % (w/v) SDS

10 x Western transfer buffer (10 x WTB): 60.57 g Tris/L and 150.14 g glycine/L

1 x WTB: 100 mL 10 x WTB, 200 mL methanol, 1 mL 10 % SDS, 700 mL ddH_2O

Tris-buffered Saline (TBS): 2.42 g/L Tris, 11.69 g/L NaCl, pH 7.4

TBST: 0.5 % (v/v) Tween® 20 in TBS

Blocking buffer: 5 % (w/v) skim milk powder (Diploma) in TBS

Primary antibody solution: 10 mL TBS, 51 mg BSA, 2 µL mouse anti-histidine tag antibody (AbD Serotec)

Secondary antibody solution: 5 mL TBS, 25 mg BSA, 1 µL sheep anti-mouse antibody, (H+L) peroxidase conjugated (Chemicon)

ECL™ Western blotting detection reagents (Amersham)

2.6.2 Procedure

After SDS PAGE (section 2.5.3) the proteins were transferred for 1 h at a constant current of 250 mA in ice-cold 1 x WTB onto a Protran® nitrocellulose transfer membrane (0.2 µm pore size, Whatman) using a Mini Trans-Blot® system (Bio-Rad). After the transfer, the membrane was incubated in blocking buffer overnight with shaking at 4 °C and then washed three times for 10 min with TBST. The membrane was then incubated in the primary antibody solution for 1 h shaking at room temperature and washed three times for 10 min with TBST. The membrane was then incubated in the secondary antibody solution for 1 h shaking at room temperature and washed three times for 10 min with TBST. The membrane was then incubated in ECL™ reagent for 1 min and subsequently a Super RX medical x-ray film (Fujifilm) was exposed to the membrane for usually 30 s or longer if appropriate in a Hypercassette™ (Amersham). The film was developed in a SRX-101A film processor (Konica Minolta).

2.7 Size exclusion chromatography

The monomeric state of the α_1AT single tryptophan mutants was confirmed by size exclusion chromatography (SEC). SEC was carried out on an Äkta FPLC (Amersham) using a Superose™ 12 HR 10/30 column (Amersham). The column was pre-equilibrated with 90 mM NaCl and 50 mM Tris, pH 8.0 and elution was monitored following the absorbance at 214 nm (A_{214}).

2.8 Preparation of BSA coated plates

Using a multi-channel pipette (Finnpipette), 200 µL of coating solution (0.5 % (w/v) PEG-8000, 0.05 % (w/v) BSA, 0.01 % (v/v) Triton®X-100) were added to each well of a MICROTEST™ 96 tissue culture plate (BD Biosciences) and incubated overnight at room temperature. The solution was removed by tapping several times on absorbent towels and the wells were washed at least twice with ddH_2O. Plates were left to dry at room temperature and then covered for storage.

2.9 Determination of α_1AT inhibitory properties

2.9.1 Solutions

Chymotrypsin assay buffer: 10 mM $CaCl_2$, 50 mM NaCl, 0.2 % PEG-8000, 100 mM Tris, pH 7.8

HNE assay buffer: 0.5 M NaCl, 0.1 % PEG-8000, 100 mM Tris, pH7.5

Chromogenic substrate for chymotrypsin: N-Succinyl-Ala-Ala-Pro-Phe p-nitroanilide

Chromogenic substrate for HNE: N-Methoxysuccinyl-Ala-Ala-Pro-Val p-nitronailide

2.9.2 Stoichiometry of inhibition (SI)

SI measurements were performed in BSA coated plates (section 2.8) using a ThermoMax Microplate Reader (Molecular Devices) and the SoftMax Pro software. Each well was set up to contain a fixed concentration of enzyme whereas the α_1AT concentration was varied. In assays using chymotrypsin the sample volume was 40 µL at a chymotrypsin concentration of 100 nM and in assays using HNE the sample volume was 50 µL at an HNE concentration of 60 nM. The α_1AT concentration was varied to result in inhibitor to enzyme ratios ranging from 0 to 20 depending on the α_1AT variant analysed. The samples were incubated at 37 °C for 30 min to allow inhibitor-enzyme complex formation. The assay was then initiated by adding 100 µL of either 200 µM substrate for chymotrypsin or 100 µM substrate for HNE. The appearance of p-nitroanilide (pNA) was followed by measuring the absorbance at 405 nm (A_{405}) for 10 min at 37 °C with readings taken every 6 s. The change in A_{405} was plotted as fractional activity of the enzyme against the ratio of inhibitor to enzyme. The data were then fit to a linear regression equation and the SI determined as the point of intersection between the fit and the x-axis.

2.9.3 Continuous association rate constant determination (k_{ass})

Slow tight binding kinetics were performed in BSA coated plates (section 2.8) using a ThermoMax Microplate Reader (Molecular Devices) and the SoftMax Pro software. The reactions were carried out at a final volume of 200 µL. Up to eight assays were performed simultaneously and the concentration of α_1AT was varied between 0 and 10 nM for reactions with chymotrypsin and 0 and 1.8 nM for reactions with HNE. The substrate was present at a concentration of 200 µM and the reactions were initiated by adding the enzyme to a final concentration of 0.2 nM for chymotrypsin and 0.6 nM for HNE. The appearance of pNA was followed by measuring A_{405} for 4 h with readings taken at least every 30 s. The data were analysed using an equation which describes slow tight binding (Equation 2.2) [217],

$$P = V_s t + \frac{V_i - V_s}{k'}(1 - e^{-k't}) \qquad \text{Equation 2.2}$$

where P is the amount of product formed over time, t. k' is the apparent first order rate constant and V_i and V_s are the initial and steady state velocities respectively. k' was then plotted against inhibitor concentration and the resulting data fit to a linear function. The gradient of the linear function is the second order rate constant, k'_{app}, in units of $M^{-1}s^{-1}$. To account for the competition between the serpin and the chromogenic substrate for the active site of the proteinase, the k'_{app} was adjusted using Equation 2.3 to produce k_{app},

$$k_{app} = k'_{app} \left(1 + \frac{[S]}{K_m}\right) \qquad \text{Equation 2.3}$$

where [S] is the substrate concentration and K_m is the Michaelis Menten constant. For N-Succinyl-Ala-Ala-Pro-Phe p-nitroanilide a K_m of 0.087 mM and for N-Methoxysuccinyl-Ala-Ala-Pro-Val p-nitronailide a K_m of 0.07 mM was used. As a serpin's inhibitory pathway is a competitive event between complex formation and substrate behaviour the k_{app} values were multiplied by the according SI values to take into account any error in k_{app} attributed to the substrate pathway and to thus obtain the more accurate association rate constant k_{ass} (Equation 2.4).

$$k_{\mathrm{ass}} = k_{\mathrm{app}} \text{ x SI} \qquad \text{Equation 2.4}$$

2.10 Circular Dichroism (CD) techniques

CD spectra were recorded on a J-815 CD spectrometer (Jasco) and thermal denaturation measurements on a J-810 spectropolarimeter (Jasco) unless otherwise noted. The temperature within the cuvette was monitored by a sensor located in the cuvette holders and maintained by a computer-controlled peltier unit. All measurements were performed in 90 mM NaCl and 50 mM Tris, pH 8.0.

2.10.1 Spectral measurements

Spectral measurements were conducted at 20 °C using a data pitch of 0.1 nm and a scan speed of 100 nm/min. Near-UV CD spectra (250-310 nm) were recorded in a 1-cm path-length quartz cell at a protein concentration of 2 mg/mL. Far-UV CD spectra (250-190 nm) were recorded in a 0.1-cm path-length quartz cell at a protein concentration of 0.25 mg/mL. CD signal at 222 nm (θ_{222}) single point measurements were recorded at 20 °C, with the signal averaged over 15 s at a protein concentration of 0.2 mg/mL.

2.10.2 Thermal denaturation

Thermal denaturation measurements were performed in a 0.1-cm path-length quartz cell at a protein concentration of 0.25 mg/mL. A heating rate of 15 °C/h was applied and the changes in θ_{222} were recorded. The melting temperature (T_{m}) was determined as the midpoint of the aggregation transition as no unfolding transition could be observed.

2.11 Fluorescence spectroscopy

Fluorescence measurements were performed in a 1-cm path-length quartz cell at room temperature unless otherwise noted. The buffer used was 90 mM NaCl and 50 mM Tris, pH 8.0.

2.11.1 Tryptophan fluorescence

Intrinsic tryptophan fluorescence measurements were performed either on a LS50B luminescence spectrometer (Perkin Elmer) or on a FluoroMax®-4 Spectrofluorometer (HORIBA Jobin Yvon). Measurements on the LS50B luminescence spectrometer were performed at an excitation wavelength of 290 nm and emission spectra (290-500 nm) were recorded at a scan speed of 50 nm/min. Unless otherwise noted, the excitation and emission slit widths were 5 nm and the protein concentration was 1 µM. Measurements on the FluoroMax®-4 Spectrofluorometer were performed at an excitation wavelength of 295 nm and emission spectra were collected from 300 to 500 nm (increment 0.5 nm). Excitation and emission slit widths were set at 5 nm and the integration time at 0.1 s. The protein concentration was 2 µM.

2.11.2 bis-ANS fluorescence

bis-ANS fluorescence was measured either on a LS50B luminescence spectrometer (Perkin Elmer) or on a FluoroMax®-4 Spectrofluorometer (HORIBA Jobin Yvon). Measurements were performed at an excitation wavelength of 390 nm and emission spectra were collected from 400 to 600 nm. For measurements on the LS50B luminescence spectrometer a scan speed of 50 nm/min was applied and the excitation and emission slit widths were set at 5 nm. For measurements on the FluoroMax®-4 Spectrofluorometer the excitation and emission slit widths were set at 3 nm and the integration time at 0.1 s with 0.5-nm increments. The protein concentration was 1 µM with bis-ANS present at a five-fold molar excess to the protein.

2.11.3 Quenching

Aliquots of a 2 M KI stock containing 10 mM $Na_2S_2O_3$ were added to the protein samples (1.5-2 µM) and the changes in fluorescence emission intensity of the tryptophan residues (λ_{em} = 330 nm) were measured after each addition. KCl was used to correct for any salt-induced effects. The protein solution had an absorbance at the excitation wavelength of lower than 0.05 after the final addition of KI. Fluorescence values were therefore only corrected for the dilution factors. The quenching data were analysed using the one-component and two-component Stern–Volmer equation as described previously [218, 219].

2.11.4 Anisotropy

Anisotropy measurements were performed on a FluoroMax®-4 Spectrofluorometer (HORIBA Jobin Yvon) in a 1-cm path-length quartz cell at 20 °C with a protein concentration of 2 µM in 90 mM NaCl and 50 mM Tris, pH 8.0. The excitation and emission slit widths were 5 nm and an integration time of 0.1 s was applied. The anisotropy was measured with an excitation at 290 nm and emission between 330 and 400 nm using the L-format method [218] incorporating *G*-factor correction.

2.12 Equilibrium unfolding studies

Stock solutions of GdnHCl or urea were prepared in 90 mM NaCl and 50 mM Tris, pH 8.0 and filtered through 0.22-µm membranes. The concentrations of GdnHCl and urea were determined from the refractive index of the prepared solutions [214]. Equilibrium unfolding curves were obtained by adding a concentrated solution of native protein in 90 mM NaCl and 50 mM Tris, pH 8.0 to a series of denaturant concentrations. The samples were incubated for 2 h at room temperature before analysis. Equilibrium unfolding as a function of denaturant concentration was monitored by measuring either the tryptophan emission spectra to determine the change in the centre of spectral mass wavelength (COSM) as described previously [220] or by following the change in the far-UV CD signal θ_{222}. The unfolding data were fit to either a two-state or a three-state unfolding model using a non-linear least-squares fitting algorithm as described previously [221-

223]. The three-state model recognises the presence of one stable intermediate structure, I, populated during the transition from the native state, N, to the unfolded state, U.

2.13 Kinetic unfolding studies

All kinetic unfolding experiments were performed at room temperature on a SF.18MV stopped-flow apparatus (Applied Photophysics). Protein unfolding was monitored by changes in fluorescence at wavelengths greater than 320 nm using a cutoff filter and an excitation wavelength of 280 nm. Experiments were performed by rapidly mixing 1 volume of protein solution at a protein concentration of 0.5 mg/mL with 10 volumes of concentrated GdnHCl at room temperature. Both solutions contained 90 mM NaCl and 50 mM Tris, pH 8.0. Data were fitted to a double-exponential function with a term included for baseline instability using the manufacturer's software.

2.14 Polymerisation studies

α_1AT polymerisation at 60 °C was carried out at a protein concentration of 10 µM in 90 mM NaCl and 50 mM Tris, pH 8.0. Samples were incubated in a 60 °C water bath and put on ice at various time points to quench the reaction. Ice-cold non-denaturing sample buffer was added and samples were analysed using continuous Native PAGE (section 2.5.1). α_1AT polymerisation at 42 °C was carried out at a protein concentration of 10 µM in 90 mM NaCl, 50 mM Tris, 1 mM EDTA and 5 mM β-mercaptoethanol, pH 8.0. Samples were incubated in a 42 °C water bath and put on ice at various time points to quench the reaction. Ice-cold non-denaturing loading buffer was added and samples were analysed using discontinuous Native PAGE (section 2.5.2).

Chapter 3

Kinetic Instability of the Serpin Z α_1-Antitrypsin Promotes Aggregation

Declaration for Thesis Chapter 3

Declaration by candidate

In the case of Chapter 3, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Experimental design, protein preparation, carried out experiments, data analysis, wrote the first draft of the manuscript and edited the final version.	70

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Vita Levina	Determination of the K_d for bis-ANS binding (Figure 3b and c), TUG gel analysis (Figure 6) and manuscript preparation.	10
Amy L. Robertson	Technical expertise and comments on the manuscript.	5
Mary C. Pearce	Technical expertise and preparation of the ribbon diagram of α_1AT (Figure 1).	
Stephen P. Bottomley	Supervision, intellectual input and manuscript preparation.	

Candidate's Signature: ....[redacted]............ **Date:** 29.5.2012

Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.

(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;

(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

(4) there are no other authors of the publication according to these criteria;

(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and

(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location: Department of Biochemistry and Molecular Biology, Monash University

Signature: ………[redacted]……………………… **Date:** 29/5/12
Vita Levina

Signature: …[redacted]…… **Date:** 29·5·12
Amy L. Robertson

Signature: ……………[redacted]……………… **Date:** 29/5/12
Mary C. Pearce

Signature: ……[redacted]……………… **Date:** 29/5/12
Stephen P. Bottomley

doi:10.1016/j.jmb.2009.11.048 *J. Mol. Biol.* (2010) **396**, 375–383

JMB

Available online at www.sciencedirect.com
ScienceDirect

ELSEVIER

Kinetic Instability of the Serpin Z α_1-Antitrypsin Promotes Aggregation

Anja S. Knaupp, Vita Levina, Amy L. Robertson, Mary C. Pearce and Stephen P. Bottomley*

Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia

Received 7 October 2009; received in revised form 18 November 2009; accepted 19 November 2009
Available online
26 November 2009

Edited by F. Schmid

The serpinopathies encompass a large number of diseases caused by inappropriate conformational change and self-association (polymerization) of a serpin (serine proteinase inhibitor) molecule. The most common serpinopathy is α_1-antitrypsin (α_1AT) deficiency, which is associated with an increased risk for liver cirrhosis, hepatocellular carcinoma and early-onset emphysema. The Z variant of α_1AT, which accounts for 95% of all cases of α_1AT deficiency, polymerizes during synthesis and after secretion. Here, we show using intrinsic and extrinsic fluorescence probes that Z α_1AT exists in a non-native conformation. We examined the thermodynamic stability by transverse urea gradient gel electrophoresis, thermal denaturation and equilibrium guanidine hydrochloride unfolding and found that, despite structural differences between the two proteins, wild-type α_1AT and Z α_1AT display similar unfolding pathways and thermodynamic stabilities. Far-UV circular dichroism and bis-ANS (4,4′-dianilino-1,1′-binaphthyl-5,5′-disulfonic acid, dipotassium salt) fluorescence suggest that the intermediate ensembles formed during unfolding of wild-type α_1AT and Z α_1AT are characterized by similar structural features. Kinetic analysis of the unfolding transition showed that Z α_1AT unfolds at least 1.5-fold faster than the wild type. The biological implications of these data are discussed.

© 2009 Elsevier Ltd. All rights reserved.

Keywords: serpin; polymerization; misfolding; aggregation; metastability

Introduction

α_1-Antitrypsin (α_1AT), the archetypal member of the serpin (serine proteinase inhibitor) superfamily, is the most abundant proteinase inhibitor found in human plasma. It is expressed in hepatocytes and secreted into circulation to control the proteolytic activity of neutrophil elastase in the lower respiratory tract.[1–4] Several mutations that cause misfolding and aggregation of α_1AT at its place of synthesis leading to a reduction in circulating α_1AT have been identified.[5–10] The decreased plasma levels of α_1AT result in a protease–antiprotease imbalance in the lung and early-onset emphysema, whereas the inclusions in the liver lead to an increased risk of developing transient juvenile hepatitis, cirrhosis and hepatocellular carcinoma.[11,12]

The most common pathological mutation is the Z mutation, a Glu342Lys substitution, which accounts for 95% of all cases of α_1AT deficiency.[13] The Z mutation lies at the top of strand 5 of the central β-sheet A and the base of the reactive center loop (RCL) and removes a salt bridge to Lys290.[14] It has been proposed that the disruption of this interaction leads to an expansion of the upper part of β-sheet A, therefore promoting the conformational transitions involved in polymerization.[15]

Spectroscopic analysis of the early polymerization stages of the serpins α_1AT, α_1-antichymotrypsin and PAI-1 revealed the formation of a non-native monomeric species termed M*.[16–20] It has been proposed that the conformational transition to M* involves the opening of β-sheet A, which thus

---

*Corresponding author. E-mail address: steve.bottomley@med.monash.edu.au.

Abbreviations used: serpin, serine proteinase inhibitor; α_1AT, α_1-antitrypsin; bis-ANS, 4,4′-dianilino-1,1′-binaphthyl-5,5′-disulfonic acid, dipotassium salt; RCL, reactive center loop; TUG, transverse urea gradient; GdnHCl, guanidine hydrochloride; COSM, center of spectral mass.

0022-2836/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.

becomes more receptive to the RCL of another molecule, allowing the formation of successive intermolecular loop–sheet linkages (loop A–sheet polymers).[15,21–24] Recently, an alternative structural model for α_1AT polymers has been proposed.[25] This model requires residues of both s5A and the RCL to insert into the same position within the adjoining serpin molecule. Importantly, this linkage, which can be identified through limited proteolysis, has not been observed in any pathological serpin aggregates. Despite the differences in the proposed structural models of serpin polymers, they both rely on the formation of M*, a polymerogenic intermediate, for polymerization to occur.

It remains a question why the Z mutation renders the protein prone to polymerization. It is therefore essential to analyze the effect of the Z mutation on the structure and thermodynamic stability of the protein to further understand the molecular basis of the misfolding mechanism of the most common serpinopathy. The high propensity of Z α_1AT to polymerize has hindered the production of recombinant protein and limited research to plasma-derived Z α_1AT. Not much information exists on the stability of Z α_1AT, and the data available are controversial. For example, unfolding of Z α_1AT in transverse urea gradient (TUG) gels implies that the pathological mutant has similar stabilities as the wild type;[26] a different study estimating the stability of Z α_1AT based on its midpoint of thermal denaturation suggested that the Z mutation significantly decreases the thermodynamic stability of α_1AT.[17] To date, equilibrium or kinetic data on the solution denaturation of Z α_1AT have not been published.

We have recently succeeded in producing recombinant monomeric Z α_1AT.[27] We have now conducted extensive biophysical and biochemical studies to understand the effect of the Z mutation on the structure and the stability of α_1AT. Spectroscopic analysis of the pathological mutant showed that Z α_1AT exists in a different native conformation compared with the wild-type protein with more hydrophobic regions exposed. Interestingly, this more open fold does not affect the protein's thermodynamic stability. Our results indicate that both wild-type α_1AT and Z α_1AT unfold via one intermediate ensemble populated at similar denaturant concentrations. Kinetic characterization of Z α_1AT unfolding revealed a difference in the native-to-intermediate transition, whereas the intermediate-to-unfolded transition remained unaffected, suggesting a decrease in the kinetic barrier in the first unfolding transition of Z α_1AT compared with α_1AT.

Results

Structural properties of Z α_1AT

The fluorescence emission spectrum of tryptophan is sensitive to its environment and can be used to monitor conformational changes within proteins. α_1AT contains two tryptophan residues that are situated at the top of the third strand of β-sheet A (Trp194) and on the second strand of β-sheet B (Trp238) (Fig. 1). Trp194 lies in close proximity (5 Å) to Glu342, which is replaced by a lysine in Z α_1AT. The location of Trp194 allows it to act as a reporter for structural changes around the top part of β-sheet A and thus for its opening, which is proposed to be the initial event in serpin polymerization.[17,21,22]

Denatured M α_1AT and Z α_1AT have almost identical fluorescence emission spectra, both displaying similar emission intensities and maxima (λ_{max}) at 350 nm. However, they exhibit different spectral properties in their native conformation (Fig. 2a). While M α_1AT has a λ_{max} of 338 nm, the Z mutation leads to a small red shift of λ_{max} to 341 nm and a 47% (λ_{em} = 330 nm) increase in emission intensity. These spectral data suggest that there is a conformational change around at least one of the two tryptophan residues in Z α_1AT.

In order to obtain additional information on the changes in the environment of the tryptophan residues, we performed fluorescence quenching experiments (Fig. 2b and c). Incremental amounts of potassium iodide were added to native and unfolded M α_1AT and Z α_1AT, and the reductions in fluorescence emission intensity (λ_{em} = 330 nm) were measured and fit to the Stern–Volmer equation (Fig. 2b and c; Table 1). The Stern–Volmer plots of denatured M α_1AT and Z α_1AT were linear and similar, indicating equal accessibility of the two tryptophan residues to iodide (Fig. 2c). The Stern–Volmer plots of native M α_1AT and Z α_1AT both showed downward curvature, indicating different Stern–Volmer constants for the two tryptophan residues in either protein (Fig. 2b). Analysis of these data with a two-component model revealed that the quenchable tryptophan residue is more accessible to the quencher in Z α_1AT (f_1 = 0.70) than in M α_1AT (f_1 = 0.55) (Table 1).

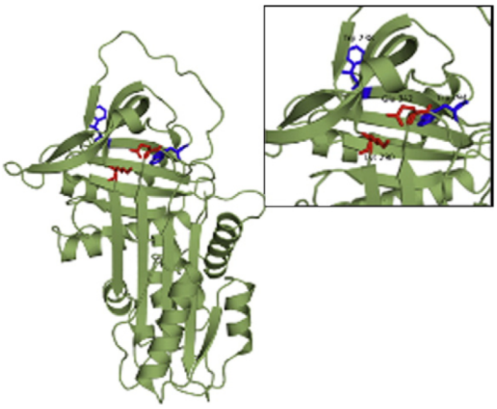


Fig. 1. Schematic representation of α_1AT. Ribbon diagram of α_1AT (Protein Data Bank code 1QLP) with the side chains of Trp194, Trp238, Lys220 and Glu342 highlighted as sticks. The inset shows a close-up view of the region around Glu342.

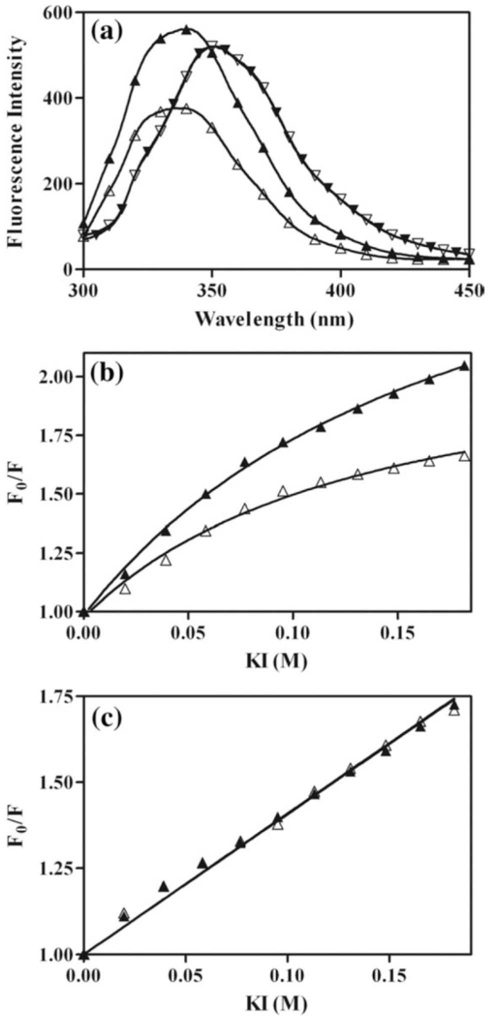


Fig. 2. Intrinsic fluorescence properties of α_1AT and Z α_1AT. (a) Tryptophan emission spectra ($\lambda_{ex}=290$ nm) of native α_1AT (△) and Z α_1AT (▲) and α_1AT (▽) and Z α_1AT (▼) unfolded in 6 M GdnHCl. Each tryptophan emission spectrum represents the average of five scans. (b, c) Stern–Volmer plots for iodide quenching of native (b) and unfolded (c) α_1AT (△) and Z α_1AT (▲). Data of native and unfolded samples were fitted to a two-component model and a one-component model, respectively, and results are shown in Table 1.

These results provide direct evidence that there are structural differences between M α_1AT and Z α_1AT and suggest that Z α_1AT has a more open native fold.

In addition, we used the fluorescent probe bis-ANS (4,4′-dianilino-1,1′-binaphthyl-5,5′-disulfonic acid, dipotassium salt), whose spectral properties depend on the hydrophobicity of the binding sites in the protein. Incubation of bis-ANS with native Z α_1AT resulted in an approximately 11-fold increase in bis-ANS fluorescence intensity ($\lambda_{em}=480$ nm) compared with α_1AT (Fig. 3a). To determine whether the increase in bis-ANS fluorescence intensity observed for Z α_1AT was due to an increase in the quantum yield of bis-ANS or an increase in the number of bis-ANS molecules bound (i.e., larger stoichiometry of bis-ANS binding to Z α_1AT than to α_1AT), we determined the equilibrium dissociation constant, K_d, and the number of bis-ANS binding sites per α_1AT molecule (Fig. 3b). The K_d was found to be 0.7 μM for Z α_1AT and a Job plot indicated that there is only one bis-ANS binding site in Z α_1AT (Fig. 3c). Due to the much weaker binding affinity of bis-ANS to α_1AT, a wider bis-ANS concentration range was used; however, we were unable to reach saturation and determine K_d for α_1AT reliably.

We further utilized circular dichroism (CD) to analyze structural differences between α_1AT and Z α_1AT. Far-UV CD analysis revealed only minor differences in secondary structure (Fig. 4a). At identical protein concentrations, Z α_1AT showed on average a 12% decrease in signal intensity at 222 nm. Near-UV CD analysis also suggested differences in the tertiary structure (Fig. 4b). Z α_1AT had a decreased ellipticity in the regions between 280 and 300 nm and between 270 and 290 nm compared with α_1AT. These spectral changes can be attributed to differences in the environment of tryptophan and tyrosine residues. Smaller changes were detected in the environment of phenylalanine residues, which are responsible for peaks in the region between 250 and 270 nm. Therefore, the CD spectra are in accord with our intrinsic fluorescence data suggesting conformational changes around the tryptophan residues.

Unfolding pathway and thermodynamic stability of Z α₁AT

Controversy exists as to why the Glu342Lys substitution leads to an increased propensity of Z α_1AT to self-associate. To address the possibility that an altered thermodynamic stability is responsible for the polymerization-prone nature of Z α_1AT, we performed TUG gel electrophoresis, thermal denaturation and equilibrium unfolding studies. The midpoints of thermal unfolding, as assessed by far-UV CD, were 57.2 °C for α_1AT and

Table 1. Fluorescence quenching parameters of α_1AT and Z α_1AT

	Conformation	K_1 (M^{-1})	K_2 (M^{-1})	f_1	f_2
α_1AT	Native	16.65	1.0×10^{-7}	0.55	0.45
	Unfolded	4.09	—	—	—
Z α_1AT	Native	17.11	1.0×10^{-7}	0.70	0.30
	Unfolded	4.08	—	—	—

Iodide quenching data of native and unfolded M α_1AT and Z α_1AT were fitted to a two-component model and a one-component model, respectively. Subscript 1 refers to the quenchable population; subscript 2, the unquenchable population. *K* represents the dynamic quenching constant; *f*, the fraction contribution.

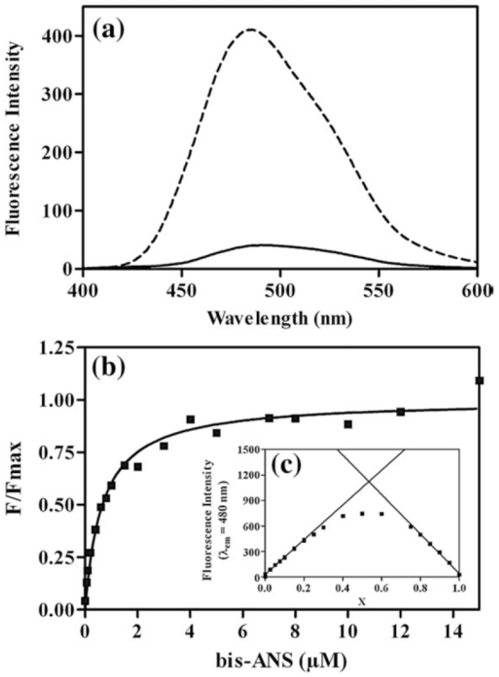


Fig. 3. bis-ANS fluorescence in the presence of α_1AT and Z α_1AT. (a) bis-ANS emission spectra ($\lambda_{ex} = 390$ nm) in the presence of native α_1AT (continuous line) and Z α_1AT (broken line). (b) Determination of the dissociation constant (K_d) for bis-ANS binding to Z α_1AT. (c) Job plot of bis-ANS binding to Z α_1AT. The fluorescence intensity at different molar ratios of bis-ANS to Z α_1AT was measured and plotted against the molar fraction of Z α_1AT (X). The intersection of the limiting slope lines as $X \rightarrow 0$ and $X \rightarrow 1$ is located at $X = 0.53$, indicating a binding stoichiometry of bis-ANS to Z α_1AT of 1:1.

56.1 °C for Z α_1AT (Table 2). Guanidine hydrochloride (GdnHCl) unfolding was also monitored by following changes in tryptophan fluorescence, monitored by changes in the center of spectral mass (COSM) (Fig. 5a; Table 2) and changes in secondary structure measured at 222 nm by far-UV CD (Fig. 5b; Table 2). Refolding was also attempted, but, in common with many other serpins,[28,29] Z α_1AT did not refold to active material. As a result of aggregation during refolding, a full thermodynamic analysis of the equilibrium unfolding pathway could not be performed. Hence, transition midpoints of the unfolding curves were used as a relative measure of protein stability. The well-characterized unfolding of α_1AT follows a three-state pathway.[21,30,31] Under identical conditions, Z α_1AT undergoes the same three-state unfolding reaction with the formation of one stable intermediate ensemble at similar GdnHCl concentrations as α_1AT (Fig. 5 and Table 2). Changes in solvent exposure of the tryptophan residues and loss of secondary structure upon unfolding coincided with a first transition centered at 0.8 M GdnHCl and a second transition at approximately 3 M GdnHCl for both α_1AT and Z α_1AT.

bis-ANS binding was also monitored during equilibrium unfolding (Fig. 5c) and showed an increase in bis-ANS fluorescence upon intermediate formation of both α_1AT and Z α_1AT with maximum fluorescence emission intensity at approximately 1 M GdnHCl, suggesting the formation of structurally similar intermediate ensembles. The thermodynamic stability of Z α_1AT was further assessed by TUG gel electrophoresis (Fig. 6). There was a single native-to-unfolded transition for both α_1AT and Z α_1AT that occurred with similar midpoints of approximately 2.2 and 2.1 M urea, respectively.

In summary, the consistency of all the equilibrium unfolding data and the similar midpoints of α_1AT and Z α_1AT unfolding during thermal denaturation and TUG gel electrophoresis suggest that the Z mutation does not significantly perturb the stability of α_1AT.

Unfolding kinetics of Z α_1AT

Stopped-flow analyses of the folding and unfolding reactions of α_1AT and Z α_1AT were carried out in order to determine differences in the unfolding/folding kinetics between the two proteins. Kinetic analysis of the refolding reaction was hindered by

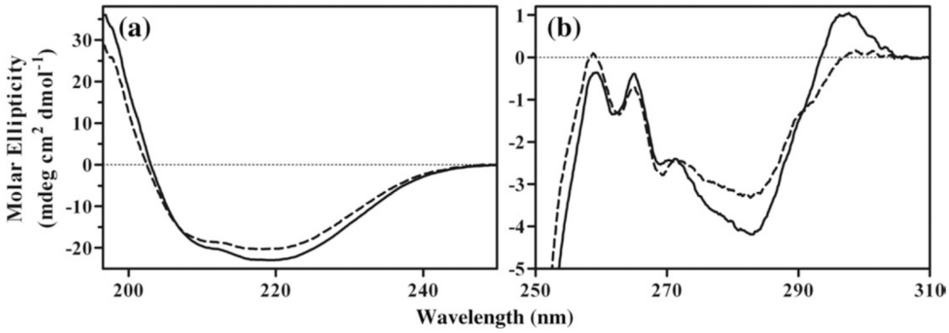


Fig. 4. CD analysis of M α_1AT and Z α_1AT. (a, b) CD spectra in the far-UV range (a) and near-UV range (b) of native α_1AT (continuous line) and Z α_1AT (broken line). Each CD spectrum represents the average of five scans.

Table 2. Stability characteristics of α$_1$AT and Z α$_1$AT

	Analysis parameter	$D_{\mathrm{mN}\to\mathrm{I}}$ (M)	$D_{\mathrm{mI}\to\mathrm{U}}$ (M)	T_m (°C)
α$_1$AT	COSM	0.7±0.1	3.5± 0.2	—
Z α$_1$AT		0.9± 0.1	3.3± 0.2	—
α$_1$AT	Far-UV CD	0.8± 0.1	2.7± 0.2	57.2± 0.2
Z α$_1$AT		0.8± 0.1	2.9± 0.2	56.1± 0.6

Results from the GdnHCl-induced equilibrium unfolding (Fig. 5a and b) and the thermal denaturation were normalized and fitted to three-state and two-state unfolding curves, respectively. The results presented for the equilibrium analysis are the average of five individual curves, whereas the results for the thermal denaturation are the average of four individual curves. $D_{\mathrm{mN}\to\mathrm{I}}$ and $D_{\mathrm{mI}\to\mathrm{U}}$ represent the midpoints of denaturation for the N→I and I→U transitions, respectively. T_m represents the midpoint of thermal denaturation.

the formation of aggregates that prevented the collection of accurate data for either protein. Unfolding of α$_1$AT and Z α$_1$AT was monitored by changes in tryptophan fluorescence upon unfolding in different GdnHCl concentrations, and these data were fitted to a double-exponential equation (Fig. 7). Whereas the intermediate-to-unfolded transition was found to be characterized by similar rate constants at all GdnHCl concentrations for α$_1$AT and Z α$_1$AT, the native-to-intermediate transition was found to occur at a 1.5-fold faster rate for Z α$_1$AT. These results suggest that the Z mutation decreases the kinetic barrier of the first unfolding transition but has no effect on the second transition. Kinetic analysis of α$_1$AT and Z α$_1$AT further confirmed the tryptophan spectral data that native Z α$_1$AT is characterized by a higher fluorescence intensity than α$_1$AT.

Discussion

Z α$_1$AT, the most common pathological mutant of α$_1$AT, predisposes the carrier to a number of liver- and lung-related diseases due to polymerization and thus inactivation. Our data clearly demonstrate that Z α$_1$AT adopts an altered but stable native conformation. Glu342, located on top of β-sheet A at the base of the RCL (P17), forms a salt bridge to Lys290.[14] The presence of the Z mutation removes this interaction, which is predicted to open β-sheet A, making the protein more receptive to the RCL of another molecule.[15,24] Recent indirect evidence supporting an alternative conformation of Z α$_1$AT came from an elegant peptide study;[15] however, these findings could also be explained by different dynamics of the two proteins. Therefore, we analyzed the spectroscopic properties of α$_1$AT and Z α$_1$AT to obtain direct information on the structural differences. α$_1$AT has two tryptophan residues: Trp194 and Trp238. Trp194, one of the most conserved residues among the serpin superfamily, is located at the top of the third strand of β-sheet A (s3A) in close proximity to Glu342. Trp238 is situated on the stable β-sheet B.[32] The spectroscopic analysis of Z α$_1$AT revealed a red shift in $\lambda_{\max}$ and a large increase in fluorescence emission intensity. These changes indicate increased solvent exposure and a change in the environment of at least one of the two tryptophan residues. Fluorescence quenching showed that in both proteins one of the two tryptophan residues is almost inaccessible to the quencher and thus extensively buried in the protein. Resolving the heterogeneous emission spectrum into two components revealed that the quenchable fraction contributes 55% and 70% to the total fluorescence emission in M α$_1$AT and Z α$_1$AT, respectively. Taking the location of Trp194 at the top of β-sheet A into consideration, the proposed opening of β-sheet A in Z α$_1$AT would most likely affect the surrounding environment and the solvent exposure of this tryptophan residue. Opening of β-sheet A would also result in the exposure of buried hydrophobic regions. This was confirmed by the use of bis-ANS, whose fluorescent properties depend on the environment surrounding the probe. bis-ANS binds to Z α$_1$AT with a significantly higher affinity than to M α$_1$AT and with a large increase in

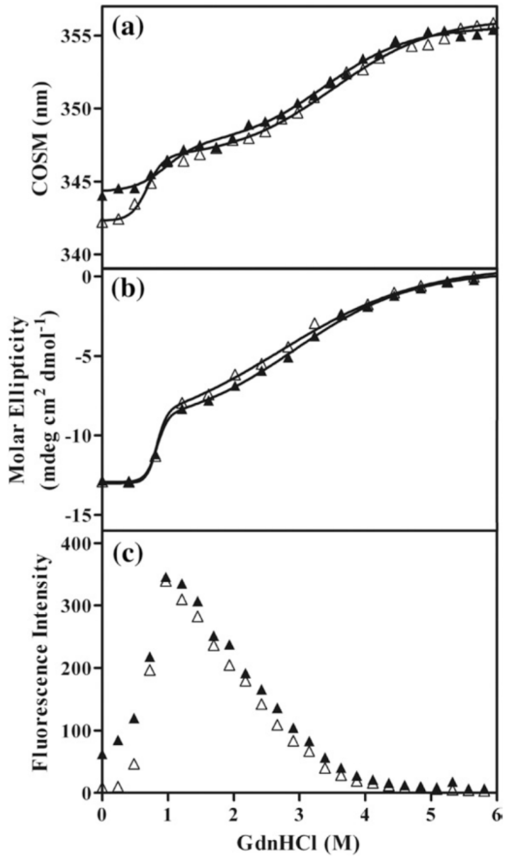


Fig. 5. GdnHCl-induced equilibrium denaturation of α$_1$AT and Z α$_1$AT. Structural changes of α$_1$AT (△) and Z α$_1$AT (▲) during unfolding were followed by changes in COSM (a), changes in the far-UV CD signal at 222 nm (b) and bis-ANS fluorescence intensity (λ_{em}=480 nm) (c). Lines represent three-state curve fits.

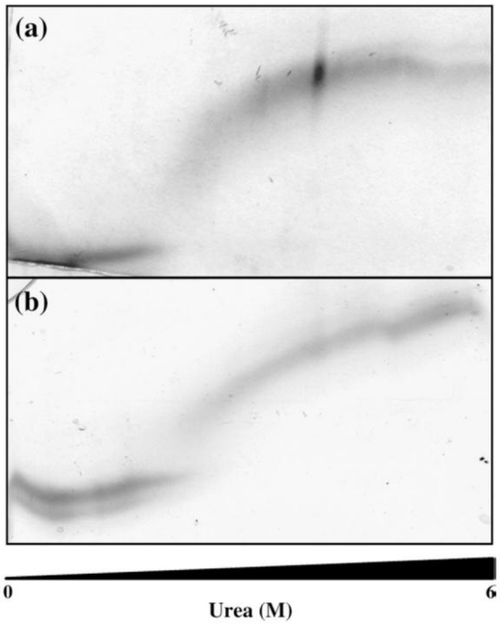


Fig. 6. TUG gel analysis of α_1AT and Z α_1AT. (a, b) Unfolding of α_1AT (a) and Z α_1AT (b) in 0 to 6 M urea gradient gel (TUG). α_1AT and Z α_1AT were applied across the top of the gradient gel with a urea gradient perpendicular to the direction of electrophoresis.

quantum yield as well as a blue shift in emission maximum indicative of a more hydrophobic environment. Both, binding affinity and spectral properties of bis-ANS bound to the native and the pathogenic conformations of α_1AT, suggest a higher degree of hydrophobic exposure of Z α_1AT compared with M α_1AT, thus again indicating differences in the tertiary structures. We have recently shown that the rate of formation of the α_1AT-proteinase complex before branching into the inhibitory and substrate pathways is over 5-fold slower for Z α_1AT compared with M α_1AT.[27] This indicates that the Z mutation affects the initial binding of the proteinase and provides further evidence that Z α_1AT exists in a different conformation than M α_1AT.

Surprisingly, this conformational change does not alter the thermodynamic stability of Z α_1AT. Several studies have shown that opening of β-sheet A plays a central role in the polymerization mechanism of Z α_1AT,[21–23] but it is controversial whether this alters the structural stability of Z α_1AT. Dafforn *et al.* proposed that Z α_1AT is characterized by a decreased thermodynamic stability compared with M α_1AT as concluded from a lower melting temperature of 8 °C.[17] A decreased thermodynamic stability would increase the concentration of M*, the monomeric intermediate species that self-associates, and therefore increase the polymerization rate of Z α_1AT. Following several distinct approaches to determine the structural stability of Z α_1AT, including TUG gel analysis, thermal denaturation and GdnHCl denaturation, we did not observe any significant difference in stability between α_1AT and Z α_1AT. We found the midpoint of thermal denaturation of Z α_1AT to be within 1 °C of the midpoint of thermal denaturation of α_1AT, suggesting similar thermal stabilities of the two proteins.

Structural changes of α_1AT and Z α_1AT during equilibrium unfolding were found to be similar for both proteins and compatible with a three-state unfolding model. α_1AT and Z α_1AT adopt an intermediate ensemble under identical conditions, 1 M GdnHCl, with similar structural and thermodynamic properties. These data indicate that the salt bridge between Glu342 and Lys290 is not present in the intermediate ensemble but only in the native state of α_1AT. Hence, it is surprising that the mutation has little effect on the stability of the native state. However, it has been shown previously that solvent-exposed salt bridges contribute little to overall protein stability.[33]

The mechanism of serpin polymerization and protein fibrillization in general has been largely attributed to destabilization of the native structure. Destabilization of the native state leads to an increased population of aggregation-prone species that under suitable conditions then self-associate. There is, however, an increasing number of proteins, such as T7 endonuclease I,[34] ataxin-3,[35,36] prion protein[37,38] and now Z α_1AT, in which aggregation-enhancing mutations do not correlate with a decreased protein stability.

This important finding obviously raises the question of how the Z mutation affects the kinetics of the unfolding/folding transition of α_1AT. It has been proposed previously from an *in vitro* translation system that the Z mutation leads to a slower folding rate and thus the accumulation of the polymerogenic intermediate M*.[26] Unfortunately, the high propensity of M* to aggregate did not allow us to obtain reliable refolding kinetic data. However,

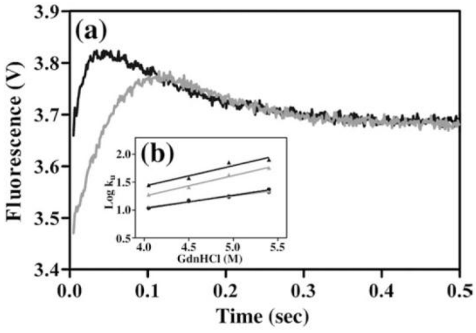


Fig. 7. Kinetics of unfolding of α_1AT and Z α_1AT. (a) Representative stopped-flow traces for α_1AT (grey) and Z α_1AT (black) unfolded in 4 M GdnHCl and corrected for baseline. (b) The observed rate constants for the N→I (▲) and I→U (●) transitions were plotted as a function of GdnHCl concentration for α_1AT (grey) and Z α_1AT (black).

we were able to follow the unfolding transition of α_1AT and Z α_1AT and found that the Z mutation affects the first transition (N→I) but that the second transition (I→U) remains unaffected. This suggests that the Z mutation lowers the kinetic barrier of the first unfolding transition, leading to a faster rate of unfolding of Z α_1AT compared with M α_1AT. Polymerization of Z α_1AT not only at its place of synthesis but also in circulation and in the lung could thus partly be due to this increased unfolding rate.

Overall, our data indicate that Z α_1AT adopts a non-native but functional conformation. Whereas the thermodynamic stability remains unaffected, the Z mutation lowers the kinetic barrier of the unfolding transition to the intermediate ensemble, which is known to readily polymerize. Therefore, our data strongly support the hypothesis that the Glu342Lys mutation alters the energy landscape and leads to a folding defect of α_1AT.

Materials and Methods

Materials

GdnHCl was obtained from Sigma, and bis-ANS was purchased from Invitrogen.

Production of recombinant α_1AT

M α_1AT and Z α_1AT were expressed in *Pichia pastoris* and purified as described previously.[27]

Spectroscopic methods

Tryptophan emission spectra were recorded at room temperature on a Perkin-Elmer LS50B luminescence spectrometer in a 1 -cm path-length quartz cell. Samples were excited at 290 nm, and spectra were collected from 290 to 500 nm. Excitation and emission slit widths were set at 5 nm, and a scan speed of 50 nm/min was used. Unless otherwise noted, the protein concentration was 1 μM in 50 mM Tris and 90 mM NaCl, pH 8.0. The absorbance at the excitation wavelength remained below 0.05. CD spectra were recorded on a Jasco J-815 CD spectrometer at 20 °C using a data pitch of 0.1 nm and a scan speed of 100 nm/min. Near-UV CD spectra (250–310 nm) were recorded in a 1 -cm path-length quartz cell at a protein concentration of 2 mg/mL. Far-UV CD spectra (250–190 nm) were recorded in a 0. 1-cm path-length quartz cell at a protein concentration of 0.25 mg/mL. Θ_{222} measurements were recorded at 20 °C, with the signal averaged over 15 s at a protein concentration of 0.2 mg/mL unless otherwise noted.

Chemical denaturation

Stock solutions of GdnHCl in 90 mM NaCl and 50 mM Tris, pH 8.0, were prepared and filtered through 0. 22-μm membranes. The GdnHCl concentration was determined by refractive index measurements as described previously.[39] Unfolding curves were obtained by adding a concentrated solution of native protein in 90 mM NaCl and 50 mM Tris, pH 8.0, to a series of GdnHCl concentrations. These solutions were incubated for 2 h at room temperature before analysis. Equilibrium unfolding as a function of GdnHCl concentration was monitored by measuring either the tryptophan emission spectra to determine the change in COSM wavelength as described previously[32,40] or the change in signal at 222 nm in the far-UV CD spectra. All equilibrium unfolding data were analyzed as described previously[21,31,41] using a three-state model that recognizes the population of one stable intermediate structure (I) during the transition from the folded state (N) to the unfolded state (U).

Conformational stability was also investigated by TUG gel analysis. The urea gradient used was 0 to 6 M in 7.5% acrylamide as described previously.[42] A total of 0.04 μg of each protein was loaded, and the gels were run at a constant current of 15 mA per gel for 105 min.

Thermal denaturation

Thermal denaturation measurements were performed in a 0.1 -cm path-length quartz cell at a protein concentration of 0.25 mg/mL in 90 mM NaCl and 50 mM Tris, pH 8.0. A heating rate of 15 °C/h was applied, and the change in signal at 222 nm in the far-UV spectra was determined. The thermal denaturation data were fit to a two-state unfolding model using a non-linear least-squares fitting algorithm as described previously to obtain the midpoint of denaturation T_m.[43]

bis-ANS binding studies

bis-ANS emission spectra (400–600 nm) were obtained in the presence of M α_1AT and Z α_1AT. bis-ANS was added in 5-fold excess to both proteins, and the emission spectra ($\lambda_{ex}=390$ nm) were recorded. Excitation and emission slit widths were set at 5 nm, and a scan speed of 50 nm/min was used.

To determine the apparent dissociation constant K_d for the interaction between bis-ANS and α_1AT, we measured the fluorescence emission intensities ($\lambda_{ex}=410$ nm, $\lambda_{em}=480$ nm) of samples containing 100 nM protein and increasing concentrations of bis-ANS in 50 mM Tris and 100 mM NaCl, pH 7.4, in a 1-cm path-length quartz cuvette. Where required, the final fluorescence intensities were corrected for the inner-filter effect as described previously.[44] The emission intensities at 480 nm were plotted against the bis-ANS concentration, and the data were fitted to a binding model using a least-squares analysis to determine the K_d.

The binding stoichiometry was determined as described previously.[45] The fluorescence of samples containing various molar ratios of Z α_1AT and bis-ANS was measured and plotted against the molar fraction of Z α_1AT. For all samples, the total concentration of Z α_1AT and bis-ANS was kept constant at 5 μM ($\sim 7\times K_d$). The stoichiometry of bis-ANS binding to Z α_1AT was determined from the intersection of the limiting slope lines.

Fluorescence quenching experiments

Fluorescence quenching measurements were performed in 90 mM NaCl and 50 mM Tris, pH 8.0, at 0 and 6 M GdnHCl. Aliquots of a 2 M KI stock containing 10 mM $Na_2S_2O_3$ were added to the protein samples, and the changes in fluorescence emission intensity of the tryptophan residues (λ_{em} = 330 nm) were measured after each step. The quenching data of three measurements for unfolded protein and of five measurements for native protein were averaged. The protein solution had an absorbance at the excitation wavelength of lower than 0.05 after the final addition of KI. Fluorescence values were therefore only corrected for the dilution factors. The quenching data were analyzed using the one-component and two-component Stern–Volmer equations as described previously.[46]

Unfolding kinetics

Protein unfolding was monitored by changes in fluorescence at wavelengths >320 nm using a cutoff filter and an excitation wavelength of 280 nm. Experiments were performed on an Applied Photophysics SF.18MV stopped-flow apparatus. Unfolding was performed by rapidly mixing 1 volume of protein solution at a protein concentration of 0.5 mg/mL with 10 volumes of concentrated GdnHCl at room temperature. Both solutions contained 90 mM NaCl and 50 mM Tris, pH 8.0. Data were collected from at least seven experiments that were averaged and fitted to a double-exponential function with a term included for baseline instability using the manufacturer's software.

Acknowledgements

This work was supported by the National Health and Medical Research Council (of which S.P.B. is a Senior Research Fellow) through a program grant. We thank members of the Bottomley laboratory for helpful discussions.

References

1. Fregonese, L. & Stolk, J. (2008). Hereditary alpha-1-antitrypsin deficiency and its clinical consequences. *Orphanet J. Rare Dis.* **3**, 16.
2. Gadek, J. E., Fells, G. A., Zimmerman, R. L., Rennard, S. I. & Crystal, R. G. (1981). Antielastases of the human alveolar structures. Implications for the protease–antiprotease theory of emphysema. *J. Clin. Invest.* **68**, 889–898.
3. Knaupp, A. S. & Bottomley, S. P. (2009). Serpin polymerization and its role in disease—the molecular basis of alpha$_1$-antitrypsin deficiency. *IUBMB Life*, **61**, 1–5.
4. Stoller, J. K. & Aboussouan, L. S. (2005). alpha$_1$-Antitrypsin deficiency. *Lancet*, **365**, 2225–2236.
5. Curiel, D. T., Chytil, A., Courtney, M. & Crystal, R. G. (1989). Serum alpha 1-antitrypsin deficiency associated with the common S-type (Glu264–Val) mutation results from intracellular degradation of alpha 1-antitrypsin prior to secretion. *J. Biol. Chem.* **264**, 10477–10486.
6. Elliott, P. R., Stein, P. E., Bilton, D., Carrell, R. W. & Lomas, D. A. (1996). Structural explanation for the deficiency of S alpha 1-antitrypsin. *Nat. Struct. Biol.* **3**, 910–911.
7. Lomas, D. A., Elliott, P. R., Sidhar, S. K., Foreman, R. C., Finch, J. T., Cox, D. W. *et al.* (1995). alpha 1-Antitrypsin Mmalton (Phe52-deleted) forms loop–sheet polymers in vivo. Evidence for the C sheet mechanism of polymerization. *J. Biol. Chem.* **270**, 16864–16870.
8. Lomas, D. A., Finch, J. T., Seyama, K., Nukiwa, T. & Carrell, R. W. (1993). alpha 1-Antitrypsin Siiyama (Ser53 → Phe). Further evidence for intracellular loop–sheet polymerization. *J. Biol. Chem.* **268**, 15333–15335.
9. Roberts, E. A., Cox, D. W., Medline, A. & Wanless, I. R. (1984). Occurrence of alpha-1-antitrypsin deficiency in 155 patients with alcoholic liver disease. *Am. J. Clin. Pathol.* **82**, 424–427.
10. Seyama, K., Nukiwa, T., Takabe, K., Takahashi, H., Miyake, K. & Kira, S. (1991). Siiyama (serine 53 (TCC) to phenylalanine 53 (TTC)). A new alpha 1-antitrypsin-deficient variant with mutation on a predicted conserved residue of the serpin backbone. *J. Biol. Chem.* **266**, 12627–12632.
11. Lomas, D. A. & Parfrey, H. (2004). alpha$_1$-Antitrypsin deficiency: 4. Molecular pathophysiology. *Thorax*, **59**, 529–535.
12. Stoller, J. K. & Aboussouan, L. S. (2005). Antitrypsin deficiency. *Lancet*, **365**, 2225–2236.
13. Ioachimescu, O. C. & Stoller, J. K. (2005). A review of alpha-1 antitrypsin deficiency. *COPD*, **2**, 263–275.
14. Loebermann, H., Tokuoka, R., Deisenhofer, J. & Huber, R. (1984). Human alpha 1-proteinase inhibitor. Crystal structure analysis of two crystal modifications, molecular model and preliminary analysis of the implications for function. *J. Mol. Biol.* **177**, 531–557.
15. Mahadeva, R., Dafforn, T. R., Carrell, R. W. & Lomas, D. A. (2002). 6-mer peptide selectively anneals to a pathogenic serpin conformation and blocks polymerization. Implications for the prevention of Z alpha(1)-antitrypsin-related cirrhosis. *J. Biol. Chem.* **277**, 6771–6774.
16. James, E. L. & Bottomley, S. P. (1998). The mechanism of alpha 1-antitrypsin polymerization probed by fluorescence spectroscopy. *Arch. Biochem. Biophys.* **356**, 296–300.
17. Dafforn, T. R., Mahadeva, R., Elliott, P. R., Sivasothy, P. & Lomas, D. A. (1999). A kinetic mechanism for the polymerization of alpha$_1$-antitrypsin. *J. Biol. Chem.* **274**, 9548–9555.
18. Zhou, A., Faint, R., Charlton, P., Dafforn, T. R., Carrell, R. W. & Lomas, D. A. (2001). Polymerization of plasminogen activator inhibitor-1. *J. Biol. Chem.* **276**, 9115–9122.
19. Devlin, G. L., Chow, M. K., Howlett, G. J. & Bottomley, S. P. (2002). Acid denaturation of alpha$_1$-antitrypsin: characterization of a novel mechanism of serpin polymerization. *J. Mol. Biol.* **324**, 859–870.
20. Crowther, D. C., Serpell, L. C., Dafforn, T. R., Gooptu, B. & Lomas, D. A. (2003). Nucleation of alpha 1-antichymotrypsin polymerization. *Biochemistry*, **42**, 2355–2363.
21. James, E. L., Whisstock, J. C., Gore, M. G. & Bottomley, S. P. (1999). Probing the unfolding pathway of alpha$_1$-antitrypsin. *J. Biol. Chem.* **274**, 9482–9488.
22. Dong, A., Meyer, J. D., Brown, J. L., Manning, M. C. & Carpenter, J. F. (2000). Comparative Fourier transform infrared and circular dichroism spectroscopic analysis of alpha$_1$-proteinase inhibitor and ovalbumin in aqueous solution. *Arch. Biochem. Biophys.* **383**, 148–155.
23. Zhou, A., Stein, P. E., Huntington, J. A., Sivasothy, P., Lomas, D. A. & Carrell, R. W. (2004). How small peptides block and reverse serpin polymerisation. *J. Mol. Biol.* **342**, 931–941.

24. Lomas, D. A., Evans, D. L., Stone, S. R., Chang, W. S. & Carrell, R. W. (1993). Effect of the Z mutation on the physical and inhibitory properties of alpha 1-antitrypsin. *Biochemistry*, **32**, 500–508.
25. Yamasaki, M., Li, W., Johnson, D. J. & Huntington, J. A. (2008). Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature*, **455**, 1255–1258.
26. Yu, M. H., Lee, K. N. & Kim, J. (1995). The Z type variation of human alpha 1-antitrypsin causes a protein folding defect. *Nat. Struct. Biol.* **2**, 363–367.
27. Levina, V., Dai, W., Knaupp, A. S., Kaiserman, D., Pearce, M. C., Cabrita, L. D. *et al.* (2009). Expression, purification and characterization of recombinant Z alpha(1)-antitrypsin—the most common cause of alpha(1)-antitrypsin deficiency. *Protein Expression Purif.* **68**, 226–232.
28. Cabrita, L. D. & Bottomley, S. P. (2004). How do proteins avoid becoming too stable? Biophysical studies into metastable proteins. *Eur. Biophys. J.* **33**, 83–88.
29. Cabrita, L. D., Irving, J. A., Pearce, M. C., Whisstock, J. C. & Bottomley, S. P. (2007). Aeropin from the extremophile *Pyrobaculum aerophilum* bypasses the serpin misfolding trap. *J. Biol. Chem.* **282**, 26802–26809.
30. Powell, L. M. & Pain, R. H. (1992). Effects of glycosylation on the folding and stability of human, recombinant and cleaved alpha 1-antitrypsin. *J. Mol. Biol.* **224**, 241–252.
31. Cabrita, L. D., Dai, W. & Bottomley, S. P. (2004). Different conformational changes within the F-helix occur during serpin folding, polymerization, and proteinase inhibition. *Biochemistry*, **43**, 9834–9839.
32. Tew, D. J. & Bottomley, S. P. (2001). Probing the equilibrium denaturation of the serpin alpha(1)-antitrypsin with single tryptophan mutants; evidence for structure in the urea unfolded state. *J. Mol. Biol.* **313**, 1161–1169.
33. Sali, D., Bycroft, M. & Fersht, A. R. (1991). Surface electrostatic interactions contribute little of stability of barnase. *J. Mol. Biol.* **220**, 779–788.
34. Guo, Z. & Eisenberg, D. (2007). The mechanism of the amyloidogenic conversion of T7 endonuclease I. *J. Biol. Chem.* **282**, 14968–14974.
35. Chow, M. K., Ellisdon, A. M., Cabrita, L. D. & Bottomley, S. P. (2004). Polyglutamine expansion in ataxin-3 does not affect protein stability: implications for misfolding and disease. *J. Biol. Chem.* **279**, 47643–47651.
36. Ellisdon, A. M., Thomas, B. & Bottomley, S. P. (2006). The two-stage pathway of ataxin-3 fibrillogenesis involves a polyglutamine-independent step. *J. Biol. Chem.* **281**, 16888–16896.
37. Liemann, S. & Glockshuber, R. (1999). Influence of amino acid substitutions related to inherited human prion diseases on the thermodynamic stability of the cellular prion protein. *Biochemistry*, **38**, 3258–3267.
38. Swietnicki, W., Petersen, R. B., Gambetti, P. & Surewicz, W. K. (1998). Familial mutations and the thermodynamic stability of the recombinant human prion protein. *J. Biol. Chem.* **273**, 31048–31052.
39. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
40. Cabrita, L. D., Gilis, D., Robertson, A. L., Dehouck, Y., Rooman, M. & Bottomley, S. P. (2007). Enhancing the stability and solubility of TEV protease using in silico design. *Protein Sci.* **16**, 2360–2367.
41. Pearce, M. C., Rubin, H. & Bottomley, S. P. (2000). Conformational change and intermediates in the unfolding of alpha 1-antichymotrypsin. *J. Biol. Chem.* **275**, 28513–28518.
42. Creighton, T. E. (Ed.). (1995). IRL Press, Oxford, UK.
43. Dafforn, T. R., Pike, R. N. & Bottomley, S. P. (2004). Physical characterization of serpin conformations. *Methods*, **32**, 150–158.
44. Lakowicz, J. R. (1999). *Principles of Fluorescence Spectroscopy*, 2nd edit. Kluwer Academic/Plenum Press, New York.
45. Huang, C. Y. (1982). Determination of binding stoichiometry by the continuous variation method: the Job plot. *Methods Enzymol.* **87**, 509–525.
46. Lehrer, S. S. (1971). Solute perturbation of protein fluorescence. The quenching of the tryptophyl fluorescence of model compounds and of lysozyme by iodide ion. *Biochemistry*, **10**, 3254–3263.

Chapter 4

Structural Change in β-Sheet A of Z α_1-Antitrypsin Is Responsible for Accelerated Polymerisation and Disease

Declaration for Thesis Chapter 4

Declaration by candidate

In the case of Chapter 4, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Experimental design, protein construction and preparation, carried out all experiments, data analysis, wrote the first draft of the manuscript and edited the final version.	80

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Stephen P. Bottomley	Supervision, intellectual input and manuscript preparation.	

Candidate's Signature: ..[redacted]................ **Date:** 29.5.2012

Anja S. Knaupp

Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.

(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;

(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

(4) there are no other authors of the publication according to these criteria;

(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and

(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location: Department of Biochemistry and Molecular Biology, Monash University

Signature: ……[redacted]……………… **Date:** …29/5/12…

doi:10.1016/j.jmb.2011.09.013 *J. Mol. Biol.* (2011) **413**, 888–898

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Structural Change in β-Sheet A of Z α_1-Antitrypsin Is Responsible for Accelerated Polymerization and Disease

Anja S. Knaupp and Stephen P. Bottomley*

Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia

Received 21 April 2011;
received in revised form
7 September 2011;
accepted 11 September 2011
Available online
16 September 2011

Edited by S. Radford

Keywords:
serpin;
misfolding;
aggregation;
protein inhibitor;
conformation

The presence of the Z mutation (Glu342Lys) is responsible for more than 95% of α_1-antitrypsin (α_1AT) deficiency cases. It leads to increased polymerization of the serpin α_1AT during its synthesis and in circulation. It has been proposed that the Z mutation results in a conformational change within the folded state of antitrypsin that enhances its polymerization. In order to localize the conformational change, we have created two single tryptophan mutants of Z α_1AT and analyzed their fluorescence properties. α_1AT contains two tryptophan residues that are located in distinct regions of the molecule: Trp194 at the top of β-sheet A and Trp238 on β-sheet B. We have replaced each tryptophan residue individually with a phenylalanine in order to study the local environment of the remaining tryptophan residue in both M and Z α_1AT. A detailed fluorescence spectroscopic analysis of each mutant was carried out, and we detected differences in the emission spectrum, the Stern–Volmer constant for potassium iodide quenching and the anisotropy of only Trp194 in Z α_1AT compared to M α_1AT. Our data reveal that the Z mutation results in a conformational change at the top of β-sheet A but does not affect the structural integrity of β-sheet B.

© 2011 Elsevier Ltd. All rights reserved.

Introduction

Z α_1-antitrypsin (α_1AT) is the most common pathological variant of the serine protease inhibitor (serpin) α_1AT. In healthy individuals, α_1AT is produced in the liver and secreted into circulation where its primary role is to control the proteolytic activity of neutrophil elastase in the lower respiratory tract.[1–4] The Z mutation (Glu342Lys) leads to retention of polymerized α_1AT in the endoplasmic reticulum (ER), which promotes a protease–antiprotease imbalance in the lung.[5] The subsequent increase in neutrophil elastase activity eventually leads to early-onset emphysema, and the presence of Z α_1AT polymers in the ER of hepatocytes can culminate in severe liver disease.[4,6]

Unlike Z α_1AT, the structure and inhibitory properties of wild-type α_1AT have been characterized extensively. α_1AT consists of 394 amino acid residues that adopt the native serpin fold consisting of three β-sheets (A–C) and nine α-helices (A–I) (Fig. 1).[7] Extending from the body of the serpin is the reactive center loop (RCL) that connects strand five of the central β-sheet A (s5A) with strand one of β-sheet C (s1C). The RCL contains the scissile bond, which is recognized by the target proteinase. Upon cleavage of the scissile bond, α_1AT undergoes an extensive conformational change during which β-sheet A accepts the proximal end of the RCL as an additional strand. The target proteinase, which is covalently attached to the RCL, is translocated to the

---

**Corresponding author.* E-mail address: steve.bottomley@monash.edu.

Abbreviations used: α_1AT, α_1-antitrypsin; ER, endoplasmic reticulum; RCL, reactive center loop; HNE, human neutrophil elastase; COSM, center of spectral mass; bis-ANS, 4,4′-dianilino-1,1′-binaphthyl-5,5′-disulfonic acid; SI, stoichiometry of inhibition.

0022-2836/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.

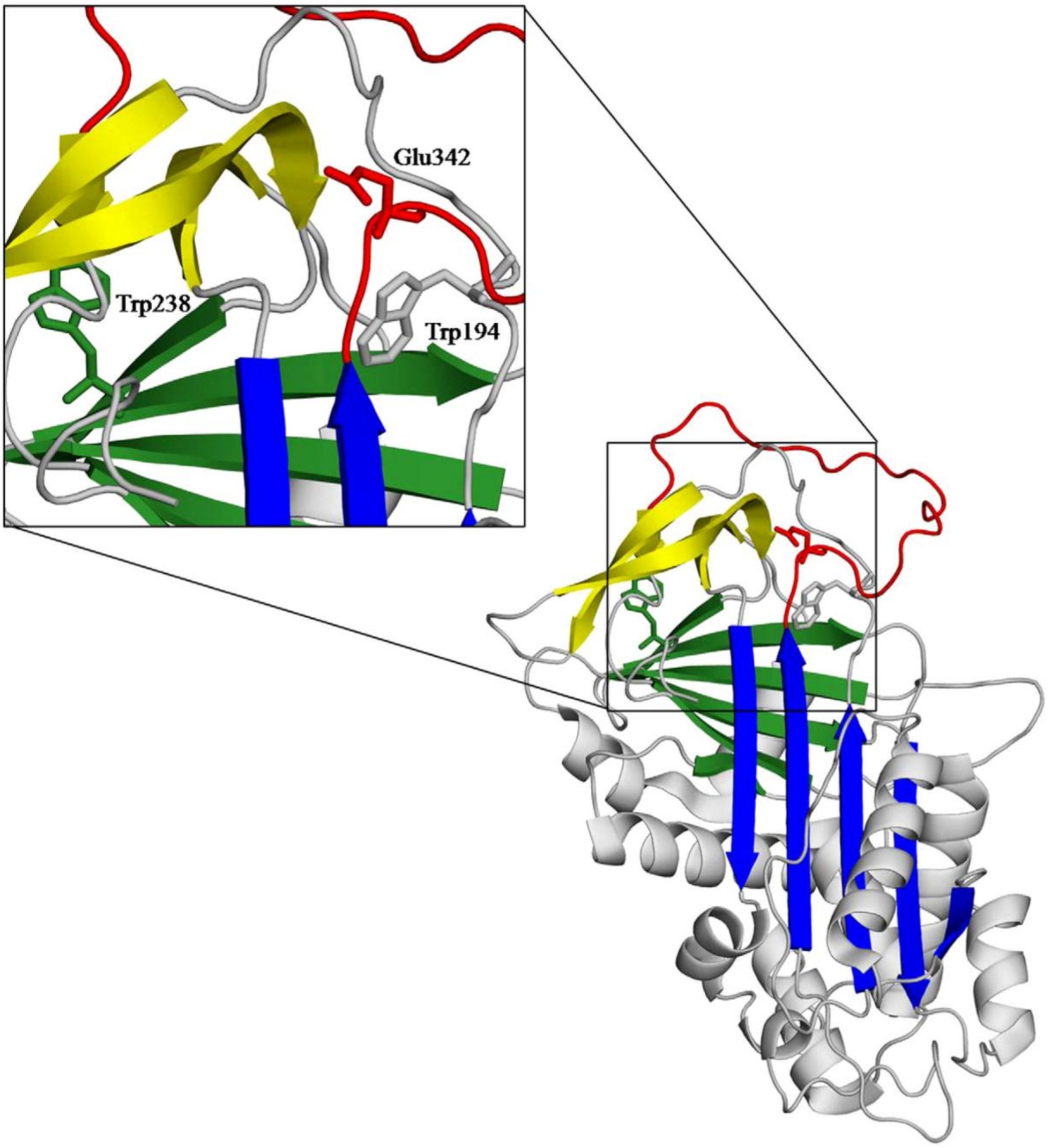


Fig. 1. Schematic representation of α_1AT. Ribbon diagram of α_1AT (Protein Data Bank ID: 1QLP) with the side chains of Trp194, Trp238 and Glu342 shown as sticks. β-Sheet A is in blue, β-sheet B is in green, β-sheet C is in yellow and the RCL is in red. The inset shows a close-up view of the region around Trp194, Trp238 and Glu342.

distal end of the serpin. The translocation process promotes a structural change in the proteinase and, therefore, its inactivation.[8–10]

Several regions within α_1AT have been identified that control the conformational change that occurs during proteinase inhibition.[11–13] One of these regions is the breach region, which is the area at the top of β-sheet A where the RCL first inserts after cleavage by the proteinase.[13] The breach region is characterized by a high degree of sequence conservation across the serpin superfamily, which suggests that the hydrogen bond network in this area is particularly important for the inhibitory process.[11] Glu342, which is located at the top of s5A and the base of the RCL, is one of these highly conserved residues forming a salt bridge to Lys290 and a hydrogen bond to Thr203.[11,14] Removal of these interactions by the introduction of a positively charged lysine at position 342 leads to the pathological phenotype of Z α_1AT.

The exact nature of the pathological phenotype of Z α_1AT, however, remains to be elucidated. There is evidence that the Z mutation results in a decrease in inhibitory activity[15,16] and in an altered native conformation of α_1AT.[17–19] Why the Z mutation leads to increased polymerization of α_1AT and which role the alternative native conformation plays during its polymerization process are still not clear though. *In vitro* studies suggest that wild-type and Z α_1AT polymerize from the folding intermediate (I)[5,18,20] and that the Z mutation slows the final folding step, thus accumulating I.[21] This might explain why, in the cell, Z α_1AT polymers are retained in the ER. However, there is also evidence that Z α_1AT can polymerize once it has attained its native state (N) under *in vitro* conditions under which wild-type α_1AT does not form polymers.[5,15,18] *In vivo* this is supported by the detection of pathological Z α_1AT polymers in circulation[22,23] and in the lung.[24] In order to obtain information about the native conformation of Z α_1AT, we have previously compared the intrinsic fluorescence properties of Z and wild-type α_1AT and showed that the Z mutation affects the environment of at least one of the two naturally occurring tryptophan residues.[17] α_1AT contains two tryptophan residues, Trp194 and Trp238, which are located in two distinct regions of the molecule (Fig. 1). Trp194 is situated above strand three of β-sheet A (s3A), whereas Trp238 is situated on the opposite side of the molecule on the second strand of β-sheet B (s2B). In the current study, we have individually replaced one of the two naturally occurring tryptophan residues in Z α_1AT with a phenylalanine to study the conformational integrity of the environment of the remaining tryptophan residue. Our fluorescence data indicate that the top of β-sheet A is disrupted by the Z mutation, leading to an increased solvent exposure of Trp194. Additionally, we were able to define an area of the Z α_1AT molecule that is structurally unaffected by the Z mutation: β-sheet B. Together, these data are the first two pieces in the puzzle of the structure of Z α_1AT.

Results

Effects of the single tryptophan mutations on structure

In our previous study, we constructed two single tryptophan mutants of M α_1AT by individually replacing the two naturally occurring tryptophan residues, Trp194 and Trp238, with phenylalanine.[25] Their distinct location and intrinsic fluorescence spectroscopic properties make the two tryptophan residues ideal probes for structural changes in β-sheet A (Trp194) and β-sheet B (Trp238). We have now introduced the same mutations into the most common pathological mutant of α_1AT, Z α_1AT, to create the two single tryptophan mutants Z $\alpha_1\mathrm{AT}_{\mathrm{W194}}$ and Z $\alpha_1\mathrm{AT}_{\mathrm{W238}}$, which contain Trp194 and Trp238, respectively. Far-UV circular dichroism (CD) and 4,4′-dianilino-1,1′-binaphthyl-5,5′-disulfonic acid (bis-ANS) were used to analyze the effect of the single tryptophan mutations on secondary structure content and exposure of hydrophobic regions (Fig. 2). In agreement with our previously published data, the far-UV CD and bis-ANS emission spectra indicate that Z α_1AT is in an altered native conformation characterized by a decrease in secondary structure content and an increased exposure of hydrophobic regions in comparison to M α_1AT.[17] Analysis of the secondary structural contribution using the CDSSTR algorithm[26–28] on the Web server DICHROWEB[29,30] suggests that wild-type α_1AT is approximately 30% helical, 26% β-sheet, 19% turns and 25% unordered, which agrees well with the DSSP secondary structure assignment[31] for the

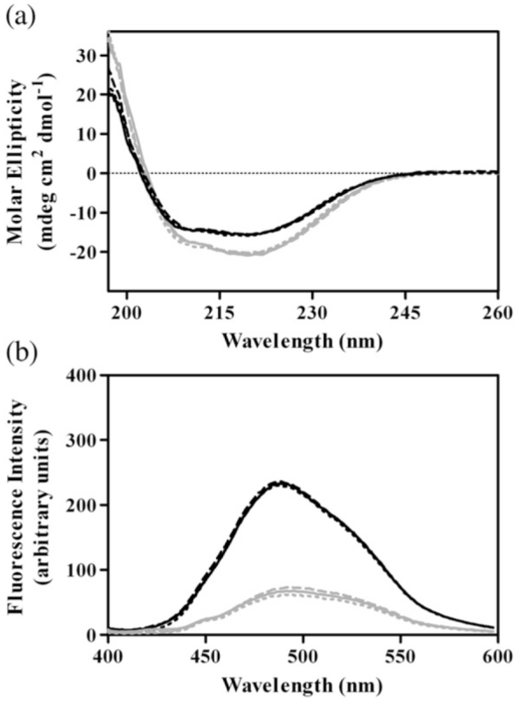


Fig. 2. Spectroscopic analysis of the single tryptophan mutants of M and Z α_1AT. (a) Far-UV CD spectra of M (gray) and Z (black) α_1AT (—), $\alpha_1\mathrm{AT}_{\mathrm{W194}}$ (---) and $\alpha_1\mathrm{AT}_{\mathrm{W238}}$ (····). (b) bis-ANS fluorescence in the presence of M (gray) and Z (black) α_1AT (—), $\alpha_1\mathrm{AT}_{\mathrm{W194}}$ (---) and $\alpha_1\mathrm{AT}_{\mathrm{W238}}$ (····).

crystal structure of α_1AT (Protein Data Bank ID: 1QLP).[32] Z α_1AT was calculated to be approximately 27% helical, 24% β-sheet, 19% turns and 29% unordered. Thus, our data suggest that the Z mutation leads to an increase in the unordered content concomitant with a decrease in the helical and β-sheet contribution in α_1AT. In terms of the single tryptophan mutations, the far-UV CD and bis-ANS emission spectra suggest that the tryptophan-to-phenylalanine mutations do not affect the native fold of M or Z α_1AT. A detailed fluorescence spectroscopic analysis of each mutant was conducted in order to determine areas of conformational difference between M and Z α_1AT.

Fluorescence analysis of the single tryptophan mutants

Differences in the tryptophan fluorescence emission spectra and the Stern–Volmer plots for potassium iodide quenching of M and Z α_1AT suggest that at least one of the two tryptophan residues is more solvent exposed in the pathological mutant.[17] In order to determine to which extent the local environment of each of the two tryptophan residues is affected by the Z mutation, we performed a detailed fluorescence spectroscopic analysis of each single tryptophan mutant (Table 1). M and Z $\alpha_1\mathrm{AT_{W238}}$ were found to have similar emission spectra, with emission maxima (λ_{max}) of 335 ± 1.0 nm and 334 ± 1.0 nm, respectively (Fig. 3a and Table 1). Differences were, however, detected in the fluorescence emission spectra of M and Z $\alpha_1\mathrm{AT_{W194}}$. While M $\alpha_1\mathrm{AT_{W194}}$ has a λ_{max} of 324 ± 1.0 nm, a large red-shift to 334 ± 1.0 nm and an approximately 16% increase in emission intensity ($\lambda_{em} = 330$ nm) were observed for Z $\alpha_1\mathrm{AT_{W194}}$ (Fig. 3b and Table 1). These spectral results suggest that the Z mutation induces a conformational change in the region around Trp194 that leads to Trp194 being situated in a more polar environment. Potassium iodide quenching experiments were carried out to determine the solvent accessibility of each tryptophan residue (Fig. 3c and Table 1). In accord with the emission spectra, the Stern–Volmer constant (K_{SV}) was significantly increased ($p < 0.001$) for Z $\alpha_1\mathrm{AT_{W194}}$ ($K_{SV} = 1.90 \pm 0.04\ \mathrm{M^{-1}}$) in comparison to M $\alpha_1\mathrm{AT_{W194}}$ ($K_{SV} = 0.39 \pm 0.04\ \mathrm{M^{-1}}$) but similar for M $\alpha_1\mathrm{AT_{W238}}$ ($K_{SV} = 3.45 \pm 0.12\ \mathrm{M^{-1}}$) and Z $\alpha_1\mathrm{AT_{W238}}$ ($K_{SV} = 3.56 \pm 0.07\ \mathrm{M^{-1}}$). Further, we analyzed the anisotropy (r) of the proteins. Whereas there was no difference in the anisotropy of M $\alpha_1\mathrm{AT_{W238}}$ ($r = 0.084 \pm 0.011$) and Z $\alpha_1\mathrm{AT_{W238}}$ ($r = 0.084 \pm 0.022$), the anisotropy was significantly decreased ($p < 0.0001$) from 0.141 ± 0.024 for M $\alpha_1\mathrm{AT_{W194}}$ to 0.092 ± 0.010 for Z $\alpha_1\mathrm{AT_{W194}}$. This decrease in anisotropy suggests an increase in the side-chain mobility of Trp194 in Z α_1AT in comparison to M α_1AT, consistent with the higher degree of solvent exposure observed. Taking the location of Trp194 into consideration, these results clearly show that there are conformational differences at the top of β-sheet A of Z and M α_1AT.

Effects of the single tryptophan mutations on inhibitory function

We have previously shown that mutating either Trp194 or Trp238 to a phenylalanine does not affect the inhibitory properties of M α_1AT.[25] The effects of the same mutations on Z α_1AT were, however, strikingly different. The inhibitory properties of Z $\alpha_1\mathrm{AT_{W194}}$ and Z $\alpha_1\mathrm{AT_{W238}}$ were determined against bovine chymotrypsin and human neutrophil elastase (HNE) (Table 2). The stoichiometry of inhibition (SI), which describes the ratio of substrate and inhibition reaction after the formation of the covalent acyl ester intermediate, was significantly increased for both Z $\alpha_1\mathrm{AT_{W194}}$ and Z $\alpha_1\mathrm{AT_{W238}}$ when compared to either M or Z α_1AT. Similar to previously published data, our results show that approximately 40% of Z α_1AT follows the substrate pathway with SI values of 1.9 ± 0.3 and 1.7 ± 0.18 for chymotrypsin and HNE, respectively.[15,16] The replacement of Trp238 with a phenylalanine (Z $\alpha_1\mathrm{AT_{W194}}$) led to SI values of 4.2 ± 0.2 and 7.2 ± 0.51, and the replacement of Trp194 (Z $\alpha_1\mathrm{AT_{W238}}$) led to even higher SI values with 7.8 ± 0.3 and 10.5 ± 0.3 for chymotrypsin and HNE, respectively.

Table 1. Fluorescence spectroscopic properties of the native state and during equilibrium unfolding of the single tryptophan mutants

	λ_{max} emission (nm)	K_{SV} ($\mathrm{M^{-1}}$)	r[a]	$D_{\mathrm{mN}\rightarrow\mathrm{I}}$[b] (M)	$D_{\mathrm{mI}\rightarrow\mathrm{U}}$[b] (M)	$D_{\mathrm{mN}\rightarrow\mathrm{U}}$[c] (M)
M $\alpha_1\mathrm{AT_{W194}}$	324 ± 1.0	0.39 ± 0.04	0.141 ± 0.024	2.1 ± 0.02	5.3 ± 0.07	—
M $\alpha_1\mathrm{AT_{W238}}$	335 ± 1.0	3.45 ± 0.12	0.084 ± 0.011	—	—	6.3 ± 0.12
Z $\alpha_1\mathrm{AT_{W194}}$	334 ± 1.0	1.90 ± 0.04	0.092 ± 0.010	1.8 ± 0.08	5.3 ± 0.09	—
Z $\alpha_1\mathrm{AT_{W238}}$	334 ± 1.0	3.56 ± 0.07	0.084 ± 0.022	—		6.4 ± 0.05

[a] Average anisotropy values ($\lambda_{em} = 340$ nm).

[b] The urea-induced unfolding curves obtained following the changes in COSM of M and Z $\alpha_1\mathrm{AT_{W194}}$ were analyzed using the three-state unfolding function to determine the midpoints of the N→I transition ($D_{\mathrm{mN}\rightarrow\mathrm{I}}$) and the I→U transition ($D_{\mathrm{mI}\rightarrow\mathrm{U}}$).

[c] The urea-induced unfolding curves obtained following the change in COSM of M and Z $\alpha_1\mathrm{AT_{W238}}$ were analyzed using the two-state unfolding function to determine the midpoint of the N→U transition ($D_{\mathrm{mN}\rightarrow\mathrm{U}}$).

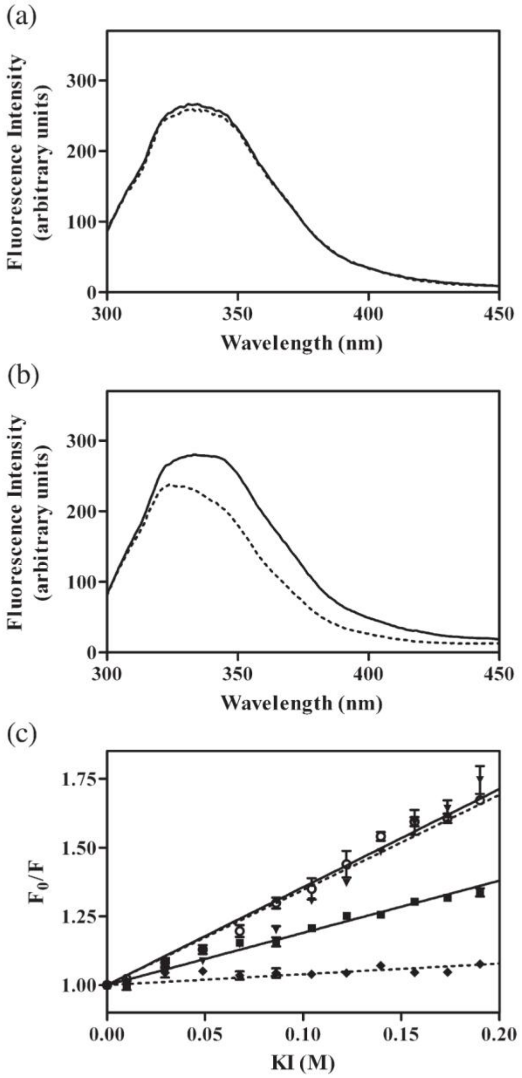


Fig. 3. Intrinsic fluorescence properties of the single tryptophan mutants of M and Z α_1AT. (a) Fluorescence emission spectra ($\lambda_{ex}=290$ nm) of M α_1AT$_{W238}$ (---) and Z α_1AT$_{W238}$ (—). (b) Fluorescence emission spectra ($\lambda_{ex}=290$ nm) of M α_1AT$_{W194}$ (---) and Z α_1AT$_{W194}$ (—). (c) Stern–Volmer plots for potassium iodide quenching of M α_1AT$_{W194}$ (◆), M α_1AT$_{W238}$ (▼), Z α_1AT$_{W194}$ (■) and Z α_1AT$_{W238}$ (○). Lines represent one-component fits.

Further, the k_{assapp} was decreased by approximately 2-fold for Z α_1AT, up to 6-fold for Z α_1AT$_{W194}$ and up to 11-fold for Z α_1AT$_{W238}$ compared to M α_1AT and its single tryptophan mutants (Table 2). However, when the k_{assapp} values were corrected for the according SI values to obtain the overall rate constant k_{ass}, all values were similar to the values of M α_1AT. Taken together, these data indicate that the Z mutation and the single tryptophan mutations affect the translocation of the proteinase after docking to α_1AT.

Effects of the single tryptophan mutations on unfolding

The thermal stability of each α_1AT mutant was assessed using far-UV CD (Table 3). The replacement of Trp194 with a phenylalanine did not affect the thermal stability of either M or Z α_1AT. Mutating Trp238 resulted in a slight decrease in the midpoints of thermal unfolding (T_m) from 57.2 ± 0.2 to 56.5 ± 0.4 for M α_1AT and from 56.1 ± 0.6 to 54.6 ± 0.4 for Z α_1AT.

Urea-induced unfolding of the single tryptophan mutants was monitored using far-UV CD (Fig. 4). The unfolding pathway of α_1AT has been studied extensively and shown to follow a reversible three-state unfolding reaction with an intermediate ensemble being populated at low denaturant concentrations.[21,25,33–37] It has been shown that Z α_1AT also unfolds via a three-state mechanism with similar midpoints of transition but leads to the formation of aggregates during refolding.[17,21] The increased aggregate formation during refolding does not allow a full thermodynamic analysis, and we therefore only note the midpoints of the unfolding transitions as an indication of protein stability (Table 3). All proteins analyzed displayed a similar three-state unfolding pathway, and in agreement with the T_m values (Table 3), the single tryptophan mutations did not significantly alter the thermodynamic stability of native M and Z α_1AT. All proteins analyzed displayed a first transition centered around 2 M urea and a second transition centered around 5 M urea, suggesting that also the stability of the intermediate ensemble remained unaffected by the single tryptophan mutations.

Analysis of the local conformational changes during equilibrium unfolding of M and Z α_1AT

To obtain information on the site-specific conformational changes that occur in the environment of Trp194 and Trp238 during the unfolding of M and Z α_1AT, we followed the changes in the center of spectral mass (COSM) of the single tryptophan mutants as a function of urea concentration (Fig. 5 and Table 1). The unfolding of M α_1AT$_{W194}$ monitored by fluorescence shows a three-state transition with an intermediate ensemble populated at a similar urea concentration as the ensemble detected by far-UV CD (Fig. 4 and Table 3). The first transition is centered around 2.1 ± 0.02 M urea, and the second transition, around 5.3 ± 0.07 M urea. The unfolding of M α_1AT$_{W238}$ does not show the formation of the intermediate species as the protein unfolds via a single transition with a midpoint of 6.3 ± 0.12 M urea.

Table 2. Inhibitory properties of the single tryptophan mutants

	SI		k_{assapp}[a] ($\times 10^6$ M^{-1} s^{-1})		k_{ass}[b] ($\times 10^6$ M^{-1} s^{-1})	
	Chymotrypsin	HNE	Chymotrypsin	HNE	Chymotrypsin	HNE
M α_1AT	1.0 ± 0.1	1.0 ± 0.03	1.26 ± 0.17	6.35 ± 0.18	1.26	6.35
M α_1AT$_{W194}$	1.3 ± 0.2	1.1 ± 0.06	0.97 ± 0.04	5.41 ± 0.41	1.26	5.95
M α_1AT$_{W238}$	1.1 ± 0.1	1.1 ± 0.09	1.35 ± 0.09	4.69 ± 0.64	1.49	5.16
Z α_1AT	1.9 ± 0.3	1.7 ± 0.18	0.75 ± 0.05	2.99 ± 0.41	1.43	5.08
Z α_1AT$_{W194}$	4.2 ± 0.2	7.2 ± 0.51	0.42 ± 0.04	1.13 ± 0.07	1.76	8.14
Z α_1AT$_{W238}$	7.8 ± 0.3	10.5 ± 0.3	0.29 ± 0.03	0.56 ± 0.09	2.26	5.88

[a] Calculated using the total concentration of α_1AT.

[b] Calculated using the fractional concentration of α_1AT that forms an inhibitory complex with the proteinase.

Together, these results suggest that Trp194, which is more buried in native M α_1AT than Trp238,[32] becomes more solvent exposed in the folding intermediate, whereas the area around Trp238 remains unchanged in the folding intermediate and only unfolds at higher urea concentrations.

Fluorescence-monitored unfolding of the single tryptophan mutants of Z α_1AT suggests that the Z mutation has no effect on the unfolding profile of the area around Trp238. Z α_1AT$_{W238}$ unfolding shows a two-state transition with a midpoint of 6.4 ± 0.05 M urea, which is superimposable with the unfolding curve of M α_1AT$_{W238}$. Similar to M α_1AT$_{W194}$, the unfolding of Z α_1AT$_{W194}$ shows a three-state unfolding curve with midpoints of 1.8 ± 0.08 and 5.3 ± 0.09 M urea for the first and the second transitions, respectively. As there is no significant difference in the midpoint of the unfolding transitions of M and Z α_1AT$_{W194}$, these data suggest that the Z mutation does not lead to destabilization of β-sheet A in the native or intermediate state. Further, the elevated pre-transition baseline of Z α_1AT$_{W194}$ in comparison to M α_1AT$_{W194}$ supports our findings that the top of β-sheet A is structurally disrupted in native Z α_1AT and suggests that the transition to the folding intermediate involves a more extensive conformational change in this area in M α_1AT than in Z α_1AT. The unfolding curves therefore suggest that the Z mutation leads to the disruption of the structural integrity of β-sheet A but it does not significantly decrease its thermodynamic stability.

Effects of the single tryptophan mutations on the rate of polymerization

The effects of the single tryptophan mutations on the rate of polymerization of M and Z α_1AT were analyzed under conditions that trigger polymer formation. Protein aliquots were incubated at 60 °C, and the loss of monomeric protein was followed using native PAGE (Fig. 6a). The rate of polymerization as a function of monomer loss (k_{agg}) was determined using an exponential decay function (Fig. 6b and Table 3). The results suggest that Z α_1AT ($k_{agg} = 6.8 \pm 0.2 \times 10^{-3}$ s^{-1}) polymerizes approximately 2-fold faster than M α_1AT ($k_{agg} = 3.5 \pm 0.2 \times 10^{-3}$ s^{-1}) under these experimental conditions. Further, the replacement of Trp194 with a phenylalanine had no effect on the polymerization rate of either protein, whereas mutating Trp238 resulted in an increased polymerization rate with k_{agg} values of $7.1 \pm 0.3 \times 10^{-3}$ s^{-1} for M α_1AT$_{W194}$ and $9.7 \pm 0.7 \times 10^{-3}$ s^{-1} for Z α_1AT$_{W194}$.

Discussion

The Z mutation is a single amino acid substitution (Glu342Lys) that leads to increased polymerization of the archetypal serpin α_1AT at its place of

Table 3. Stability and polymerization properties of the single tryptophan mutants

	T_m[a] (°C)	$D_{mN \rightarrow I}$[b] (M)	$D_{mI \rightarrow U}$[b] (M)	k_{agg}[c] ($\times 10^{-3}$ s^{-1})
M α_1AT	57.2 ± 0.2	2.3 ± 0.01	4.7 ± 0.42	3.5 ± 0.2
M α_1AT$_{W194}$	56.5 ± 0.4	2.2 ± 0.02	5.2 ± 0.04	7.1 ± 0.3
M α_1AT$_{W238}$	57.3 ± 0.4	2.1 ± 0.02	5.2 ± 0.39	4.2 ± 0.2
Z α_1AT	56.1 ± 0.6	2.0 ± 0.01	4.9 ± 0.40	6.8 ± 0.2
Z α_1AT$_{W194}$	54.6 ± 0.4	2.2 ± 0.02	5.1 ± 0.14	9.7 ± 0.7
Z α_1AT$_{W238}$	56.3 ± 0.4	2.0 ± 0.06	4.4 ± 0.31	6.4 ± 0.5

[a] The thermal denaturation curves obtained using far-UV CD were analyzed using the two-state unfolding function to determine the midpoint of transition.

[b] The urea-induced unfolding curves obtained using far-UV CD were analyzed using the three-state unfolding function to determine the midpoints of the N→I transition ($D_{mN \rightarrow I}$) and the I→U transition ($D_{mI \rightarrow U}$).

[c] The rate of polymerization as a function of monomer loss was determined using the equation for exponential decay.

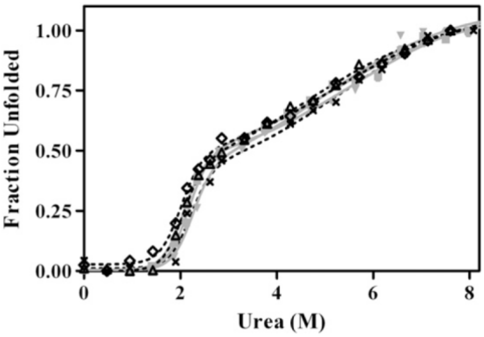


Fig. 4. Far-UV CD unfolding profiles of the single tryptophan mutants of M and Z α$_1$AT. The urea-induced unfolding of M α$_1$AT (▼), M α$_1$AT$_{W194}$ (■), M α$_1$AT$_{W238}$ (●), Z α$_1$AT (◊), Z α$_1$AT$_{W194}$ (×) and Z α$_1$AT$_{W238}$ (△) was followed by changes in the far-UV CD signal at 222 nm. Lines represent three-state curve fits.

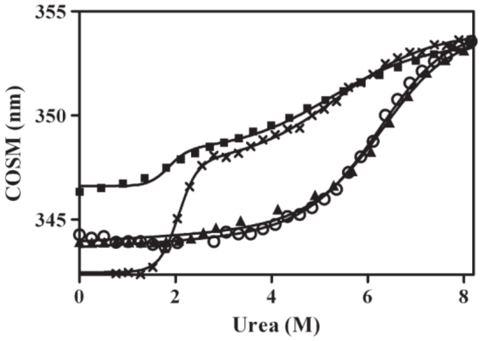


Fig. 5. Intrinsic fluorescence unfolding profiles of the single tryptophan mutants of M and Z α$_1$AT. The urea-induced unfolding of M α$_1$AT$_{W194}$ (×), M α$_1$AT$_{W238}$ (○), Z α$_1$AT$_{W194}$ (■) and Z α$_1$AT$_{W238}$ (▲) was followed by changes in COSM. M and Z α$_1$AT$_{W194}$ unfolding was analyzed using a three-state model, and M and Z α$_1$AT$_{W238}$ unfolding was analyzed using a two-state model (lines).

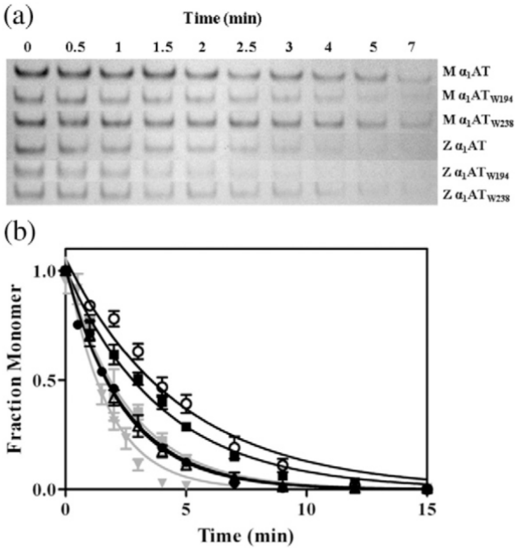


Fig. 6. Polymerization of the single tryptophan mutants of M and Z α$_1$AT. (a) Samples of the single tryptophan mutants (10 μM) were incubated at 60 °C for the time spans indicated and then subjected to native PAGE analysis. (b) The rate of polymerization of M α$_1$AT (○), M α$_1$AT$_{W194}$ (△), M α$_1$AT$_{W238}$ (■), Z α$_1$AT (●), Z α$_1$AT$_{W194}$ (▼) and Z α$_1$AT$_{W238}$ (■) was determined as a function of monomer loss using the equation for exponential decay (lines).

synthesis, the ER of hepatocytes.[5,22,38] It is not clear why the Z mutation leads to an increased polymerization propensity of α$_1$AT, but there is evidence suggesting that Z α$_1$AT can polymerize during folding[18,21] and once the protein has attained its native state.[5,15,18] Controversy exists about how α$_1$AT polymerization occurs,[5,39] but disruption of β-sheet A seems to play a central role, as several mutations in the shutter region of α$_1$AT lead to increased polymerization.[40–42] Disruption of β-sheet A has also been associated with the Z mutation. An exogenous 6-mer peptide homologous to P$_{7\text{-}2}$ of the RCL was found to selectively anneal to Z α$_1$AT in preference to a full-length 12-mer peptide.[19] Assuming that the 6-mer peptide interacts with the lower part of β-sheet A, the authors concluded that the RCL is partially inserted at the top, therefore hindering the insertion of the full-length peptide. Additional evidence for partial RCL insertion in Z α$_1$AT comes from studies with papaya proteinase IV, which cleaves the P$_9$–P$_{10}$ bond of wild type but not of Z α$_1$AT, suggesting that this peptide bond is inaccessible in the pathological mutant.[18] The hypothesis that Z α$_1$AT adopts an alternative conformation to the wild type is further supported by differences observed by CD[17,18] and fluorescence spectroscopy.[17] That this conformational change involves opening of β-sheet A, however, is mainly based on the interactions of Z α$_1$AT with exogenous markers whose exact binding sites are unknown at this stage.

The fact that Z α$_1$AT shows differences in the tryptophan fluorescence emission spectrum and the Stern–Volmer constant for potassium iodide quenching in comparison to the wild type suggests that the conformational change involves the environment of at least one of the two tryptophan residues.[17] α$_1$AT contains two naturally occurring tryptophan residues, Trp194 and Trp238, which are both located at the top of the molecule separated from residue 342 by approximately 5 Å and 21 Å, respectively (Fig. 1). Trp194 is situated at the top of strand three of β-sheet A (s3A), and Trp238 lies on strand two of β-sheet B (s2B). In order to localize the conformational change via intrinsic means and to determine to which extent the environment of each of the two tryptophan residues is affected by the Z mutation, we have created two single tryptophan mutants of Z α_1AT. We have previously replaced each of the two tryptophan residues with phenylalanine in the most common allelic variant of α_1AT, M α_1AT, to obtain information on the local conformational changes that occur during α_1AT folding.[25] We have now introduced the same mutations into Z α_1AT to create Z $\alpha_1\mathrm{AT}_{\mathrm{W194}}$ and Z $\alpha_1\mathrm{AT}_{\mathrm{W238}}$ containing Trp194 and Trp238, respectively.

The single tryptophan mutations did not alter the structure (Fig. 2) or the far-UV CD-monitored unfolding pathway (Fig. 4) of M or Z α_1AT and had little effect on the thermodynamic stability of either protein (Table 3). We therefore conducted a detailed fluorescence spectroscopic analysis of each single tryptophan mutant (Fig. 3 and Table 1). The fluorescence emission maximum and the emission intensity of Trp238 were similar for M and Z α_1AT, suggesting that the environment of this residue is not affected by the Z mutation. This was further confirmed by potassium iodide quenching and steady-state fluorescence anisotropy, which did not show any differences between M and Z α_1AT in the solvent accessibility and the side-chain motion of Trp238, respectively. Significant differences between M and Z α_1AT were, however, detected in the environment of Trp194. The crystal structure of α_1AT suggests that the side chain of Trp194 is buried in the hydrophobic core, whereas Trp238 is located in a partially surface exposed pocket.[32] This is reflected in the approximately 11-nm lower emission maximum, the 9-fold lower Stern–Volmer constant for potassium iodide quenching and the 1.7-fold increase in the steady-state fluorescence anisotropy that were observed for Trp194 in comparison to Trp238 in M α_1AT (Table 1). In contrast, for Z α_1AT, we observed an approximately 10-nm red-shift in the emission maximum of Trp194 compared to M α_1AT, which suggests that the Z mutation induces a conformational change that causes this residue to reside in a more polar environment. This red-shift is accompanied by an approximately 5-fold increase in the Stern–Volmer constant for potassium iodide quenching and a 1.5-fold decrease in the anisotropy, which are indicative of an increase in the solvent exposure and the rotational freedom of Trp194, respectively (Table 1). Taking the location of the two tryptophan residues into consideration, our fluorescence data therefore strongly suggest that the Z mutation induces opening of the top of β-sheet A but does not affect the conformation of β-sheet B.

These findings are further supported by the fluorescence unfolding curves of the single tryptophan mutants (Fig. 5). We have previously shown that both M and Z α_1AT unfold via the formation of one intermediate ensemble populated under similar denaturant concentrations[17] and that β-sheet A is disrupted, whereas β-sheet B is intact in the folding intermediate of M α_1AT.[25] In the current study, we could now show that, similar to the wild type, the environment of Trp238 is intact in the folding intermediate of Z α_1AT (Fig. 5). No difference between M and Z α_1AT could be detected in the unfolding of the environment of Trp238, which suggests that the Z mutation has no effect on the structural and thermodynamic integrity of β-sheet B. As expected from the fluorescence analysis of native M and Z α_1AT (Fig. 3 and Table 1), the unfolding of the area around Trp194 during the transition to the intermediate state proceeds from an already disrupted β-sheet A in Z α_1AT (Fig. 5). The midpoints of this unfolding transition are, however, similar for M and Z α_1AT, which suggests that, despite leading to a conformational rearrangement of β-sheet A, the Z mutation does not alter the thermodynamic stability of this area. This is consistent with previously published data that show that the Z mutation does not affect the global thermodynamic stability of α_1AT.[17,21] The simplest explanation for this is that any loss of interactions between amino acid residues that is caused by the Z mutation is presumably balanced by new interactions formed in the altered conformation.

Further, our data suggest that the Z mutation interferes with the inhibitory function of α_1AT by affecting the translocation process of the proteinase after docking to α_1AT and that this effect is enhanced by replacing Trp194 or Trp238 with a phenylalanine (Table 2). Several regions in α_1AT have been identified that play an important role in controlling the conformational change that occurs during proteinase inhibition.[11–13] One of these regions is the breach region, which is the area at the top of β-sheet A where the RCL first inserts after cleavage by the proteinase.[13] The breach region is characterized by a high degree of sequence conservation, which suggests that the hydrogen bond network in this area is particularly important for the inhibitory process.[11] All residues mutated in this study (Trp194, Trp238 and Glu342) are located in the breach region, but the detrimental effect of the single tryptophan mutations on the function of α_1AT only occurs on the background of the Z mutation. The fact that the single tryptophan mutations have no effect on the activity of M α_1AT but are detrimental to the activity of Z α_1AT suggests that the interactions in the breach region are already compromised by the Z mutation so that this area cannot accommodate the disruptions caused by the single tryptophan mutations without a significant loss in activity. Intriguingly, we did not observe a difference in the second-order rate constants k_{ass} for inhibition of two proteinases by Z α_1AT and its single tryptophan mutants in comparison to M α_1AT and its mutants, which one might expect if the RCL was partially inserted in the pathological variant as previously proposed.[18,19]

The effects of the single tryptophan mutations on the polymerization rate of M and Z α_1AT were determined using native PAGE (Fig. 6 and Table 3). Whereas mutating Trp194 had no effect on the rate of polymer formation of either α_1AT variant, the replacement of Trp238 with a phenylalanine led to approximately 2-fold and 1.4-fold increases in the polymerization rate of M and Z α_1AT, respectively. We have previously discussed the importance of the environment of Trp238 for α_1AT folding and proposed that, together with helix G and helix H, it might serve as a folding nucleus.[25] Considering that polymerization seems to occur as an off-pathway reaction during α_1AT folding, it is therefore not surprising that the disruption of interactions in this area leads to an increased polymerization rate.

In conclusion, our results clearly demonstrate that the Z mutation leads to a structural rearrangement in the top of β-sheet A of α_1AT that is most likely stabilized by new interactions and, hence, does not lead to a change in thermodynamic stability.

Materials and Methods

Materials

Ultrapure-grade urea was purchased from Sigma, and bis-ANS was obtained from Invitrogen.

Production of the recombinant α_1AT variants

The Blunt vectors (Invitrogen) carrying the M or Z α_1AT gene[15] were used as templates. For creation of the single tryptophan mutants, Trp194 and Trp238 were replaced individually with a phenylalanine residue using the QuikChange site-directed mutagenesis approach (Stratagene), and all mutations were verified with DNA sequencing. The different mutants were then ligated into the pHILD2 vector (Invitrogen) for expression. The α_1AT variants were expressed in *Pichia pastoris* and purified as described previously.[15] The proteins were greater than 95% pure and monomeric as determined by size-exclusion chromatography.

Spectroscopic methods

Fluorescence emission spectra were recorded at room temperature on a PerkinElmer LS50B luminescence spectrometer in a 1-cm-path-length quartz cell. For tryptophan emission spectra (290–500 nm), the samples were excited at 290 nm, and a scan speed of 50 nm/min was applied. Unless otherwise noted, the excitation and emission slit widths were set at 5 nm, and the protein concentration was 1 μM in 50 mM Tris and 90 mM NaCl, pH 8.0. bis-ANS emission spectra (400–600 nm) were obtained in the presence of 5 μM bis-ANS at an excitation wavelength of 390 nm. Unless otherwise noted, CD measurements were performed on a Jasco J-815 CD spectrometer at 20 °C in a 0.1-cm-path-length quartz cell. Far-UV CD spectra (250–190 nm) were recorded at a protein concentration of 0.25 mg/mL in 50 mM Tris and 90 mM NaCl, pH 8.0, using a data pitch of 0.1 nm and a scan speed of 100 nm/min. Θ_{222} measurements were made with the signal averaged over 15 s.

Fluorescence quenching experiments

Fluorescence quenching measurements were performed in 50 mM Tris and 90 mM NaCl, pH 8.0, at a protein concentration of 2 μM. Aliquots of a 2-M KI stock containing 10 mM $Na_2S_2O_3$ were added to the protein samples, and the changes in fluorescence emission intensity of the tryptophan residues ($\lambda_{em}=330$ nm) were measured after each step. KCl was used to correct for any salt-induced effects. The excitation slit width was set at 5 nm, and the emission slit width, at 10 nm. The protein solution had an absorbance at the excitation wavelength of lower than 0.05 after the final addition of KI. Fluorescence values were therefore only corrected for the dilution factors. The quenching data were analyzed as described previously.[43]

Fluorescence anisotropy

Anisotropy measurements were performed in a 1-cm-path-length quartz cell at 20 °C with a protein concentration of 2 μM in 90 mM NaCl and 50 mM Tris, pH 8.0. The excitation and emission slit widths were 5 nm, and an integration time of 0.1 s was applied. The anisotropy was measured with an excitation at 290 nm and emission between 330 and 400 nm using the L-format method[44] incorporating *G*-factor correction.

Characterization of the inhibitory properties

The SI and the association rate constant (k_{ass}) against bovine chymotrypsin and HNE were determined as described previously.[15]

Thermal denaturation

Thermal denaturation measurements were performed in a 0.1-cm-path-length quartz cell at a protein concentration of 0.25 mg/mL in 90 mM NaCl and 50 mM Tris, pH 8.0. A heating rate of 15 °C/h was applied, and the change in signal at 222 nm in the far-UV spectra was determined. The thermal denaturation data were fit to a two-state unfolding model using a nonlinear least-squares fitting algorithm as described previously to obtain the midpoint of denaturation (T_m).[45]

Chemical denaturation

Stock solutions of urea were prepared in 90 mM NaCl and 50 mM Tris, pH 8.0, and filtered through 0.22-μm membranes. The urea concentration was determined by refractive index measurements as described previously.[46] Equilibrium unfolding curves were obtained by adding a concentrated solution of native protein in 90 mM NaCl and 50 mM Tris, pH 8.0, to a series of urea concentrations. The samples were incubated for 2 h at room temperature before analysis by measuring either the fluorescence emission spectra or the change in signal at 222 nm in the far-UV spectra as a function of denaturant concentration. The fluorescence emission spectra were recorded to determine the change in COSM wavelength between 310 and 400 nm during equilibrium unfolding.[47] The final protein concentration in the fluorescence unfolding curves was 1.5 μM with the excitation slit width set at 5 nm and the emission slit width set at 10 nm. The final protein concentration in the far-UV CD unfolding curves was 0.2 mg/mL, and measurements were performed on a Jasco 810 spectropolarimeter. The unfolding data were fit to either a two-state or a three-state unfolding model using a nonlinear least-squares fitting algorithm as described previously.[25] The three-state unfolding analysis recognizes the presence of one stable intermediate structure (I) populated during the transition from the native (N) to the unfolded (U) state.

Determination of the rate of polymerization

The rate of polymerization was determined using continuous native PAGE analysis according to the Mini-PROTEAN® 3 Cell Instruction Manual (Bio-Rad). Samples of protein (10 μM) were incubated at 60 °C in 90 mM NaCl and 50 mM Tris, pH 8.0, and put on ice at various time points to quench the reaction. Ice-cold non-denaturing sample buffer was added, and samples were separated at 4 °C on 6% continuous native gels. The rate of polymerization was determined by following the loss of monomer using an exponential decay function.

Acknowledgements

This work was supported by the National Health and Medical Research Council (of which S.P.B. is a Senior Research Fellow) through a program grant. We thank the Protein Production Unit at Monash University for their help in protein purification and members of the Bottomley laboratory for helpful discussions.

References

1. Fregonese, L. & Stolk, J. (2008). Hereditary alpha-1-antitrypsin deficiency and its clinical consequences. *Orphanet J. Rare Dis.* **3**, 16.
2. Gadek, J. E., Fells, G. A., Zimmerman, R. L., Rennard, S. I. & Crystal, R. G. (1981). Antielastases of the human alveolar structures. Implications for the protease–antiprotease theory of emphysema. *J. Clin. Invest.* **68**, 889–898.
3. Knaupp, A. S. & Bottomley, S. P. (2009). Serpin polymerization and its role in disease—the molecular basis of alpha1-antitrypsin deficiency. *IUBMB Life*, **61**, 1–5.
4. Stoller, J. K. & Aboussouan, L. S. (2005). Alpha1-antitrypsin deficiency. *Lancet*, **365**, 2225–2236.
5. Lomas, D. A., Evans, D. L., Finch, J. T. & Carrell, R. W. (1992). The mechanism of Z alpha 1-antitrypsin accumulation in the liver. *Nature*, **357**, 605–607.
6. Lomas, D. A. & Parfrey, H. (2004). Alpha1-antitrypsin deficiency. 4: Molecular pathophysiology. *Thorax*, **59**, 529–535.
7. Elliott, P. R., Abrahams, J. P. & Lomas, D. A. (1998). Wild-type alpha 1-antitrypsin is in the canonical inhibitory conformation. *J. Mol. Biol.* **275**, 419–425.
8. Huntington, J. A., Read, R. J. & Carrell, R. W. (2000). Structure of a serpin–protease complex shows inhibition by deformation. *Nature*, **407**, 923–926.
9. Stratikos, E. & Gettins, P. G. (1999). Formation of the covalent serpin–proteinase complex involves translocation of the proteinase by more than 70 Å and full insertion of the reactive center loop into beta-sheet A. *Proc. Natl Acad. Sci. USA*, **96**, 4808–4813.
10. Tew, D. J. & Bottomley, S. P. (2001). Intrinsic fluorescence changes and rapid kinetics of proteinase deformation during serpin inhibition. *FEBS Lett.* **494**, 30–33.
11. Irving, J. A., Pike, R. N., Lesk, A. M. & Whisstock, J. C. (2000). Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. *Genome Res.* **10**, 1845–1864.
12. Stein, P. E. & Carrell, R. W. (1995). What do dysfunctional serpins tell us about molecular mobility and disease? *Nat. Struct. Biol.* **2**, 96–113.
13. Whisstock, J. C., Skinner, R., Carrell, R. W. & Lesk, A. M. (2000). Conformational changes in serpins: I. The native and cleaved conformations of alpha(1)-antitrypsin. *J. Mol. Biol.* **296**, 685–699.
14. Loebermann, H., Tokuoka, R., Deisenhofer, J. & Huber, R. (1984). Human alpha 1-proteinase inhibitor. Crystal structure analysis of two crystal modifications, molecular model and preliminary analysis of the implications for function. *J. Mol. Biol.* **177**, 531–557.
15. Levina, V., Dai, W., Knaupp, A. S., Kaiserman, D., Pearce, M. C., Cabrita, L. D. *et al.* (2009). Expression, purification and characterization of recombinant Z alpha(1)-antitrypsin—the most common cause of alpha(1)-antitrypsin deficiency. *Protein Expression Purif.* **68**, 226–232.
16. Ogushi, F., Fells, G. A., Hubbard, R. C., Straus, S. D. & Crystal, R. G. (1987). Z-type alpha 1-antitrypsin is less competent than M1-type alpha 1-antitrypsin as an inhibitor of neutrophil elastase. *J. Clin. Invest.* **80**, 1366–1374.
17. Knaupp, A. S., Levina, V., Robertson, A. L., Pearce, M. C. & Bottomley, S. P. (2010). Kinetic instability of the serpin Z alpha1-antitrypsin promotes aggregation. *J. Mol. Biol.* **396**, 375–383.
18. Lomas, D. A., Evans, D. L., Stone, S. R., Chang, W. S. & Carrell, R. W. (1993). Effect of the Z mutation on the physical and inhibitory properties of alpha 1-antitrypsin. *Biochemistry*, **32**, 500–508.
19. Mahadeva, R., Dafforn, T. R., Carrell, R. W. & Lomas, D. A. (2002). 6-mer peptide selectively anneals to a

pathogenic serpin conformation and blocks polymerization. Implications for the prevention of Z alpha(1)-antitrypsin-related cirrhosis. *J. Biol. Chem.* **277**, 6771–6774.

20. Koloczek, H., Guz, A. & Kaszycki, P. (1996). Fluorescence-detected polymerization kinetics of human alpha 1-antitrypsin. *J. Protein Chem.* **15**, 447–454.
21. Yu, M. H., Lee, K. N. & Kim, J. (1995). The Z type variation of human alpha 1-antitrypsin causes a protein folding defect. *Nat. Struct. Biol.* **2**, 363–367.
22. Cox, D. W., Billingsley, G. D. & Callahan, J. W. (1986). Aggregation of plasma Z type alpha 1-antitrypsin suggests basic defect for the deficiency. *FEBS Lett.* **205**, 255–260.
23. Janciauskiene, S., Dominaitiene, R., Sternby, N. H., Piitulainen, E. & Eriksson, S. (2002). Detection of circulating and endothelial cell polymers of Z and wild type alpha 1-antitrypsin by a monoclonal antibody. *J. Biol. Chem.* **277**, 26540–26546.
24. Elliott, P. R., Bilton, D. & Lomas, D. A. (1998). Lung polymers in Z alpha1-antitrypsin deficiency-related emphysema. *Am. J. Respir. Cell Mol. Biol.* **18**, 670–674.
25. Tew, D. J. & Bottomley, S. P. (2001). Probing the equilibrium denaturation of the serpin alpha(1)-antitrypsin with single tryptophan mutants; evidence for structure in the urea unfolded state. *J. Mol. Biol.* **313**, 1161–1169.
26. Compton, L. A. & Johnson, W. C., Jr (1986). Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal. Biochem.* **155**, 155–167.
27. Manavalan, P. & Johnson, W. C., Jr (1987). Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem.* **167**, 76–85.
28. Sreerama, N. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260.
29. Whitmore, L. & Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32**, W668–W673.
30. Whitmore, L. & Wallace, B. A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers*, **89**, 392–400.
31. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
32. Elliott, P. R., Pei, X. Y., Dafforn, T. R. & Lomas, D. A. (2000). Topography of a 2.0 Å structure of alpha1-antitrypsin reveals targets for rational drug design to prevent conformational disease. *Protein Sci.* **9**, 1274–1281.
33. Cabrita, L. D., Dai, W. & Bottomley, S. P. (2004). Different conformational changes within the F-helix occur during serpin folding, polymerization, and proteinase inhibition. *Biochemistry*, **43**, 9834–9839.
34. Cabrita, L. D., Whisstock, J. C. & Bottomley, S. P. (2002). Probing the role of the F-helix in serpin stability through a single tryptophan substitution. *Biochemistry*, **41**, 4575–4581.
35. James, E. L., Whisstock, J. C., Gore, M. G. & Bottomley, S. P. (1999). Probing the unfolding pathway of alpha1-antitrypsin. *J. Biol. Chem.* **274**, 9482–9488.
36. Pearce, M. C., Rubin, H. & Bottomley, S. P. (2000). Conformational change and intermediates in the unfolding of alpha 1-antichymotrypsin. *J. Biol. Chem.* **275**, 28513–28518.
37. Powell, L. M. & Pain, R. H. (1992). Effects of glycosylation on the folding and stability of human, recombinant and cleaved alpha 1-antitrypsin. *J. Mol. Biol.* **224**, 241–252.
38. Le, A., Ferrell, G. A., Dishon, D. S., Le, Q. Q. & Sifers, R. N. (1992). Soluble aggregates of the human PiZ alpha 1-antitrypsin variant are degraded within the endoplasmic reticulum by a mechanism sensitive to inhibitors of protein synthesis. *J. Biol. Chem.* **267**, 1072–1080.
39. Yamasaki, M., Li, W., Johnson, D. J. & Huntington, J. A. (2008). Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature*, **455**, 1255–1258.
40. Dafforn, T. R., Mahadeva, R., Elliott, P. R., Sivasothy, P. & Lomas, D. A. (1999). A kinetic mechanism for the polymerization of alpha1-antitrypsin. *J. Biol. Chem.* **274**, 9548–9555.
41. Lomas, D. A., Finch, J. T., Seyama, K., Nukiwa, T. & Carrell, R. W. (1993). Alpha 1-antitrypsin Siiyama (Ser53 → Phe). Further evidence for intracellular loop-sheet polymerization. *J. Biol. Chem.* **268**, 15333–15335.
42. Zhou, A., Stein, P. E., Huntington, J. A. & Carrell, R. W. (2003). Serpin polymerization is prevented by a hydrogen bond network that is centered on His-334 and stabilized by glycerol. *J. Biol. Chem.* **278**, 15116–15122.
43. Lehrer, S. S. (1971). Solute perturbation of protein fluorescence. The quenching of the tryptophyl fluorescence of model compounds and of lysozyme by iodide ion. *Biochemistry*, **10**, 3254–3263.
44. Lakowicz, J. R. (2006). *Principles of Fluorescence Spectroscopy*, 3rd edit Springer, New York, NY.
45. Dafforn, T. R., Pike, R. N. & Bottomley, S. P. (2004). Physical characterization of serpin conformations. *Methods*, **32**, 150–158.
46. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
47. Cabrita, L. D., Gilis, D., Robertson, A. L., Dehouck, Y., Rooman, M. & Bottomley, S. P. (2007). Enhancing the stability and solubility of TEV protease using in silico design. *Protein Sci.* **16**, 2360–2367.

Chapter 5

Conformational Properties of the Disease-Causing Z Variant of α_1-Antitrypsin Revealed by Theory and Experiment

Declaration for Thesis Chapter 5

Declaration by candidate

In the case of Chapter 5, the nature and extent of my contribution to the work was the following:

Nature of contribution	**Extent of contribution (%)**
Experimental design for *in vitro* experiments, protein construction and preparation, *in vitro* experiments, *in vitro* data analysis and manuscript preparation.	40

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

Name	**Nature of contribution**	**Extent of contribution (%) for student co-authors only**
Itamar Kass	Molecular dynamics simulations and manuscript preparation.	
Stephen P. Bottomley	Design of project and manuscript preparation.	
Ashley M. Buckle	Design of project and manuscript preparation.	

Candidate's Signature: ....[redacted].............. **Date:** 29.5.2012

Anja S. Knaupp

Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.

(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;

(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

(4) there are no other authors of the publication according to these criteria;

(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and

(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location: Department of Biochemistry and Molecular Biology, Monash University

[redacted]

Signature: .... **Date:** 29/05/2012
Itamar Kass

Signature: ....[redacted]................... **Date:** ...24/5/12...
Stephen P. Bottomley

[redacted]

Signature: .... **Date:** 24/5/2012
Ashley M. Buckle

2856 Biophysical Journal Volume 102 June 2012 2856–2865

Conformational Properties of the Disease-Causing Z Variant of α1-Antitrypsin Revealed by Theory and Experiment

Itamar Kass,[△] Anja S. Knaupp,[△] Stephen P. Bottomley, and Ashley M. Buckle*
Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, Australia

ABSTRACT The human serine protease inhibitor (serpin) α-1 antitrypsin (α1-AT) protects tissues from proteases of inflammatory cells. The most common disease-causing mutation in α1-AT is the Z-mutation (E342K) that results in an increased propensity of α1-AT to polymerize in the ER of hepatocytes, leading to a lack of secretion into the circulation. The structural consequences of this mutation, however, remain elusive. We report a comparative molecular dynamics investigation of the native states of wild-type and Z α1-AT, revealing a striking contrast between their structures and dynamics in the breach region at the top of β-sheet A, which is closed in the wild-type simulations but open in the Z form. Our findings are consistent with experimental observations, notably the increased solvent exposure of buried residues in the breach region in Z, as well as polymerization via domain swapping, whereby the reactive center loop is rapidly inserted into an open A-sheet before proper folding of the C-terminal β-strands, allowing C-terminal domain swapping with a neighboring molecule. Taken together, our experimental and simulation data imply that mutations at residue 342 that either stabilize an open form of the top of β-sheet A or increase the local flexibility in this region, may favor polymerization and hence aggregation.

INTRODUCTION

Alpha 1-antitrypsin (α1-AT) belongs to the serine protease inhibitor superfamily. The native fold of α1-AT is composed of nine α-helices and three β-sheets and an extended reactive center loop (RCL) (Fig. 1 *A*) (1). The native serpin fold is a kinetically trapped metastable conformation that uses the extended RCL as bait for serine proteases. Upon protease binding the RCL is cleaved and the serpin undergoes a major irreversible structural rearrangement, where the cleaved RCL is inserted between strands 3 and 5 of the A β-sheet (2). As a result, the active site of the bound protease is critically distorted, leaving it inactive and trapped (2). Although this mechanism allows a high degree of regulation of serpin activity, it renders members of this superfamily vulnerable to misfolding and aggregation (3).

α1-AT is synthesized mainly by hepatocytes (4) and protects lungs cells from excessive elastase activity (5). Several disease-causing mutations of α1-AT have been identified, the most common of them being the Z-variant (E342K) located at the breach region, at the top of the A β-sheet (6) (Fig. 1 *A*). The Z-variant has an increased propensity to polymerize in the endoplasmic reticulum of hepatocytes leading to cell death and liver damage (6). In addition, α1-AT aggregation causes a reduction in the overall inhibitory capabilities of the plasma that results in unhindered elastase activity and subsequently emphysema (7). Importantly, ~15% of Z α1-AT does fold correctly and is secreted from the hepatocytes. This Z α1-AT, unlike the WT form, polymerizes under physiological conditions and pathological Z α1-AT polymers have been detected in circulation (8) and in the lung (9). Understanding the molecular mechanisms of serpin polymerization has been an intensely active area of investigation over the past 15 years. Studies have shown that α1-AT polymerization can form a heterogeneous mixture of polymers (10), is temperature-dependent (6), pH-dependent (11), and concentration-dependent (12), and can be blocked by RCL homologous peptides (6,12,13) as well as citrate ions (14). Although the structure of Z α1-AT and its associated polymer are not known, recent structural and biochemical data suggest that polymerization occurs via a domain-swapping event (6,15).

Unfortunately, the structural mechanism by which the mutation E342K facilitates α1-AT polymerization remains unknown. One possible mechanism is that the stability of the already folded protein is lost, leading to a conformational change that results in a polymerogenic species (11,16,17). However, it was shown that although Z α1-AT polymerized faster than wild-type α1-AT, it is as thermodynamically stable as the wild-type (19). Observations that in vivo polymerization of α1-AT occurs mainly in the ER suggest that folding-intermediates, not the native state, play the dominant role in polymerization. Due to the intrinsic short lifespan of the species involved in folding and polymerization, experimental methods such as crystallography are limited. Therefore, theoretical methods such as molecular dynamics (MD) are necessary to understand, at an atomic level, the conformational differences between wild-type α1-AT and the Z-variant. In this article, we use MD simulations and fluorescence spectroscopy to investigate the structural and dynamical differences among wild-type, E342K, E342Q, and E342R mutated α1-AT.

Submitted November 20, 2011, and accepted for publication May 16, 2012.

[△] Itamar Kass and Anja S. Knaupp contributed equally to this work.

*Correspondence: ashley.buckle@monash.edu

Editor: Michael Feig.

© 2012 by the Biophysical Society

0006-3495/12/06/2856/10 $2.00

doi: 10.1016/j.bpj.2012.05.023

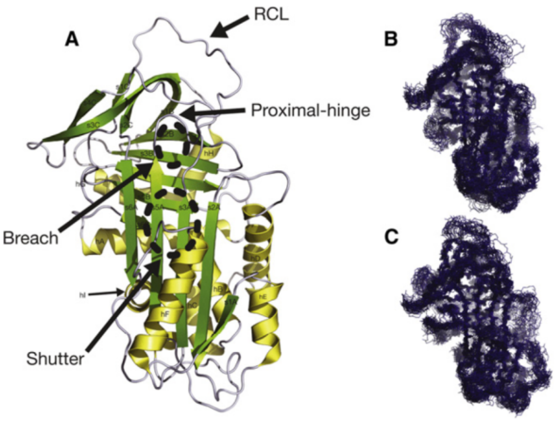


FIGURE 1 α1-AT structure and representative overviews of the dynamics of different simulated systems. (*A*) The crystal structure of wild-type α1-AT (PDB 1QLP (1)) indicating the nomenclature of the helices and strands. The positions of the breach, shutter, proximal-hinge, and RCL are indicated. (*B* and *C*) Superposition of 30 structures sampled every 10 ns from a 300-ns simulation of wild-type and E342K mutated α1-AT, respectively.

MATERIALS AND METHODS

System setup

Six systems were simulated: wild-type, E342K, E342Q, E342R, and K290E/E342K mutated α1-AT in solution, and wild-type α1-AT in a crystalline-like environment. The atomic coordinates of wild-type α1-AT were taken from the crystal structure of Elliott et al. (PDB:1QLP (1)). E342K, E342R, E342Q, and K290E/E342K mutated α1-AT atomic coordinates were achieved by in silico mutation of wild-type using PyMOL version 1.3r2 (20). Two models of E342K were built, differing in K342 side chain conformation: E342K_conf1 in which the K342 side chain is pointing toward T203, and E342K_conf2 in which the side chain points directly to solvent. Ionizable residues were in their standard protonation states at pH 7.

Each system was first put into a suitably sized box, of which the minimal distance from the protein to the box wall was 1.4 nm, except in the case of wild-type α1-AT in a crystalline-like environment where no minimal distance was used. This was done to mimic the crystal lattice interactions of 1QLP. Each box was then subjected to energy-minimization before commencing dynamics. To avoid unnecessary distortion of the protein when the molecular dynamics simulation is started, a gradual positional-restraints procedure was used. Three consecutive equilibration runs in which all heavy protein atoms are restrained to their starting positions (using a force constant of 1000, 100, and 10 kJ mol^{-1} nm^{-2} respectively) while the water is relaxing around the structure. Each system was then subjected to free simulation with configurations stored every 100 ps for analysis.

Simulation parameters

All simulations were performed using the GROMACS package Ver. 4.0.7 (21) in conjunction with the GROMOS 53A6 united-atom force field (22). Water was represented explicitly using the simple-point-charge model (23). Each system was simulated under periodic conditions in a cubical box. Simulation conditions were roughly the same as described in Oostenbrink et al. (22). Briefly, the temperature of the simulation systems in solution was maintained by independently coupling the protein and the solvent to an external temperature bath at 298 Kelvin (K) with a coupling constant of $\tau_T = 0.1$ps using a Berendsen thermostat (24). The pressure was maintained at 1 bar by weakly coupling the system to an isotropic pressure bath (24), using an isothermal compressibility of 4.6×10^{-5} bar^{-1} and a coupling constant of $\tau_P = 1$ ps. During the simulations, the length of all bonds within the protein was constrained using the LINCS algorithm (25). The time step for integrating the equations of motion was 2 fs. Nonbonded interactions were evaluated using a twin-range cut-off scheme: interactions falling within the 0.8-nm short-range cutoff were calculated every step whereas interactions within the 1.4-nm long cutoff were updated every three steps, together with the pair list. A reaction-field correction was applied to the electrostatic interactions beyond the long-range cutoff (26), using a relative dielectric permittivity constant of $\varepsilon_{RF} = 62$ as appropriate for simple-point-charge water (27). Wild-type α1-AT and mutants were simulated three times independently, each system starting with a different distribution of initial velocities.

Simulations of wild-type α1-AT under conditions mimicking the crystalline environment were performed at 100 K, i.e., the temperature of x-ray data collection for PDB:1QLP, and under constant pressure. The temperature of the system was maintained by independently coupling the protein and the solvent to an external temperature bath at 100 K with a coupling constant of $\tau_T = 0.1$ps using a Berendsen thermostat (24). Long-range electrostatics were calculated using the particle-mesh Ewald (28), with a 1.4-nm cutoff for the direct space calculation. These conditions, although different to those used for parameterizing the force field, were used to mimic the conditions under which x-ray data were collected. Specifically, we used particle-mesh Ewald for the calculation of electrostatic interactions, as this is known to induce order in small simulation boxes (29,30).

Root mean-square deviation

The position deviations of each Cα atom with respect to its initial structure (as described in System Setup, above) were calculated every 100 ps, after performing a least-squares fit to its initial structure.

H-bonds

H-bonds were calculated as a function of time. Two atoms are being defined as having an H-bond if the distance between the donor and the acceptor atoms is ≤ 0.35 nm and the acceptor-donor-hydrogen angle is $< 30°$.

Electrostatic surface potential calculations

The electrostatic potentials of models were calculated using APBS Ver. 1.3 (31). Atom parameters for the calculation were taken from the GROMOS 53A6 force field (22). Electrostatic potential was visualized using the PyMOL Ver. 1.3r2 (20) with positive potential in blue and negative potential in red in a range between -1 and $+1$ k_BT/e_c, where k_b is the Boltzmann constant, T is the temperature (set to 298°K), and e_c is electron charge.

Materials

4,4′-Dianilino-1,1′-binaphthyl-5,5′-disulfonic acid, dipotassium salt (bis-ANS) was purchased from Invitrogen (Carlsbad, CA).

Production of recombinant α1-AT mutants

The α1-AT mutants were generated using the pHILD2 vector (Invitrogen) carrying the α1-AT gene as a template (32). To introduce the single amino-acid substitution, the E342 residue was mutated with KOD DNA polymerase (Novagen, Madison, WI) using the Quik-Change site-directed mutagenesis approach (Stratagene, La Jolla, CA) and all mutations were verified with DNA sequencing. The α1-AT variants generated were electroporated into *Pichia pastoris* and expressed and purified as described previously in Levina et al. (32).

Characterization of inhibitory properties

The stoichiometry of inhibition (SI) of the α1-AT variants against bovine chymotrypsin was determined as described previously in Levina et al. (32).

Spectroscopic methods

All spectra were recorded at room temperature on a FluoroMax-4 spectrofluorometer (HORIBA Jobin-Yvon, Edison, NJ) in a 1-cm path-length quartz cell. For tryptophan emission, spectra samples were excited at 295 nm and spectra were collected from 300 to 500 nm (increment 0.5 nm). Excitation and emission slit widths were set at 5 nm and the integration time at 0.1 s. The protein concentration was 2 μM in 50 mM Tris and 90 mM NaCl, pH 8.0. The absorbance at the excitation wavelength remained below 0.05. For bis-ANS emission spectra, samples were excited at 390 nm and spectra were collected from 400 to 600 nm (increment 0.5 nm). Excitation and emission slit widths were set at 3 nm and the integration time at 0.1 s. The protein concentration was 1 μM and the bis-ANS concentration 5 μM in 50 mM Tris and 90 mM NaCl, pH 8.0.

Determination of the rate of polymerization

The rate of polymerization was determined by native polyacrylamide gel electrophoresis (PAGE). Samples of protein (10 μM) were incubated at 42°C in 90 mM NaCl and 50 mM Tris, 1 mM EDTA, and 5 mM β-mercaptoethanol at pH 8.0 and put on ice at various time-points to quench the reaction. Ice-cold nondenaturing sample buffer was added and samples were analyzed using 10% nondenaturing polyacrylamide gel electrophoresis. The rate of polymerization was determined by following the loss of monomer using an exponential decay function.

RESULTS

Wild-type and Z α1-AT are stable throughout simulations

Molecular simulations of both wild-type and modeled E342K α1-AT were performed. The wild-type α1-AT simulation started from the crystal structure of α1-AT in the native state (PDB:1QLP), whereas the simulation of Z started from the E342K mutation modeled into the wild-type α1-AT 1QLP coordinates, in two conformations (see below). For each system, three independent simulations of 300 ns were performed. In all cases, the majority of secondary structure, including helices hA and hB, β-sheets A and most of B (except s1B), remained stable throughout the simulation (Fig. 1, *B* and *C*; and see Movie S1 and Movie S2 in the Supporting Material). In contrast, hF, the top of hD as well as the RCL, all showed a high degree of flexibility, consistent with high crystallographic B factors (1) and susceptibility to H/D exchange (33).

Calculated root mean-square deviations (RMSDs) (see Fig. S1 in the Supporting Material) indicate system equilibration times ranging from 25 to 40 ns, with the shifts in Cα values reaching a plateau within the productive stages of the simulations (typically after 50 ns of simulations); the average Cα RMSDs were found to stabilize at 0.37 ± 0.02 nm for wild-type α1-AT, and at 0.43 ± 0.04 nm and 0.41 ± 0.04 nm for E342K-mutated α1-AT starting conformations 1 and 2, respectively. The RMSD values and inspection of snapshots, suggests that despite small shifts, both wild-type and the Z α1-AT are stable during simulations, and that the E342K mutation does not affect the overall stability, conformation, and dynamics (Fig. 1, *B* and *C*). The RCL in wild-type and Z α1-AT shows a high degree of movement during simulations (Fig. 2). This is consistent

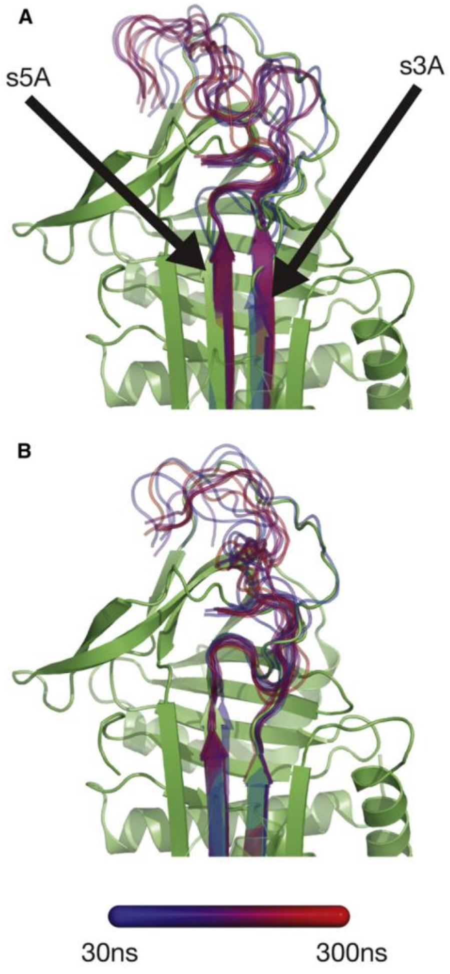


FIGURE 2 Representative structure of β-sheet A and proximal-hinge of (*A*) wild-type and (*B*) E342K mutated α1-AT. Superposition of snapshots taken every 30 ns from 300-ns simulations aligned to the crystal structure of wild-type α1-AT (*green*, PDB 1QLP (1)), shown as a cartoon representation. For clarity, only s3A, s5A, and the RCL of the snapshots are shown, superposed on the crystal structure. The color coding is indicated in the legend, as a gradient going from early (*blue*) to late (*red*), except for the initial structure (*green*).

with RCL conformational diversity and high mobility observed in crystallographic structures (1,34,35).

The starting conformation of K342 does not affect the dynamics of E342K mutated α1-AT

Two different starting side chain conformations of K342 were modeled to prevent sampling bias from the starting configuration (see Fig. S3). In E342K_{conf1}, the side chain of K342 forms an H-bond with the side chain of T203 and the backbone carbonyl of V200. In contrast, in E342K_{conf2}, the side chain of K342 was modeled to extend out toward solvent, having no H-bond interactions with the protein. Nevertheless, after <5 ns of simulation, the conformation of K342 in E342K_{conf2} shifts toward that in E342K_{conf1} (where K342 forms stable H-bonds with T203 and V200; see Fig. S3). Due to the conformational convergence of E342K_{conf1} and E342K_{conf2}, both systems share a high degree of similarity in structure and dynamics throughout simulations.

The top of β-sheet A zips up in wild-type α1-AT but not in the Z variant

Comparison of the structure at the top of β-sheet A for both wild-type α1-AT and the Z variant during simulations reveals that wild-type and Z α1-AT differ strikingly in the number and length of H-bonds between s3A and s5A, the expulsion of the proximal-hinge from the top of β-sheet A, and the conformation of the RCL (Figs. 2 and 3). In the early stages (20–40 ns) of the wild-type simulations the proximal-hinge (residues 342–347 of the RCL;P_{17}-P_{12}) is expelled from the top of β-sheet A, followed by a zipping-up of β-strands s3A and s5A at the top of β-sheet A, which occurs between 30 and 50 ns (Figs. 2 *A* and 3 *B*). As such, two new H-bonds are formed, between the backbone atoms of T339 to G192 and G192 to D341 (Fig. 3 *B*; and see Table S2 in the Supporting Material).

In contrast, the conformation of the proximal-hinge region in Z α1-AT is found to be stable throughout simulation, with no structural change observed at the top of β-sheet A (Figs. 2 *B* and 3 *C*). Thus, whereas during the simulation of wild-type α1-AT the top of β-sheet A zips up, it remains partially open during the simulation of Z α1-AT.

The breach region is partially solvent-exposed in Z but not in wild-type α1-AT

The breach region, the area at the top of β-sheet A where the RCL first inserts after cleavage by the protease, contains several highly conserved residues, in particular W194 that is present in >95% of all serpins (36). We observed differences in the solvent exposure of W194 side chain during wild-type and Z α1-AT simulations (see Fig. S4). As a result of opening the top of the A-sheet, the solvent-exposed surface area of W194 during the last 100 ns of the simulations was found to be 0.44 ± 0.15 nm^2 and 0.78 ± 0.18 nm^2 for wild-type and Z α1-AT, respectively. In contrast, the solvent exposure of W238, which is situated on β-sheet B, is similar for both proteins (1.21 ± 0.18 nm^2 and 1.20 ± 0.20 nm^2, respectively). In the crystal structure of wild-type α1-AT, the side chain of W194 forms a H-bond with the backbone carbonyl of D341, but closure of the top of sheet-A in the breach region early in the simulation of wild-type α1-AT results in the loss of this interaction, with the W194 backbone amine forming a H-bond with the carbonyl of D341 (see Table S1 and Fig. 4 *A*). In contrast, in Z α1-AT, the H-bond between D341° to $W194^{N\varepsilon1}$ is maintained during the productive stage of simulations (Fig. 4 *B*). These observations are consistent with recent experimental data that indicate conformational differences at the top of β-sheet A between wild-type and Z α1-AT, in addition to increased solvent exposure and polarity of surroundings for W194 in Z compared to wild-type α1-AT (37).

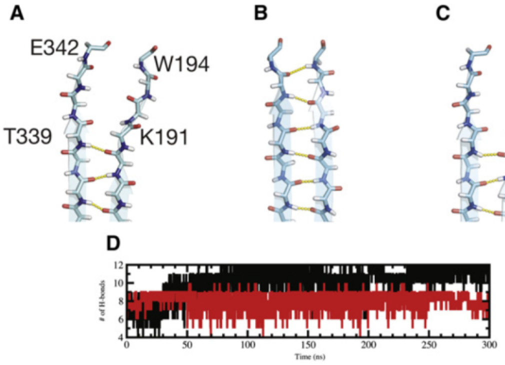


FIGURE 3 Interstrand H-bonds between strands s3A and s5A during simulations. Interstrand H-bonds between strands s3A and s5A of (*A*) crystal structure, and representative time-averaged structure (over the last 50 ns of simulation) of (*B*) wild-type and (*C*) E342K-mutated α1-AT. Differences in H-bond patterns can be seen, as wild-type α1-AT has H-bond interactions between T339 → G192, G192 → D341, and D341 → W194 that are formed during simulations. (*D*) Number of H-bonds between s3A and s5A main-chain atoms as a function of simulation time for wild-type and E342K-mutated α1-AT (*black* and *red*, respectively). The average number of H-bonds for the last 200 ns of the simulations between s3A and s5A is 10 ± 1 and 8 ± 1 for wild-type and E342K mutated α1-AT, respectively.

Electrostatic interactions of K342 favor a partially open conformation at the top of β-sheet A in Z α1-AT

In simulations of wild-type α1-AT, E342 maintains a conformation similar to that seen in the crystal structure—being stabilized by a salt-bridge with $K290^{NZ}$, which is also present in the crystal structure (Fig. 4 *A*). The H-bond between E342 and T203, present in the crystal structure, is not maintained during the simulations. In contrast, in

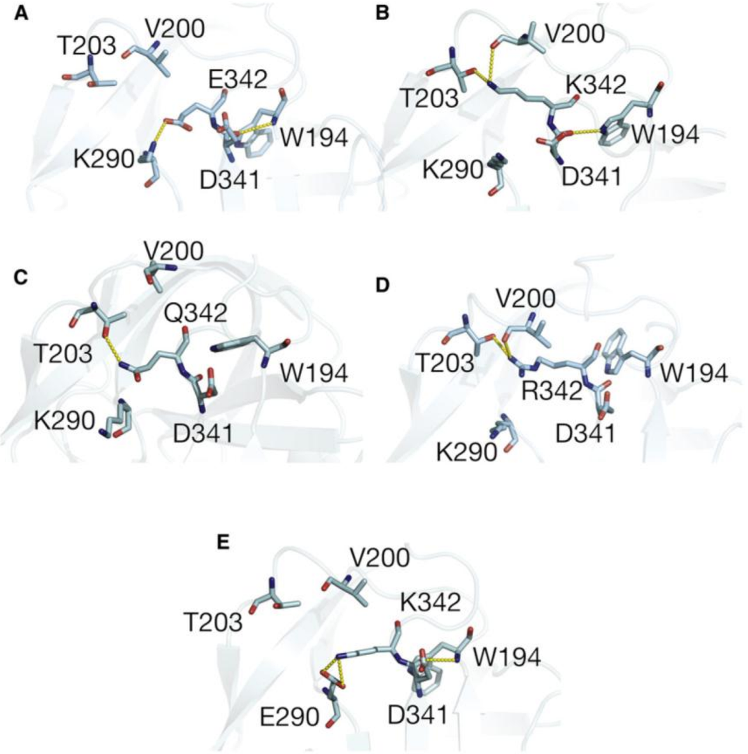


FIGURE 4 Representative conformations of residue 342 of (*A*) wild-type, (*B*) E342K, (*C*) E342Q, (*D*) E342R, and (*E*) K290E/E342K-mutated α1-AT. Structures are time-averaged over the last 50 ns of the simulation. Whereas residues E342 and Q342 (*A* and *C*) formed a salt-bridge with K290, K342 and R342 formed H-bonds with T203 and V200 (*B* and *D*). In addition, a salt-bridge between D341 and W194 is found in E342K-mutated α1-AT but not in the wild-type protein.

simulations of Z α1-AT K342 that is electrostatically repelled by the positive potential of this region, particularly K290 (Fig. S5 *B*), adopts a different conformation in which its side chain forms H-bonds with the backbone of T203 and V200. As a result, the difference in conformations between Glu and Lys at position 342 contribute to the dynamic and structural differences at the top of β-sheet A. Specifically, the conformation of K342 is dictated by the surrounding electropositive potential and stabilized by H-bonds to T203 and V200, thus preventing the expulsion of the RCL hinge-region from the top of β-sheet A. In contrast, in the wild-type structure, E342 can adopt a conformation that allows the expulsion of the proximal-hinge and the closure of the top of β-sheet A.

Stereochemical and physicochemical properties of residues at position 342 dictate the conformation at the top of β-sheet A in α1-AT

To investigate the role of electrostatics, polarity, and the physicochemical nature of residue 342, we performed MD studies of both E342Q and E342R mutations modeled into α1-AT. These mutations were deemed the most stereochemically conservative substitutions while allowing us to alter the chemical properties of the side chain. Both mutations introduce a positive electrostatic potential similar to that in the E342K Z-variant (see Fig. S5, *C* and *D*) and thus their proximity to K290 is energetically unfavorable. Both E342Q and E342R were found to be stable during simulations, with average Cα RMSDs of 0.39 ± 0.02 nm and 0.37 ± 0.02 nm, respectively, after 50 ns of simulation (see Fig. S1). However, in contrast to Z α1-AT we found that E342Q and E342R adopt two distinct conformations throughout simulations. Initially, the proximal-hinge of both proteins is expelled after 40–60 ns of simulation (see Fig. S6, *A* and *B*) and the top of β-sheet A in both cases closes (see Fig. S2 and Table S2). Thus, although the side chains of Arg and Gln at position 342 adopt a conformation similar to that of K342, they allow the expulsion of the proximal-hinge and the closure of the top of β-sheet A. However, during simulations both E342Q (150–170 ns) and E342R (180–200 ns) undergo a local conformational change, resulting in a small degree of opening of the top of β-sheet A (see Fig. S2 and Table S2). This conformation is stabilized by H-bond interactions between the side chains of Q342 and R342 to T203$^{N\gamma}$ and V200°, as seen in E342K (Fig. 4, *D* and *E*; and see Movie S3 and Movie S4).

As a result of this slight opening of the top of the A-sheet, the solvent-exposed surface area of W194 during the last 100 ns of simulations was found to be 0.66 ± 0.17 nm^2

and 0.61 ± 0.17 nm^2 for E342Q and E342R α1-AT, respectively. In contrast, the solvent exposure of W238 is similar to that of WT α1-AT and the Z-variant (1.19 ± 0.18 nm^2 and 1.22 ± 0.20 nm^2, respectively). Thus, simulations show that both E342Q and E342R can adopt open and closed conformations at the top of β-sheet A. The transition between conformations, seen in E342Q and E342R simulations, but not for Z α1-AT, is likely due to differences in the properties of Arg and Gln side chains. Although the side chains of Arg and Lys are similar, the Arg side chain harbors a bulky guanidinium moiety that may slow the rate of its structural rearrangement, allowing for some plasticity at the top of β-sheet A (see Movie S3). In contrast, the side chain of Gln is shorter compared to that of Lys, having an amide instead of an amino group. As a result, throughout simulations its side chain adopts two conformations—one in which it is close to K290 and another in which it is close to T203 (Fig. 4 *C*; and see Movie S4).

Taken together, these observations imply that removing the negative charge at position 342 (E342Q) or even replacing it with a positive one (E342R) does affect the final conformation of β-sheet A. However, whereas E342R forms stable interactions with T203 and V200, in a similar manner to E342K, E342Q does not. Rather, it seems to alternate between a conformation in which its side chain is near K290 and a conformation in which it interacts with V200.

The Z reversal mutation K290E/E342K reveals wild-type-like dynamics

The K290E mutation was shown to rescue the Z-variant of α1-AT, maintaining a wild-type-like secretion level from cells (38). Because our simulations indicate that electrostatic and H-bond interactions play an important role in the conformation and dynamics at the top of β-sheet A, we were interested to see whether MD could provide insight into how a second mutation can repair the deleterious effects of the Z mutation. We therefore modeled the double-mutant K290E/E342K and performed MD simulation. K290E/E342K was found to be stable during simulations, with average Cα RMSDs of 0.42 ± 0.02 nm after 50 ns of simulations (see Fig. S1). During simulations, K342 maintained a salt-bridge with E290, as seen in wild-type simulations (Fig. 4 *E*). As a result, the double mutant is dynamically similar to wild-type α1-AT (see Fig. S6 *C* and Movie S5), with its proximal-hinge being expelled after 20–40 ns of simulation. Unlike the Z variant, no H-bonds are seen between K290 and T203 or V200. These results indicate that the rescue mutation K290E restores a salt-bridge between position 290 and 342 of α1-AT, allowing the top of β-sheet A to zip up, as seen in the wild-type simulations. As a result of the closed conformation at the top of the A-sheet, the solvent-exposed surface area of W194 during the last 100 ns of simulations was found to be 0.50 ± 0.10 nm^2, in contrast to 1.21 ± 0.19 nm^2 for W238.

Effect of crystal packing on the conformation of the RCL and the top of β-sheet A

The effect of crystal packing on protein structure has been studied widely by comparing identical molecules crystallized in different forms (39–44). In general, it was observed that whereas more loosely organized segments of secondary structure might be affected, the main protein core is largely invariant. However, in a few cases such as T4 lysozyme, large structural changes have been reported (39). The temperature at which an experiment is performed can also affect the structure. A recent comparison of crystal structures determined at room temperature and 100 K (45) indicate that cryo-cooling can affect the energy-landscape of molecules, resulting in smaller, overpacked models. In addition, cryo-cooling significantly reduces or even eliminates motions important for protein function. We were therefore interested in whether crystal packing or temperature were related to the conformational changes we observed in our simulations.

We ran MD simulations of wild-type α1-AT in a crystalline-like environment at 100 K, simulating the crystal packing forces and the cryogenic temperatures of 1QLP. In contrast to the simulations of wild-type α1-AT in solution, wild-type α1-AT under crystal conditions displays significantly less flexibility (see Fig. S6 *D*). This is particularly apparent in the RCL, s4C-s3C turn, hD-s2A turn, and hF, all of which are in crystal packing regions (see Fig. S7). This is consistent with experimental hydrogen exchange data that show that the flexibility of hD and hF higher in solution than that predicted based on the crystal structure (33).

Interestingly, in the crystalline simulation the proximal hinge is not expelled from the top of β-sheet A, as seen in the simulations in solution. This suggests that the proximal hinge conformation observed in the crystal structure of 1QLP is stabilized by crystal packing and probably also related to the cryogenic temperature of data collection. In addition, it provides an explanation for the structural changes in this region observed early in the simulation of wild-type α1-AT.

The stoichiometry of inhibition of both E342Q and E342R mutated α1-AT is higher than that of wild-type α1-AT

Wild-type, Z variant (E342K), E342Q-, and E342R-mutated α1-AT were expressed in *P. pastoris* and purified as described previously (32). The stoichiometry of inhibition (SI) of each α1-AT variant was determined against bovine chymotrypsin (Table 1). All α1-AT variants were functional inhibitors of this proteinase, however, all mutants had elevated SI values compared to the wild-type. It has previously been shown that the Z mutation results in an increased substrate behavior with ~40% following the substrate

TABLE 1 Stoichiometry of inhibition (SI) and polymerization properties of wild-type α1-AT and mutants

α1-AT variant	SI	k_{agg} (h^{-1})
wt	1.0 ± 0.1	—
E342K	1.8 ± 0.3	$4.8 \pm 0.3 \times 10^{-2}$
E342Q	1.4 ± 0.1	$2.3 \pm 0.2 \times 10^{-2}$
E342R	1.7 ± 0.1	$4.3 \pm 0.3 \times 10^{-2}$

pathway (32,46,47). Our data suggest that both E342Q and E342R mutations have a similar effect on α1-AT, increasing the SI from 1.0 ± 0.1 for wild-type α1-AT to 1.4 ± 0.1 and 1.7 ± 0.1, respectively. This indicates that mutations at position 342 decrease the rate of RCL insertion into the A-sheet after proteinase docking to α1-AT, thereby leading to an increase in substrate behavior. Both the Z variant (SI = 1.8 ± 0.3) and E342R mutated α1-AT result in elevated SI values compared to E342Q mutated α1-AT. This suggests that whereas the loss of salt-bridge interactions with K290 impairs the inhibitory process, a positively charged amino acid at position 342 imposes greater effects.

Fluorescence analysis of E342Q and E342R mutated α1-AT

To determine the effect of the E342Q and E342R mutations on the structural integrity of α1-AT, intrinsic and extrinsic fluorescence emission scans were performed. Previous work on Z-mutated α1-AT has shown that the molecule is characterized by an increase in tryptophan fluorescence emission intensity and a small red shift in the wavelength of maximum fluorescence intensity (λ_{max}) in comparison to the wild-type (19). Additional work utilizing single tryptophan mutants could then attribute these spectral differences to a conformational change around W194 caused by the E342K mutation (37). Our data indicate that both the E342Q and E342R mutations analyzed in this study led to an increase in the tryptophan emission intensity in comparison to wild-type α1-AT (Fig. 5 *A*). The data indicate that the local environment of W194 in both mutants differs from that of wild-type α1-AT. A red shift of λ_{max} was, however, only observed if residue 342 was mutated to a positively charged residue. Taken together, the data indicate that in both E342Q and E342R mutated α1-AT, W194 is more solvent-exposed compared to wild-type α1-AT. This is consistent with our simulation data, which show some plasticity at the top of the A-sheet.

To obtain information on the hydrophobicity of the α1-AT variants, the fluorescent probe bis-ANS was used (Fig. 5 *B*). It has previously been shown that Z α1-AT is characterized by an increased exposure of hydrophobic regions, resulting in an increase in bis-ANS fluorescence emission intensity (19). Our data suggest that both E342Q and E342R mutations lead to an increase in bis-ANS fluorescence emission intensity in comparison to wild-type α1-AT. Similar to the tryptophan emission scans, the bis-ANS emission scans suggest slight differences between the E342Q variants and the E342R and the Z α1-AT variants, with the highest bis-ANS fluorescence intensity observed in the presence of the latter two.

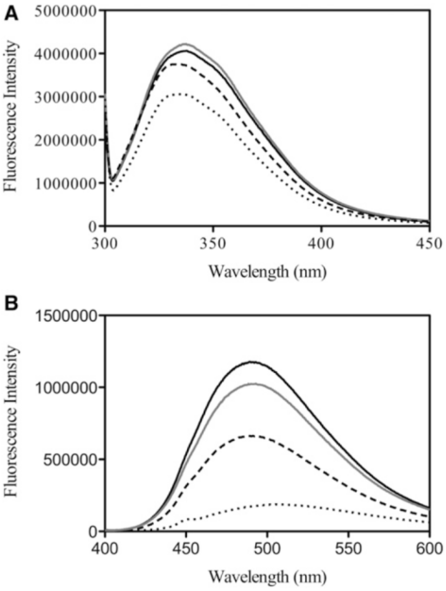


FIGURE 5 Fluorescence emission spectra of wild-type and mutated α1-AT. (*A*) Tryptophan emission spectra (λ_{ex} = 295 nm) of wild-type (*dotted line*), E342K (*solid line*), E342Q (*dashed line*), and E342R (*shaded line*) mutated α1-AT. (*B*) bis-ANS emission spectra (λ_{ex} = 390 nm) in the presence of wild-type (*dotted line*), E342K (*solid line*), E342Q (*dashed line*), and E342R (*shaded line*) -mutated α1-AT.

Both E342Q and E342R mutated α1-AT polymerize faster than wild-type α1-AT

The effect of the E342K, E342Q, and E342R mutations on the rate of α1-AT polymerization was determined at physiological temperatures. Protein samples were incubated at 42°C and the loss of monomeric protein was followed by native PAGE (Fig. 6 *A*). The rate of polymerization as a function of monomer loss (k_{agg}) was determined using an exponential decay function (Fig. 6 *B*; Table 1). Whereas wild-type α1-AT did not polymerize to a significant extent under those conditions, all of the variants did. Differences in k_{agg} were observed between the mutants, with Z α1-AT ($k_{agg} = 4.8 \pm 0.3 \times 10^{-2}$ h^{-1}) and the E342R ($k_{agg} = 4.3 \pm 0.3 \times 10^{-2}$ h^{-1}) variant polymerizing at approximately double the rate of the E342Q ($2.3 \pm 0.2 \times 10^{-2}$ h^{-1}) variant. This suggests that it is not solely the loss of the salt-bridge that results in the increased polymerization propensity of Z variant α1-AT but also the positive charge at position 342. This is in good agreement with previous work that determined the levels of α1-AT secretion upon mutation of residue 342 in a mammalian cell system (48). This study suggests that the E342Q mutation results in secretion levels of 55% of normal and the E342K in 17% of normal.

DISCUSSION

Z α1-AT is the most common pathological variant of α1-AT (49,50), and differs from the wild-type by only one residue (E342K) located in the breach region. The breach region is important for controlling the conformational change that occurs during protease inhibition (11,16,17), and is thus characterized by a high degree of sequence conservation across the serpin superfamily (36). As such, it is likely that the hydrogen-bond network in this area is particularly important for the inhibitory process (16). Because E342 is located at the top of s5A and the base of the RCL, it is one of these highly conserved residues, forming a salt-bridge to K290 and a H-bond to T203 (16,33). Mutation to K at this position is thus predicted to have significant structural consequences that lie at the cause of the pathological phenotype of Z α1AT.

However, the nature of the structural response to the Z mutation and hence the precise mechanism by which it promotes α1-AT polymerization is unclear (15,16). To investigate the effect of the Z mutation on the structure and dynamics of α1-AT, particularly within the breach region at the top of β-sheet A, we have performed a comparison between the dynamics and structure of wild-type and Z α1-AT using long MD simulations. In a previous theoretical study, done by Jezierski and Pasenkiewicz-Gierula (51), the native fold of both wild-type and E342K mutated α1-AT was modeled, followed by a short simulated annealing protocol. Based on the models, the authors conclude that the mutation introduces an overall instability into the protein. Our study is distinguished from the previous work because our simulations are based on the experimentally determined crystal structure of α1-AT, are significantly longer in timescale (allowing equilibrium to be reached), and include several mutations at position 342, allowing us to explore the role of stereochemistry and physicochemistry at this important location in the protein. Importantly, we also use fluorescence spectroscopy to characterize the properties of variants with mutations at this position.

The simulation data indicate that, in contrast to wild-type α1-AT the presence of the mutation E342K causes the top of β-sheet A to remain in a partially open configuration. Whereas in wild-type α1-AT the proximal-hinge of the RCL is expelled from the top of β-sheet A, the proximal-hinge of E342K α1-AT is partially inserted into the top of β-sheet A. Our findings are consistent with the recent characterization of Z α1-AT using intrinsic fluorescence spectroscopy that showed that W194, a residue in the breach region, is buried in wild-type α1-AT but has increased solvent exposure in the Z form (46).

These observations suggest that the Z mutation induces a conformational change in the region around W194 that leads to W194 being situated in a more polar environment. Further, unfolding data suggest that the Z mutation leads to the disruption of the structural integrity of β-sheet A but does not decrease its thermodynamic stability (19). This is consistent with the net maintenance of hydrogen bonds in Z compared to wild-type α1-AT that we observe during our simulations. Although wild-type α1-AT has two more H-bonds compared with Z α1-AT (at the top of β-sheet A) as well as a salt-bridge between K290 and E342, the Z-form compensates by forming three new H-bonds between V200, T203 to K342, and between D341 and the side chain of T194. This indicates that despite structural differences between wild-type α1-AT and the Z-form, both proteins have a similar number of H-bonds and salt-bridges, and as a result have similar thermodynamic stabilities, consistent with recent fluorescence spectroscopy studies (19,37).

Recently a mechanism for polymerization via domain swapping was proposed based upon crystallographic and biophysical studies, whereby the RCL is rapidly inserted into a conformationally open/labile β-sheet A before proper folding of the C-terminal β-strands, allowing C-terminal domain swapping with a neighboring molecule (15). Our key observation that the Z mutation stabilizes an open conformation at the top of β-sheet A, might indicate that this may also be a structural feature of folding intermediates, thus favoring polymerization.

Our fluorescence and polymerization data for E342Q and E342R mutated α1-AT, show clearly that they behave in a similar fashion to the Z variant. However, our MD results for these mutants predict that their characteristics should be more similar to those of the wild-type protein. Close inspection of the MD trajectories reveals that while the top of the A-sheet does zip up to a large degree in both mutants, unlike the wild-type some plasticity is present, with some side-chain flexibility at position 342. This structural heterogeneity may be sufficient to allow rapid insertion of the RCL into a conformationally open β-sheet A, resulting in polymerization. Taken together, our simulation and experimental results imply that mutations at residue 342 that either stabilize an open form of the top of β-sheet A (E342K) or increase the local flexibility in this region (E342Q and E342R), favor polymerization (Fig. 6).

Our simulations of α1-AT in a crystal-like environment indicate that crystal packing forces, in addition to cryogenic temperatures, may influence the specific conformation of various regions, and suggest that the structure of native wild-type α1-AT in solution may differ from that seen for the crystal form (at least in the case of 1QLP). Simulation in solution indicates that some structural reorganization may occur, specifically zipping up of the top of β-sheet A. These finding are in agreement with biophysical data for wild-type α1-AT in solution that suggest that the protein

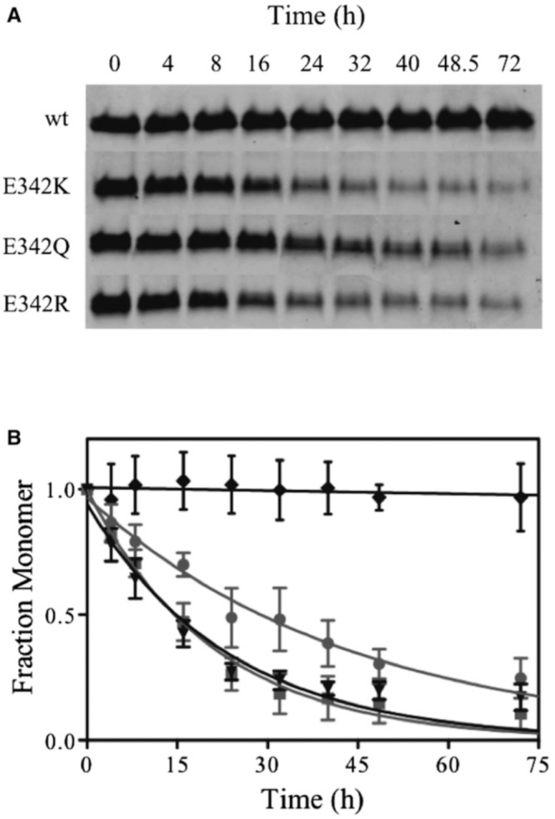


FIGURE 6 Polymerization properties of wild-type and mutated $\alpha 1$-AT. (*A*) Protein samples (10 μM) were incubated at 42°C for the time spans indicated and then stored on ice until subjected to native PAGE analysis. (*B*) The rate of polymerization of wild-type (*dash-diamond-dash*), E342K (*dash-box-dash*), E342Q (*dash-circle-dash*), and E342R (*dash-down arrow-dash*) -mutated $\alpha 1$-AT was determined as a function of monomer loss using the equation for exponential decay (*lines*).

populates more closed conformations in solution (12,19,46). One explanation for this behavior is that the wild-type protein samples a range of conformations in the breach region, perhaps partitioning between the open and closed forms of β-strands s3A and s5A at the top of β-sheet A, and the crystal lattice stabilizes the open form due to the partial insertion of the RCL and the resulting conformation of the RCL that are favorable for crystal packing. Indeed, inspection of the crystal lattice environment in wild-type $\alpha 1$-AT (1QLP) reveals lattice contacts involving s3C-loop-s4c and the RCL that may serve to stabilize this RCL conformation (Fig. S7).

CONCLUSIONS

Our investigations identify specific interactions in the breach region of the Z-variant of $\alpha 1$-AT that may play a role in its greater propensity to misfold, aggregate, and thus cause disease, and are consistent with biophysical and structural data that highlight the importance of this region in serpin folding and polymerization.

SUPPORTING MATERIAL

Two tables, seven figures, and five movies are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(12)00576-0.

We thank the Monash eResearch center and the Victorian Life Sciences Computation Initiative for computational resources and assistance.

This work was supported by grants from the National Health and Medical Research Council (NHMRC) and the Australian Research Council. A.M.B. and S.P.B. were supported by an NHMRC senior research fellowship.

REFERENCES

1. Elliott, P. R., X. Y. Pei, …, D. A. Lomas. 2000. Topography of a 2.0 Å structure of $\alpha 1$-antitrypsin reveals targets for rational drug design to prevent conformational disease. *Protein Sci.* 9:1274–1281.
2. Huntington, J. A., R. J. Read, and R. W. Carrell. 2000. Structure of a serpin-protease complex shows inhibition by deformation. *Nature.* 407:923–926.
3. Cabrita, L. D., and S. P. Bottomley. 2004. How do proteins avoid becoming too stable? Biophysical studies into metastable proteins. *Eur. Biophys. J.* 33:83–88.
4. Rogers, J., N. Kalsheker, …, S. E. Humphries. 1983. The isolation of a clone for human $\alpha 1$-antitrypsin and the detection of $\alpha 1$-antitrypsin in mRNA from liver and leukocytes. *Biochem. Biophys. Res. Commun.* 116:375–382.
5. Crystal, R. G. 1989. The $\alpha 1$-antitrypsin gene and its deficiency states. *Trends Genet.* 5:411–417.
6. Lomas, D. A., D. L. Evans, …, R. W. Carrell. 1992. The mechanism of Z $\alpha 1$-antitrypsin accumulation in the liver. *Nature.* 357:605–607.
7. Gadek, J. E., G. A. Fells, …, R. G. Crystal. 1984. Role of connective tissue proteases in the pathogenesis of chronic inflammatory lung disease. *Environ. Health Perspect.* 55:297–306.
8. Cox, D. W., G. D. Billingsley, and J. W. Callahan. 1986. Aggregation of plasma Z type $\alpha 1$-antitrypsin suggests basic defect for the deficiency. *FEBS Lett.* 205:255–260.
9. Elliott, P. R., D. Bilton, and D. A. Lomas. 1998. Lung polymers in Z $\alpha 1$-antitrypsin deficiency-related emphysema. *Am. J. Respir. Cell Mol. Biol.* 18:670–674.
10. Ekeowa, U. I., J. Freeke, …, D. A. Lomas. 2010. Defining the mechanism of polymerization in the serpinopathies. *Proc. Natl. Acad. Sci. USA.* 107:17146–17151.
11. Devlin, G. L., M. K. Chow, …, S. P. Bottomley. 2002. Acid denaturation of $\alpha 1$-antitrypsin: characterization of a novel mechanism of serpin polymerization. *J. Mol. Biol.* 324:859–870.
12. Lomas, D. A., D. L. Evans, …, R. W. Carrell. 1993. Effect of the Z mutation on the physical and inhibitory properties of $\alpha 1$-antitrypsin. *Biochemistry.* 32:500–508.
13. Mahadeva, R., T. R. Dafforn, …, D. A. Lomas. 2002. 6-mer peptide selectively anneals to a pathogenic serpin conformation and blocks polymerization. Implications for the prevention of Z $\alpha(1)$-antitrypsin-related cirrhosis. *J. Biol. Chem.* 277:6771–6774.
14. Mast, A. E., J. J. Enghild, and G. Salvesen. 1992. Conformation of the reactive site loop of $\alpha 1$-proteinase inhibitor probed by limited proteolysis. *Biochemistry.* 31:2720–2728.
15. Yamasaki, M., T. J. Sendall, …, J. A. Huntington. 2011. Molecular basis of $\alpha 1$-antitrypsin deficiency revealed by the structure of a domain-swapped trimer. *EMBO Rep.* 12:1011–1017.

16. James, E. L., and S. P. Bottomley. 1998. The mechanism of α1-antitrypsin polymerization probed by fluorescence spectroscopy. *Arch. Biochem. Biophys.* 356:296–300.
17. Dafforn, T. R., R. Mahadeva, …, D. A. Lomas. 1999. A kinetic mechanism for the polymerization of α1-antitrypsin. *J. Biol. Chem.* 274:9548–9555.
18. Reference deleted in proof.
19. Knaupp, A. S., V. Levina, …, S. P. Bottomley. 2010. Kinetic instability of the serpin Z α1-antitrypsin promotes aggregation. *J. Mol. Biol.* 396:375–383.
20. Schrodinger, LLC. 2010. The PyMOL Molecular Graphics System, Ver. 1.3r2. http://sourceforge.net/projects/pymol/files/.
21. Hess, B., C. Kutzner, …, E. Lindahl. 2008. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4:435–447.
22. Oostenbrink, C., A. Villa, …, W. F. van Gunsteren. 2004. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* 25:1656–1676.
23. Berendsen, H. J. C., J. P. M. Postma, and W. F. van Gunsteren. 1981. Interaction models for water in relation to protein hydration. *In* Intermolecular Forces. B. Pullman, editor. Reidel, Dordrecht, The Netherlands. 331–342.
24. Berendsen, H. J. C., J. P. M. Postma, …, J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.
25. Hess, B., H. Bekker, …, J. G. E. M. Fraaije. 1997. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18:1463–1472.
26. Tironi, I. G., R. Sperb, …, W. F. Vangunsteren. 1995. A generalized reaction field method for molecular-dynamics simulations. *J. Chem. Phys.* 102:5451–5459.
27. Heinz, T. N., W. F. van Gunsteren, and P. H. Hunenberger. 2001. Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *J. Chem. Phys.* 115:1125–1136.
28. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald—an *N*.log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.
29. Lins, R. D., and U. Rothlisberger. 2006. Influence of long-range electrostatic treatments on the folding of the N-terminal H4 histone tail peptide. *J. Chem. Theory Comput.* 2:246–250.
30. Figueirido, F., G. S. Del Buono, and R. M. Levy. 1995. On finite-size effects in computer simulations using the Ewald potential. *J. Chem. Phys.* 103:6133–6142.
31. Baker, N. A., D. Sept, …, J. A. McCammon. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA.* 98:10037–10041.
32. Levina, V., W. Dai, …, S. P. Bottomley. 2009. Expression, purification and characterization of recombinant Z α1-antitrypsin—the most common cause of α1-antitrypsin deficiency. *Protein Expr. Purif.* 68:226–232.
33. Tsutsui, Y., L. Liu, …, P. L. Wintrode. 2006. The conformational dynamics of a metastable serpin studied by hydrogen exchange and mass spectrometry. *Biochemistry.* 45:6561–6569.
34. Whisstock, J. C., and S. P. Bottomley. 2006. Molecular gymnastics: serpin structure, folding and misfolding. *Curr. Opin. Struct. Biol.* 16:761–768.
35. Bottomley, S. P. 2011. The structural diversity in α1-antitrypsin misfolding. *EMBO Rep.* 12:983–984.
36. Irving, J. A., R. N. Pike, …, J. C. Whisstock. 2000. Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. *Genome Res.* 10:1845–1864.
37. Knaupp, A. S., and S. P. Bottomley. 2011. Structural change in β-sheet A of Z α1-antitrypsin is responsible for accelerated polymerization and disease. *J. Mol. Biol.* 413:888–898.
38. Brantly, M., M. Courtney, and R. G. Crystal. 1988. Repair of the secretion defect in the Z form of α1-antitrypsin by addition of a second mutation. *Science.* 242:1700–1702.
39. Faber, H. R., and B. W. Matthews. 1990. A mutant T4 lysozyme displays five different crystal conformations. *Nature.* 348:263–266.
40. Ely, K. R., J. N. Herron, …, A. B. Edmundson. 1989. Three-dimensional structure of a light chain dimer crystallized in water. Conformational flexibility of a molecule in two crystal forms. *J. Mol. Biol.* 210:601–615.
41. Moult, J., A. Yonath, …, A. Saya. 1976. The structure of triclinic lysozyme at 2–5 Å resolution. *J. Mol. Biol.* 100:179–195.
42. Eigenbrot, C., M. Randal, and A. A. Kossiakoff. 1992. Structural effects induced by mutagenesis affected by crystal packing factors: the structure of a 30–51 disulfide mutant of basic pancreatic trypsin inhibitor. *Proteins.* 14:75–87.
43. Sheriff, S., W. A. Hendrickson, …, L. H. Jensen. 1985. Influence of solvent accessibility and intermolecular contacts on atomic mobilities in hemerythrins. *Proc. Natl. Acad. Sci. USA.* 82:1104–1107.
44. Phillips, Jr., G. N. 1990. Comparison of the dynamics of myoglobin in different crystal forms. *Biophys. J.* 57:381–383.
45. Fraser, J. S., H. van den Bedem, …, T. Alber. 2011. Accessing protein conformational ensembles using room-temperature x-ray crystallography. *Proc. Natl. Acad. Sci. USA.* 108:16247–16252.
46. Knaupp, A. S., and S. P. Bottomley. 2011. Structural change in β-sheet A of Z α(1)-antitrypsin is responsible for accelerated polymerization and disease. *J. Mol. Biol.* 413:888–898.
47. Ogushi, F., G. Fells, …, R. Crystal. 1986. Z-type α-1-antitrypsin (Aat) is less competent than M1-type Aat as an inhibitor of neutrophil elastase. *Am. Rev. Respir. Dis.* 133:A218.
48. Sifers, R. N., C. P. Hardick, and S. L. C. Woo. 1989. Disruption of the 290–342 salt bridge is not responsible for the secretory defect of the PiZ α1-antitrypsin variant. *J. Biol. Chem.* 264:2997–3001.
49. Fregonese, L., and J. Stolk. 2008. Hereditary α-1-antitrypsin deficiency and its clinical consequences. *Orphanet J. Rare Dis.* 3:16–24.
50. Stoller, J. K., and L. S. Aboussouan. 2005. α1-antitrypsin deficiency. *Lancet.* 365:2225–2236.
51. Jezierski, G., and M. Pasenkiewicz-Gierula. 2001. The effect of the Glu^{342}Lys mutation in α1-antitrypsin on its structure, studied by molecular modeling methods. *Acta Biochim. Pol.* 48:65–75.

Chapter 6

General Discussion

The conformational diseases comprise a set of diverse disorders caused by protein misfolding and subsequent aggregation [41]. The disease-associated protein aggregates are usually highly ordered and can be grouped into two distinct classes: fibrillar and polymeric aggregates. Fibril formation involves the transition of proteins often low in β-sheet content into well-organised filaments characterised by an extensive β-sheet structure [68]. The deposition of fibrillar aggregates is associated with several neurodegenerative disorders including Alzheimer's disease, Parkinson's disease and the spongiform encephalopathies as well as conditions that do not affect the nervous system such as Type II diabetes and dialysis-related amyloidosis [68, 224]. Polymer formation, on the other hand, is a protein misfolding event which does not usually require a significant structural rearrangement of self-associating units [41]. One protein superfamily commonly associated with disease-causing polymerisation events are the serpins (serine proteinase inhibitors). Several members of this superfamily are prone to polymerisation which is linked to a range of diverse disorders including emphysema, liver disease, angioedema, neurodegeneration and thrombosis [41, 68].

The serpins represent the largest class of serine/ cysteine proteinase inhibitors and are involved in several essential physiological processes including the coagulation, inflammatory and fibrinolytic cascades [72, 225]. Members of this superfamily share a characteristic metatstable tertiary fold which undergoes an extensive conformational change during proteinase inhibition [226]. This structural rearrangement is associated with a significant increase in the serpin's stability and is hence a thermodynamically driven process [226]. The metastable native state together with a considerable degree of structural flexibility are therefore essential for a serpin's inhibitory function. These properties, however, render the serpin architecture particularly sensitive to external stressors such as elevated temperatures [143, 154, 198], low denaturant concentrations [142, 143, 195], decreased pH values [153, 175] as well as to mutations [71, 139] that result in misfolding into dysfunctional stabilised states which in most cases are polymeric.

The Z mutation (Glu342Lys) is the most common pathological mutation of the archetypal serpin α_1-antitrypsin (α_1AT), causing disease in approximately 1 in 2000 live births in Northern Europe [163]. It leads to increased α_1AT polymerisation in the ER of hepatocytes, and accordingly in a reduction in the amount of circulating α_1AT. This plasma deficiency eventually results in a loss-of-function disease due to the protease-antiprotease imbalance and the accumulation of polymerised Z α_1AT in the liver can culminate in a gain-of-toxic-function disease [163]. The mechanism by which the Z mutation results in this increased polymerisation propensity of α_1AT and what type of polymer is formed in patients is poorly understood.

The observation that Z α_1AT polymers accumulate predominantly in the ER [142] suggests that polymerisation occurs as an off-pathway event during the α_1AT folding reaction. The α_1AT folding pathway has been studied extensively *in vitro* and is characterised by a reversible three-state transition with the formation of a single intermediate ensemble (I) which is populated in low denaturant concentrations between the native (N) and unfolded (U) states [94, 132, 186-188, 191]. The equilibrium unfolding data presented in this thesis suggest that Z α_1AT also follows this three-state folding reaction and populates I at similar denaturant concentrations as wild-type α_1AT (chapter 3 and chapter 4). Incubation of α_1AT at high protein concentrations and for prolonged periods in these low denaturant concentrations results in polymer formation [142, 143, 195] which suggests that misfolding occurs from this transiently populated species. Under these mildly denaturing conditions wild-type and Z α_1AT seem to polymerise at the same rate [143] whereas this is not the case if the proteins are incubated under physiological conditions [142-144]. This suggests that the effect of the Z mutation on the α_1AT molecule is related to the formation of I. There is evidence that it results in retardation of the second folding transition ($I \rightarrow N$) of α_1AT and consequently accumulation of the polymerogenic folding intermediate I [213]. Yet, α_1AT polymers formed *in vitro* from the folding intermediate I are not recognised by the 2C1 antibody [167, 199], a monoclonal antibody that recognises Z α_1AT polymers isolated from hepatocellular inclusions [185]. In addition to pathological Z α_1AT polymers, the 2C1 antibody interacts with *in vitro* polymers of wild-type and Z α_1AT formed at 60 °C and of Z α_1AT formed at 41 °C [167, 185, 199]. The recent crystal structures of a self-terminating antithrombin dimer [184] and a self-terminating α_1AT trimer [167] together with disulphide-trapping experiments [167] suggest the differences in 2C1 antibody binding may be due to alternative α_1AT polymerisation pathways under different conditions: α_1AT seems to predominantly polymerise via a s4A/s5A swap mechanism in low denaturant concentrations and a C-terminal swap mechanism at elevated temperatures (Figure 6.1) [167, 184].

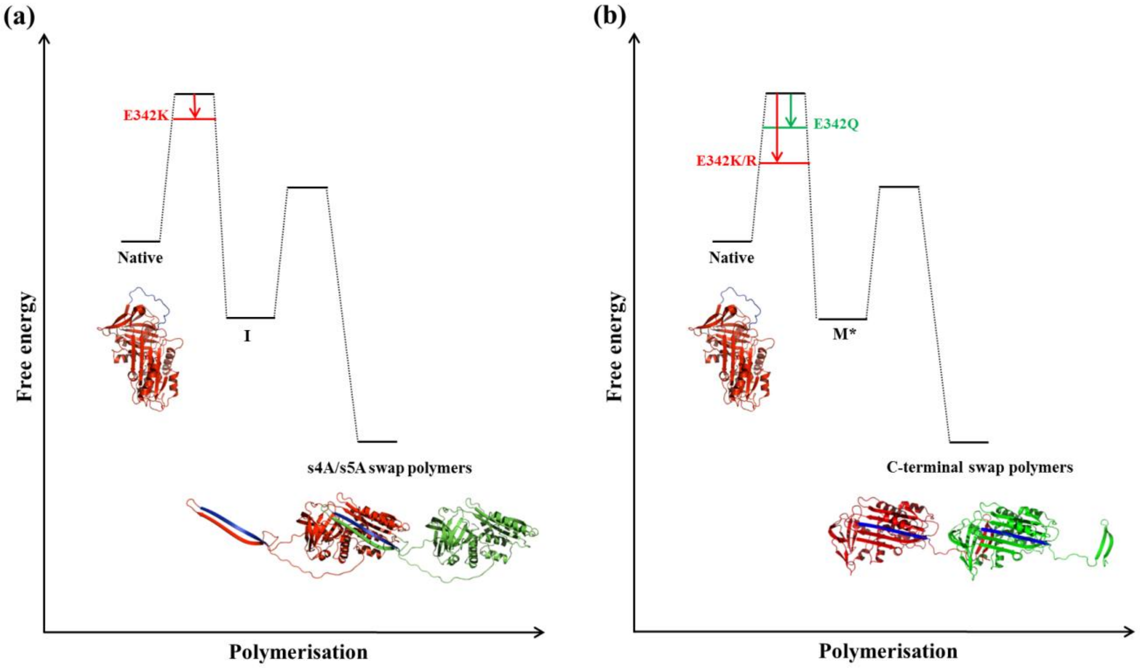


Figure 6.1. Free energy diagrams of α_1AT polymerisation proceeding from the native state. (a) α_1AT polymerises via the folding intermediate I if incubated in low denaturant concentrations. The replacement of E342 with a lysine in Z α_1AT (E342K) leads to changes in the transition state between the native state and I and lowers the kinetic barrier to the adoption of the polymerogenic folding intermediate. Accumulation of I results in the dual-strand insertion of s4A and s5A of one molecule into β-sheet A of another molecule and consequently the formation of s4A/s5A swap polymers. (b) Incubation of α_1AT at elevated temperatures leads to the adoption of the polymer precursor M* which allows polymerisation to proceed via a C-terminal swap including s1C, s4B and s5B. Mutation of E342 to a neutral amino acid such as a glutamine (E342Q) leads to changes in the transition state between the native state and M* and a decrease in the kinetic barrier to M* formation. This effect is further enhanced by replacement of E342 with a positively charged amino acid such as a lysine in Z α_1AT or an arginine (E342R). Structures were adapted from [176].

Based on its reactivity with the 2C1 antibody the C-terminal swap polymer has been proposed to be the pathological α_1AT polymer present in hepatocellular inclusions [167]. It is not clear at this stage, however, whether Z α_1AT polymerises only via one pathway *in vivo* or whether there are several alternative polymerisation pathways. Parallel polymerisation pathways of α_1AT have been observed *in vitro* [167], but whether that is also the case *in vivo* remains to be elucidated. The 2C1 antibody was raised against α_1AT polymers formed *in vitro* at 60 °C [185]. If the pathological hepatocellular inclusions were potentially composed of a heterogeneous mixture of Z α_1AT polymers these pathological structures might also react with antibodies raised against polymers formed under alternative *in vitro* conditions. The development of such alternative antibodies therefore represents an important hurdle in the understanding of Z α_1AT polymerisation and consequently in its therapeutic intervention.

Since different types of α_1AT polymers seem to be formed in low denaturant concentrations and at elevated temperatures this suggests that these two polymerisation pathways involve the formation of two alternative polymerogenic species/ intermediates. The monomeric non-native α_1AT species formed at elevated temperatures has been termed M* [198] and it has been shown that at least helix F undergoes different conformational changes during its adoption in comparison to the formation of the folding intermediate I [93]. It is unclear how M* formation of Z α_1AT relates to the formation of Z α_1AT polymers *in vivo*. The observation that Z α_1AT forms 2C1 antibody reactive polymers upon incubation at 60 °C and even at the physiological temperature 41 °C [185] suggests that polymerisation of Z α_1AT from M* is of physiological relevance. Considering that the α_1AT polymerisation reaction is concentration dependent [142, 143, 198] and the ER is not only the place of α_1AT folding but also the place of its highest concentration [142] one might expect Z α_1AT to mainly polymerise in the ER even if polymerisation was to proceed from the native state via M*. *In vivo* polymerisation of Z α_1AT from its native state could be supported by the detection of Z α_1AT polymers in circulation [207, 209] and in the lung [172].

Whether Z α_1AT polymerises *in vivo* from the folding intermediate I, from M* or from both is not known, however, this thesis suggests that the Z mutation leads to changes in the kinetic barriers between the native state and I (Figure 6.1a) as well as M* (Figure 6.1b) (chapter 3, chapter 4 and chapter 5). Whereas the rate of I formation was approximately 1.5-fold faster for Z α_1AT in comparison to the wild-type (chapter 3) the rate of M* formation seems to be more significantly affected by the pathological mutation. Incubation of Z α_1AT at 42 °C resulted in the subsequent build-up of polymeric species and the half-life of native Z α_1AT in this reaction was approximately 15 hours (chapter 5). No loss of wild-type α_1AT was observed after 72 hours of incubation under identical conditions (chapter 5). When the polymerisation reaction was conducted at 60 °C the difference in the rate of monomer loss between wild-type and Z α_1AT was only approximately 2-fold (chapter 4). It has been suggested previously that the chaotropic effect of extreme temperatures is so severe on the structure of α_1AT that it eliminates the effect of the Z mutation [143] and this notion is supported by the work presented in this thesis.

It is unclear how the Z mutation accelerates the rates of I and M* formation. A significant decrease in the thermodynamic stability of α_1AT upon Z mutation would, however, provide a satisfactory explanation. Most of the mutations that lead to serpin polymerisation and protein fibrillisation events are associated with a decrease in native state stability of the aggregating protein [61, 227]. This observation has contributed to the establishment of the conformational change hypothesis which postulates that at least partial unfolding of the native state is required to initiate the aggregation process of a globular protein [227]. If the Z mutation destabilised the native state of α_1AT this would result in an increase in the equilibrium concentration of the partially folded species I and M* and consequently enhanced polymerisation. One study suggests that Z α_1AT is indeed characterised by a decrease in the melting temperature of approximately 8 °C in comparison to the wild-type [198]. A second study, however, proposes that the stability of Z α_1AT is similar to wild-type α_1AT as estimated from transverse urea gradient (TUG) gels [213]. The work presented here is supportive of the latter study suggesting that the Z mutation does not result in a significant decrease in the thermodynamic stability of α_1AT. A difference of approximately 1 °C was observed in the melting temperature of Z α_1AT in comparison to the wild-type (chapter 3) consistent with equilibrium unfolding studies in guanidine hydrochloride (GdnHCl) (chapter 3) and urea (chapter 4) as well as TUG gels (chapter 3). This suggests that the increase in the rates of I and M* formation upon Z mutation are caused by changes in the transition states and not by a decrease in the thermodynamic stability of α_1AT (Figure 6.1).

The equilibrium unfolding curves in chapters 3 and 4 suggest that the stability of the folding intermediate I is unaffected by the Z mutation as estimated from the midpoints of the second (I → U) unfolding transition. This suggests that the Z mutation does not result in stabilisation of the folding intermediate I which would have been an appropriate explanation for the previously observed I to N folding retardation of Z α_1AT [213]. The I to U transition is, however, highly non-cooperative [192] and it is therefore possible that small changes in the reaction caused by single amino acid substitutions, such as the increase in stability of the folding intermediate upon Z mutation, might not be readily detectable.

The crystal structure of native α_1AT shows that Glu342 is located at the top of the central β-sheet A and the base of the reactive centre loop (RCL) and that it forms a hydrogen bond to Thr203 and a salt bridge to Lys290 [60]. The Z mutation replaces Glu342 with a lysine (Glu342Lys) and results in the disruption of the 290-342 salt bridge. Hence, it is surprising that the Z mutation has little effect on the stability of the native state of α_1AT. The molecular dynamics simulations presented in this thesis (chapter 5), however, suggest that the Z mutation results in the formation of new interactions which compensate for the lost interactions and hence explain why Z α_1AT is thermodynamically as stable as the wild-type.

It is controversial whether the loss of the 290-342 salt bridge is the main disruption that is responsible for the polymerisation prone nature of Z α_1AT. Whereas one publication indicates that reestablishment of the 290-342 salt bridge in Z α_1AT by introduction of a second mutation (Lys290Glu) leads to restoration of wild-type-like secretion levels [210] a second study suggests that the Lys290Glu mutation only marginally increases secretion levels of Z α_1AT [211]. Additionally, there is evidence that disruption of the 290-342 salt bridge from the other side by introducing the Lys290Glu mutation into wild-type α_1AT does not [212] or only slightly [211] decrease secretion levels. These findings led to the hypothesis that introduction of a positive charge at position 342 as a result of the Z mutation leads to a significant change in the electrostatic charge in that area of the α_1AT molecule which triggers misfolding and potentially the adoption of an alternative native conformation [211]. The work presented here supports this hypothesis as it shows that the charge of amino acid 342 plays an essential role in the polymerisation pathway of α_1AT (chapter 5) and that Z α_1AT is in an alternative native conformation in comparison to wild-type α_1AT (chapter 3, chapter 4 and chapter 5). The adoption of an alternative native structure of Z α_1AT in comparison to wild-type α_1AT has previously been proposed based on differences in near-UV circular dichroism (CD) spectra [143] and its interaction with RCL peptides whose exact binding sites are unknown [180]. The results in this thesis suggest that the Z mutation results in a structural rearrangement at the top of β-sheet A which is accompanied by an increase in the exposure of hydrophobic regions as determined by sensitive fluorescence spectroscopic measurements (chapter 3 and chapter 4) and molecular dynamics simulations (chapter 5).

The data presented in chapter 5 indicate that the side chain charge of amino acid 342 influences the rate of α_1AT polymerisation. Disruption of the 290-342 salt bridge by replacement of Glu342 with the neutral amino acid glutamine (Glu342Gln) resulted in polymerisation of α_1AT at 42 °C whereas wild-type α_1AT did not polymerise (chapter 5). The rate of polymerisation at 42 °C was, however, doubled if Glu342 was mutated to a positively charged amino acid residue (Glu342Lys and Glu342Arg) (chapter 5). Considering that disruption of the 290-342 salt bridge from the other side did not have the same devastating effect on the α_1AT molecule [211, 212] this indicates that it is the negative charge native to position 342 and not the 290-342 salt bridge that prevents polymerisation. Together, these data suggest that Glu342 forms interactions in the transition state between the native state and M* and that replacement of the negatively charged Glu342 with a neutral glutamine lowers the kinetic barrier to M* formation, an effect which is further enhanced by the replacement with a positive charge (Figure 6.1b).

In conclusion, the work presented in this thesis extends our understanding of serpin biology and clarifies several discrepancies in the field. It clearly demonstrates that Z α_1AT is in an altered but stable and functional conformation. Further, it suggests that the Z mutation lowers the kinetic barriers between the native state and I as well as M* which might explain why Z α_1AT is characterised by an increased propensity to polymerise. The conformational diseases have largely been attributed to mutations that decrease the native state stability of the aggregating proteins [61, 227]. For several proteins, however, aggregation-enhancing mutations do not result in destabilisation of the native state. These proteins include ataxin-3 [228, 229], prion protein [230, 231], T7 endonuclease I [232] and according to this thesis also Z α_1AT.

References

1. Levinthal, C., *How to Fold Graciously*. Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois, 1969: p. 22-24.
2. Anfinsen, C.B., *Principles that govern the folding of protein chains.* Science, 1973. **181**(96): p. 223-30.
3. Wetlaufer, D.B., *Nucleation, rapid folding, and globular intrachain regions in proteins.* Proc Natl Acad Sci U S A, 1973. **70**(3): p. 697-701.
4. Kim, P.S. and R.L. Baldwin, *Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding.* Annu Rev Biochem, 1982. **51**: p. 459-89.
5. Ptitsyn, O.B. and A.A. Rashin, *A model of myoglobin self-organization.* Biophys Chem, 1975. **3**(1): p. 1-20.
6. Dill, K.A., *Theory for the folding and stability of globular proteins.* Biochemistry, 1985. **24**(6): p. 1501-9.
7. Kauzmann, W., *Some factors in the interpretation of protein denaturation.* Adv Protein Chem, 1959. **14**: p. 1-63.
8. Harrison, S.C. and R. Durbin, *Is there a single pathway for the folding of a polypeptide chain?* Proc Natl Acad Sci U S A, 1985. **82**(12): p. 4028-30.
9. Karplus, M. and D.L. Weaver, *Protein-folding dynamics.* Nature, 1976. **260**(5550): p. 404-6.
10. Karplus, M. and D.L. Weaver, *Protein folding dynamics: the diffusion-collision model and experimental data.* Protein Sci, 1994. **3**(4): p. 650-68.
11. Kim, P.S. and R.L. Baldwin, *Intermediates in the folding reactions of small proteins.* Annu Rev Biochem, 1990. **59**: p. 631-60.
12. Ptitsyn, O.B., et al., *Evidence for a molten globule state as a general intermediate in protein folding.* FEBS Lett, 1990. **262**(1): p. 20-4.
13. Jackson, S.E. and A.R. Fersht, *Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition.* Biochemistry, 1991. **30**(43): p. 10428-35.
14. Otzen, D.E., et al., *Structure of the transition state for the folding/unfolding of the barley chymotrypsin inhibitor 2 and its implications for mechanisms of protein folding.* Proc Natl Acad Sci U S A, 1994. **91**(22): p. 10422-5.

15. Fersht, A.R., *Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications.* Proc Natl Acad Sci U S A, 1995. **92**(24): p. 10869-73.

16. Daggett, V. and A.R. Fersht, *Is there a unifying mechanism for protein folding?* Trends Biochem Sci, 2003. **28**(1): p. 18-25.

17. Vabulas, R.M., et al., *Protein folding in the cytoplasm and the heat shock response.* Cold Spring Harb Perspect Biol, 2010. **2**(12): p. a004390.

18. Leopold, P.E., M. Montal, and J.N. Onuchic, *Protein folding funnels: a kinetic approach to the sequence-structure relationship.* Proc Natl Acad Sci U S A, 1992. **89**(18): p. 8721-5.

19. Bartlett, A.I. and S.E. Radford, *An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms.* Nat Struct Mol Biol, 2009. **16**(6): p. 582-8.

20. Dill, K.A. and H.S. Chan, *From Levinthal to pathways to funnels.* Nat Struct Biol, 1997. **4**(1): p. 10-9.

21. Dobson, C.M. and M. Karplus, *The fundamentals of protein folding: bringing together theory and experiment.* Curr Opin Struct Biol, 1999. **9**(1): p. 92-101.

22. Stefani, M. and C.M. Dobson, *Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution.* J Mol Med, 2003. **81**(11): p. 678-99.

23. Baker, D. and D.A. Agard, *Kinetics versus thermodynamics in protein folding.* Biochemistry, 1994. **33**(24): p. 7505-9.

24. Cabrita, L.D. and S.P. Bottomley, *How do proteins avoid becoming too stable? Biophysical studies into metastable proteins.* Eur Biophys J, 2004. **33**(2): p. 83-8.

25. Baker, D., J.L. Sohl, and D.A. Agard, *A protein-folding reaction under kinetic control.* Nature, 1992. **356**(6366): p. 263-5.

26. Bullough, P.A., et al., *Structure of influenza haemagglutinin at the pH of membrane fusion.* Nature, 1994. **371**(6492): p. 37-43.

27. Eder, J., M. Rheinnecker, and A.R. Fersht, *Folding of subtilisin BPN': characterization of a folding intermediate.* Biochemistry, 1993. **32**(1): p. 18-26.

28. Strausberg, S., et al., *Catalysis of a protein folding reaction: thermodynamic and kinetic analysis of subtilisin BPN' interactions with its propeptide fragment.* Biochemistry, 1993. **32**(32): p. 8112-9.

29. Chan, D.C., et al., *Core structure of gp41 from the HIV envelope glycoprotein.* Cell, 1997. **89**(2): p. 263-73.
30. Clark, A.C., J.F. Sinclair, and T.O. Baldwin, *Folding of bacterial luciferase involves a non-native heterodimeric intermediate in equilibrium with the native enzyme and the unfolded subunits.* J Biol Chem, 1993. **268**(15): p. 10773-9.
31. Sinclair, J.F., et al., *Purified native subunits of bacterial luciferase are active in the bioluminescence reaction but fail to assemble into the alpha beta structure.* Biochemistry, 1993. **32**(19): p. 5036-44.
32. Whisstock, J., R. Skinner, and A.M. Lesk, *An atlas of serpin conformations.* Trends Biochem Sci, 1998. **23**(2): p. 63-7.
33. Ellis, R.J., *Macromolecular crowding: an important but neglected aspect of the intracellular environment.* Curr Opin Struct Biol, 2001. **11**(1): p. 114-9.
34. Sherman, M.Y. and A.L. Goldberg, *Cellular defenses against unfolded proteins: a cell biologist thinks about neurodegenerative diseases.* Neuron, 2001. **29**(1): p. 15-32.
35. Jahn, T.R. and S.E. Radford, *Folding versus aggregation: polypeptide conformations on competing pathways.* Arch Biochem Biophys, 2008. **469**(1): p. 100-17.
36. Hartl, F.U., *Molecular chaperones in cellular protein folding.* Nature, 1996. **381**(6583): p. 571-9.
37. Kostova, Z. and D.H. Wolf, *For whom the bell tolls: protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection.* EMBO J, 2003. **22**(10): p. 2309-17.
38. Romisch, K., *Endoplasmic reticulum-associated degradation.* Annu Rev Cell Dev Biol, 2005. **21**: p. 435-56.
39. Pan, T., et al., *The role of autophagy-lysosome pathway in neurodegeneration associated with Parkinson's disease.* Brain, 2008. **131**(Pt 8): p. 1969-78.
40. Ciechanover, A. and P. Brundin, *The ubiquitin proteasome system in neurodegenerative diseases: sometimes the chicken, sometimes the egg.* Neuron, 2003. **40**(2): p. 427-46.
41. Carrell, R.W. and D.A. Lomas, *Conformational disease.* Lancet, 1997. **350**(9071): p. 134-8.

42. Carrell, R.W. and D.A. Lomas, *Alpha1-antitrypsin deficiency--a model for conformational diseases.* N Engl J Med, 2002. **346**(1): p. 45-53.

43. Kruger, R., et al., *Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease.* Nat Genet, 1998. **18**(2): p. 106-8.

44. Polymeropoulos, M.H., et al., *Mutation in the alpha-synuclein gene identified in families with Parkinson's disease.* Science, 1997. **276**(5321): p. 2045-7.

45. Zarranz, J.J., et al., *The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia.* Ann Neurol, 2004. **55**(2): p. 164-73.

46. Hou, L., et al., *Solution NMR studies of the A beta(1-40) and A beta(1-42) peptides establish that the Met35 oxidation state affects the mechanism of amyloid formation.* J Am Chem Soc, 2004. **126**(7): p. 1992-2005.

47. Selkoe, D.J., *The molecular pathology of Alzheimer's disease.* Neuron, 1991. **6**(4): p. 487-98.

48. Liu, C., M.R. Sawaya, and D. Eisenberg, *beta-microglobulin forms three-dimensional domain-swapped amyloid fibrils with disulfide linkages.* Nat Struct Mol Biol, 2011. **18**(1): p. 49-55.

49. Bennett, M.J., et al., *A linear lattice model for polyglutamine in CAG-expansion diseases.* Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11634-9.

50. Li, W., et al., *Expression and characterization of full-length human huntingtin, an elongated HEAT repeat protein.* J Biol Chem, 2006. **281**(23): p. 15916-22.

51. Ellisdon, A.M., M.C. Pearce, and S.P. Bottomley, *Mechanisms of ataxin-3 misfolding and fibril formation: kinetic analysis of a disease-associated polyglutamine protein.* J Mol Biol, 2007. **368**(2): p. 595-605.

52. Masino, L., et al., *Domain architecture of the polyglutamine protein ataxin-3: a globular domain followed by a flexible tail.* FEBS Lett, 2003. **549**(1-3): p. 21-5.

53. Lotharius, J. and P. Brundin, *Pathogenesis of Parkinson's disease: dopamine, vesicles and alpha-synuclein.* Nat Rev Neurosci, 2002. **3**(12): p. 932-42.

54. Saraiva, M.J., *Transthyretin amyloidosis: a tale of weak interactions.* FEBS Lett, 2001. **498**(2-3): p. 201-3.

55. Clark, A., et al., *Islet amyloid formed from diabetes-associated peptide may be pathogenic in type-2 diabetes.* Lancet, 1987. **2**(8553): p. 231-4.

56. Kayed, R., et al., *Conformational transitions of islet amyloid polypeptide (IAPP) in amyloid formation in vitro.* J Mol Biol, 1999. **287**(4): p. 781-96.

57. Padrick, S.B. and A.D. Miranker, *Islet amyloid polypeptide: identification of long-range contacts and local order on the fibrillogenesis pathway.* J Mol Biol, 2001. **308**(4): p. 783-94.
58. Faber, J.P., et al., *The molecular basis of alpha 1-antichymotrypsin deficiency in a heterozygote with liver and lung disease.* J Hepatol, 1993. **18**(3): p. 313-21.
59. Gooptu, B., et al., *Inactive conformation of the serpin alpha(1)-antichymotrypsin indicates two-stage insertion of the reactive loop: implications for inhibitory function and conformational disease.* Proc Natl Acad Sci U S A, 2000. **97**(1): p. 67-72.
60. Elliott, P.R., et al., *Topography of a 2.0 A structure of alpha1-antitrypsin reveals targets for rational drug design to prevent conformational disease.* Protein Sci, 2000. **9**(7): p. 1274-81.
61. Lomas, D.A. and R. Mahadeva, *Alpha1-antitrypsin polymerization and the serpinopathies: pathobiology and prospects for therapy.* J Clin Invest, 2002. **110**(11): p. 1585-90.
62. Johnson, D.J., et al., *Crystal structure of monomeric native antithrombin reveals a novel reactive center loop conformation.* J Biol Chem, 2006. **281**(46): p. 35478-86.
63. Picard, V., et al., *Antithrombin Phe229Leu: a new homozygous variant leading to spontaneous antithrombin polymerization in vivo associated with severe childhood thrombosis.* Blood, 2003. **102**(3): p. 919-25.
64. Miranda, E., K. Romisch, and D.A. Lomas, *Mutants of neuroserpin that cause dementia accumulate as polymers within the endoplasmic reticulum.* J Biol Chem, 2004. **279**(27): p. 28283-91.
65. Takehara, S., et al., *The 2.1-A crystal structure of native neuroserpin reveals unique structural elements that contribute to conformational instability.* J Mol Biol, 2009. **388**(1): p. 11-20.
66. Citron, M., *Alzheimer's disease: treatments in discovery and development.* Nat Neurosci, 2002. **5 Suppl**: p. 1055-7.
67. Esler, W.P. and M.S. Wolfe, *A portrait of Alzheimer secretases--new features and familiar faces.* Science, 2001. **293**(5534): p. 1449-54.

68. Chow, M.K., D.A. Lomas, and S.P. Bottomley, *Promiscuous beta-strand interactions and the conformational diseases.* Curr Med Chem, 2004. **11**(4): p. 491-9.

69. Stromer, T. and L.C. Serpell, *Structure and morphology of the Alzheimer's amyloid fibril.* Microsc Res Tech, 2005. **67**(3-4): p. 210-7.

70. Lomas, D.A., et al., *Alpha 1-antitrypsin Siiyama (Ser53-->Phe). Further evidence for intracellular loop-sheet polymerization.* J Biol Chem, 1993. **268**(21): p. 15333-5.

71. Devlin, G.L. and S.P. Bottomley, *A protein family under 'stress' - serpin stability, folding and misfolding.* Front Biosci, 2005. **10**: p. 288-99.

72. Silverman, G.A., et al., *Serpins flex their muscle: I. Putting the clamps on proteolysis in diverse biological systems.* J Biol Chem, 2010. **285**(32): p. 24299-305.

73. Hunt, L.T. and M.O. Dayhoff, *A surprising new protein superfamily containing ovalbumin, antithrombin-III, and alpha 1-proteinase inhibitor.* Biochem Biophys Res Commun, 1980. **95**(2): p. 864-71.

74. Carrell, R. and J. Travis, *alpha-1-antitrypsin and the serpins: variation and countervariation.* Trends Biochem Sci, 1985. **10**(1): p. 20-24.

75. Schick, C., et al., *Cross-class inhibition of the cysteine proteinases cathepsins K, L, and S by the serpin squamous cell carcinoma antigen 1: a kinetic analysis.* Biochemistry, 1998. **37**(15): p. 5258-66.

76. Irving, J.A., et al., *Inhibitory activity of a heterochromatin-associated serpin (MENT) against papain-like cysteine proteinases affects chromatin structure and blocks cell proliferation.* J Biol Chem, 2002. **277**(15): p. 13192-201.

77. Pak, S.C., et al., *SRP-2 is a cross-class inhibitor that participates in postembryonic development of the nematode Caenorhabditis elegans: initial characterization of the clade L serpins.* J Biol Chem, 2004. **279**(15): p. 15448-59.

78. Hwang, S.R., et al., *The novel serpin endopin 2 demonstrates cross-class inhibition of papain and elastase: localization of endopin 2 to regulated secretory vesicles of neuroendocrine chromaffin cells.* Biochemistry, 2002. **41**(33): p. 10397-405.

79. Al-Khunaizi, M., et al., *The serpin SQN-5 is a dual mechanistic-class inhibitor of serine and cysteine proteinases.* Biochemistry, 2002. **41**(9): p. 3189-99.

80. Degterev, A., M. Boyce, and J. Yuan, *A decade of caspases.* Oncogene, 2003. **22**(53): p. 8543-67.
81. Gettins, P.G., *Serpin structure, mechanism, and function.* Chem Rev, 2002. **102**(12): p. 4751-804.
82. Zou, Z., et al., *Maspin, a serpin with tumor-suppressing activity in human mammary epithelial cells.* Science, 1994. **263**(5146): p. 526-9.
83. Nagata, K., *Hsp47: a collagen-specific molecular chaperone.* Trends Biochem Sci, 1996. **21**(1): p. 22-6.
84. Pemberton, P.A., et al., *Hormone binding globulins undergo serpin conformational change in inflammation.* Nature, 1988. **336**(6196): p. 257-8.
85. Irving, J.A., et al., *Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function.* Genome Res, 2000. **10**(12): p. 1845-64.
86. Remold-O'Donnell, E., *The ovalbumin family of serpin proteins.* FEBS Lett, 1993. **315**(2): p. 105-8.
87. Dickinson, J.L., et al., *The C-D interhelical domain of the serpin plasminogen activator inhibitor-type 2 is required for protection from TNF-alpha induced apoptosis.* Cell Death Differ, 1998. **5**(2): p. 163-71.
88. Potempa, J., E. Korzus, and J. Travis, *The serpin superfamily of proteinase inhibitors: structure, function, and regulation.* J Biol Chem, 1994. **269**(23): p. 15957-60.
89. Silverman, G.A., et al., *The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature.* J Biol Chem, 2001. **276**(36): p. 33293-6.
90. Buzza, M.S., et al., *The granzyme B inhibitor, PI-9, is present in endothelial and mesothelial cells, suggesting that it protects bystander cells during immune responses.* Cell Immunol, 2001. **210**(1): p. 21-9.
91. Sun, J., et al., *A cytosolic granzyme B inhibitor related to the viral apoptotic regulator cytokine response modifier A is present in cytotoxic lymphocytes.* J Biol Chem, 1996. **271**(44): p. 27802-9.

92. Whisstock, J.C., et al., *Serpins flex their muscle: II. Structural insights into target peptidase recognition, polymerization, and transport functions.* J Biol Chem, 2010. **285**(32): p. 24307-12.
93. Cabrita, L.D., W. Dai, and S.P. Bottomley, *Different conformational changes within the F-helix occur during serpin folding, polymerization, and proteinase inhibition.* Biochemistry, 2004. **43**(30): p. 9834-9.
94. Cabrita, L.D., J.C. Whisstock, and S.P. Bottomley, *Probing the role of the F-helix in serpin stability through a single tryptophan substitution.* Biochemistry, 2002. **41**(14): p. 4575-81.
95. Law, R.H., et al., *An overview of the serpin superfamily.* Genome Biol, 2006. **7**(5): p. 216.
96. Stein, P.E., et al., *Crystal structure of ovalbumin as a model for the reactive centre of serpins.* Nature, 1990. **347**(6288): p. 99-102.
97. Wei, A., et al., *Crystal structure of an uncleaved serpin reveals the conformation of an inhibitory reactive loop.* Nat Struct Biol, 1994. **1**(4): p. 251-8.
98. Elliott, P.R., J.P. Abrahams, and D.A. Lomas, *Wild-type alpha 1-antitrypsin is in the canonical inhibitory conformation.* J Mol Biol, 1998. **275**(3): p. 419-25.
99. Elliott, P.R., et al., *Inhibitory conformation of the reactive loop of alpha 1-antitrypsin.* Nat Struct Biol, 1996. **3**(8): p. 676-81.
100. Carrell, R.W., et al., *Biological implications of a 3 A structure of dimeric antithrombin.* Structure, 1994. **2**(4): p. 257-70.
101. Schreuder, H.A., et al., *The intact and cleaved human antithrombin III complex as a model for serpin-proteinase interactions.* Nat Struct Biol, 1994. **1**(1): p. 48-54.
102. Skinner, R., et al., *The 2.6 A structure of antithrombin indicates a conformational change at the heparin binding site.* J Mol Biol, 1997. **266**(3): p. 601-9.
103. Loebermann, H., et al., *Human alpha 1-proteinase inhibitor. Crystal structure analysis of two crystal modifications, molecular model and preliminary analysis of the implications for function.* J Mol Biol, 1984. **177**(3): p. 531-57.
104. Bjork, I., et al., *Kinetic characterization of the substrate reaction between a complex of antithrombin with a synthetic reactive-bond loop tetradecapeptide and four target proteinases of the inhibitor.* J Biol Chem, 1992. **267**(27): p. 19047-50.

105. Fa, M., et al., *The structure of a serpin-protease complex revealed by intramolecular distance measurements using donor-donor energy migration and mapping of interaction sites.* Structure, 2000. **8**(4): p. 397-405.

106. Mast, A.E., J.J. Enghild, and G. Salvesen, *Conformation of the reactive site loop of alpha 1-proteinase inhibitor probed by limited proteolysis.* Biochemistry, 1992. **31**(10): p. 2720-8.

107. Owen, W.G., *Evidence for the formation of an ester between thrombin and heparin cofactor.* Biochim Biophys Acta, 1975. **405**(2): p. 380-7.

108. Patston, P.A., et al., *Mechanism of serpin action: evidence that C1 inhibitor functions as a suicide substrate.* Biochemistry, 1991. **30**(36): p. 8876-82.

109. Schulze, A.J., et al., *Evidence for the extent of insertion of the active site loop of intact alpha 1 proteinase inhibitor in beta-sheet A.* Biochemistry, 1992. **31**(33): p. 7560-5.

110. Stratikos, E. and P.G. Gettins, *Formation of the covalent serpin-proteinase complex involves translocation of the proteinase by more than 70 A and full insertion of the reactive center loop into beta-sheet A.* Proc Natl Acad Sci U S A, 1999. **96**(9): p. 4808-13.

111. Wilczynska, M., et al., *The inhibition mechanism of serpins. Evidence that the mobile reactive center loop is cleaved in the native protease-inhibitor complex.* J Biol Chem, 1995. **270**(50): p. 29652-5.

112. Huntington, J.A., R.J. Read, and R.W. Carrell, *Structure of a serpin-protease complex shows inhibition by deformation.* Nature, 2000. **407**(6806): p. 923-6.

113. Calugaru, S.V., R. Swanson, and S.T. Olson, *The pH dependence of serpin-proteinase complex dissociation reveals a mechanism of complex stabilization involving inactive and active conformational states of the proteinase which are perturbable by calcium.* J Biol Chem, 2001. **276**(35): p. 32446-55.

114. Plotnick, M.I., et al., *Heterogeneity in serpin-protease complexes as demonstrated by differences in the mechanism of complex breakdown.* Biochemistry, 2002. **41**(1): p. 334-42.

115. Kounnas, M.Z., et al., *Cellular internalization and degradation of antithrombin III-thrombin, heparin cofactor II-thrombin, and alpha 1-antitrypsin-trypsin complexes is mediated by the low density lipoprotein receptor-related protein.* J Biol Chem, 1996. **271**(11): p. 6523-9.

116. Orth, K., et al., *Complexes of tissue-type plasminogen activator and its serpin inhibitor plasminogen-activator inhibitor type 1 are internalized by means of the low density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor.* Proc Natl Acad Sci U S A, 1992. **89**(16): p. 7422-6.
117. Poller, W., et al., *Differential recognition of alpha 1-antitrypsin-elastase and alpha 1-antichymotrypsin-cathepsin G complexes by the low density lipoprotein receptor-related protein.* J Biol Chem, 1995. **270**(6): p. 2841-5.
118. Travis, J. and G.S. Salvesen, *Human plasma proteinase inhibitors.* Annu Rev Biochem, 1983. **52**: p. 655-709.
119. Huntington, J.A. and R.W. Carrell, *The serpins: nature's molecular mousetraps.* Sci Prog, 2001. **84**(Pt 2): p. 125-36.
120. Wright, H.T. and J.N. Scarsdale, *Structural basis for serpin inhibitor activity.* Proteins, 1995. **22**(3): p. 210-25.
121. Schechter, I. and A. Berger, *On the size of the active site in proteases. I. Papain.* Biochem Biophys Res Commun, 1967. **27**(2): p. 157-62.
122. Djie, M.Z., et al., *Role of the P2 residue in determining the specificity of serpins.* Biochemistry, 1996. **35**(35): p. 11461-9.
123. Djie, M.Z., S.R. Stone, and B.F. Le Bonniec, *Intrinsic specificity of the reactive site loop of alpha1-antitrypsin, alpha1-antichymotrypsin, antithrombin III, and protease nexin I.* J Biol Chem, 1997. **272**(26): p. 16268-73.
124. Hopkins, P.C. and S.R. Stone, *The contribution of the conserved hinge region residues of alpha1-antitrypsin to its reaction with elastase.* Biochemistry, 1995. **34**(48): p. 15872-9.
125. Rubin, H., et al., *Conversion of alpha 1-antichymotrypsin into a human neutrophil elastase inhibitor: demonstration of variants with different association rate constants, stoichiometries of inhibition, and complex stabilities.* Biochemistry, 1994. **33**(24): p. 7627-33.
126. Schechter, N.M. and M.I. Plotnick, *Measurement of the kinetic parameters mediating protease-serpin inhibition.* Methods, 2004. **32**(2): p. 159-68.
127. Lawrence, D.A., et al., *Partitioning of serpin-proteinase reactions between stable inhibition and substrate cleavage is regulated by the rate of serpin reactive center loop insertion into beta-sheet A.* J Biol Chem, 2000. **275**(8): p. 5839-44.

128. Stratikos, E. and P.G. Gettins, *Mapping the serpin-proteinase complex using single cysteine variants of alpha1-proteinase inhibitor Pittsburgh.* J Biol Chem, 1998. **273**(25): p. 15582-9.
129. Shin, J.S. and M.H. Yu, *Viscous drag as the source of active site perturbation during protease translocation: insights into how inhibitory processes are controlled by serpin metastability.* J Mol Biol, 2006. **359**(2): p. 378-89.
130. Bruch, M., V. Weiss, and J. Engel, *Plasma serine proteinase inhibitors (serpins) exhibit major conformational changes and a large increase in conformational stability upon cleavage at their reactive sites.* J Biol Chem, 1988. **263**(32): p. 16626-30.
131. Herve, M. and C. Ghelis, *Conformational changes in intact and papain-modified alpha 1-proteinase inhibitor induced by guanidinium chloride.* Eur J Biochem, 1990. **191**(3): p. 653-8.
132. Powell, L.M. and R.H. Pain, *Effects of glycosylation on the folding and stability of human, recombinant and cleaved alpha 1-antitrypsin.* J Mol Biol, 1992. **224**(1): p. 241-52.
133. Carrell, R.W. and M.C. Owen, *Plakalbumin, alpha 1-antitrypsin, antithrombin and the mechanism of inflammatory thrombosis.* Nature, 1985. **317**(6039): p. 730-2.
134. Yamasaki, M., et al., *Loop-inserted and thermostabilized structure of P1-P1' cleaved ovalbumin mutant R339T.* J Mol Biol, 2002. **315**(2): p. 113-20.
135. Lee, K.N., S.D. Park, and M.H. Yu, *Probing the native strain in alpha1-antitrypsin.* Nat Struct Biol, 1996. **3**(6): p. 497-500.
136. Ryu, S.E., et al., *The native strains in the hydrophobic core and flexible reactive loop of a serine protease inhibitor: crystal structure of an uncleaved alpha1-antitrypsin at 2.7 A.* Structure, 1996. **4**(10): p. 1181-92.
137. Seo, E.J., et al., *Distribution of the native strain in human alpha 1-antitrypsin and its association with protease inhibitor function.* J Biol Chem, 2000. **275**(22): p. 16904-9.
138. Im, H., E.J. Seo, and M.H. Yu, *Metastability in the inhibitory mechanism of human alpha1-antitrypsin.* J Biol Chem, 1999. **274**(16): p. 11072-7.
139. Stein, P.E. and R.W. Carrell, *What do dysfunctional serpins tell us about molecular mobility and disease?* Nat Struct Biol, 1995. **2**(2): p. 96-113.

140. Whisstock, J.C., et al., *Conformational changes in serpins: I. The native and cleaved conformations of alpha(1)-antitrypsin.* J Mol Biol, 2000. **296**(2): p. 685-99.

141. Hopkins, P.C., R.W. Carrell, and S.R. Stone, *Effects of mutations in the hinge region of serpins.* Biochemistry, 1993. **32**(30): p. 7650-7.

142. Lomas, D.A., et al., *The mechanism of Z alpha 1-antitrypsin accumulation in the liver.* Nature, 1992. **357**(6379): p. 605-7.

143. Lomas, D.A., et al., *Effect of the Z mutation on the physical and inhibitory properties of alpha 1-antitrypsin.* Biochemistry, 1993. **32**(2): p. 500-8.

144. Levina, V., et al., *Expression, purification and characterization of recombinant Z alpha(1)-Antitrypsin-The most common cause of alpha(1)-Antitrypsin deficiency.* Protein Expr Purif, 2009.

145. Ogushi, F., et al., *Z-type alpha 1-antitrypsin is less competent than M1-type alpha 1-antitrypsin as an inhibitor of neutrophil elastase.* J Clin Invest, 1987. **80**(5): p. 1366-74.

146. Blouse, G.E., et al., *Mutation of the highly conserved tryptophan in the serpin breach region alters the inhibitory mechanism of plasminogen activator inhibitor-1.* Biochemistry, 2003. **42**(42): p. 12260-72.

147. Lomas, D.A., et al., *alpha 1-Antitrypsin Mmalton (Phe52-deleted) forms loop-sheet polymers in vivo. Evidence for the C sheet mechanism of polymerization.* J Biol Chem, 1995. **270**(28): p. 16864-70.

148. Poller, W., et al., *A leucine-to-proline substitution causes a defective alpha 1-antichymotrypsin allele associated with familial obstructive lung disease.* Genomics, 1993. **17**(3): p. 740-3.

149. Millar, D.S., et al., *Three novel missense mutations in the antithrombin III (AT3) gene causing recurrent venous thrombosis.* Hum Genet, 1994. **94**(5): p. 509-12.

150. Mottonen, J., et al., *Structural basis of latency in plasminogen activator inhibitor-1.* Nature, 1992. **355**(6357): p. 270-3.

151. Tucker, H.M., et al., *Engineering of plasminogen activator inhibitor-1 to reduce the rate of latency transition.* Nat Struct Biol, 1995. **2**(6): p. 442-5.

152. Im, H., et al., *Interactions causing the kinetic trap in serpin protein folding.* J Biol Chem, 2002. **277**(48): p. 46347-54.

153. Devlin, G.L., et al., *Acid Denaturation of alpha1-antitrypsin: characterization of a novel mechanism of serpin polymerization.* J Mol Biol, 2002. **324**(4): p. 859-70.

154. James, E.L. and S.P. Bottomley, *The mechanism of alpha 1-antitrypsin polymerization probed by fluorescence spectroscopy.* Arch Biochem Biophys, 1998. **356**(2): p. 296-300.

155. Lomas, D.A., et al., *Preparation and characterization of latent alpha 1-antitrypsin.* J Biol Chem, 1995. **270**(10): p. 5282-8.

156. Hekman, C.M. and D.J. Loskutoff, *Endothelial cells produce a latent inhibitor of plasminogen activators that can be activated by denaturants.* J Biol Chem, 1985. **260**(21): p. 11581-7.

157. Berkenpas, M.B., D.A. Lawrence, and D. Ginsburg, *Molecular evolution of plasminogen activator inhibitor-1 functional stability.* EMBO J, 1995. **14**(13): p. 2969-77.

158. Sancho, E., et al., *Conformational studies on plasminogen activator inhibitor (PAI-1) in active, latent, substrate, and cleaved forms.* Biochemistry, 1995. **34**(3): p. 1064-9.

159. Chang, W.S. and D.A. Lomas, *Latent alpha1-antichymotrypsin. A molecular explanation for the inactivation of alpha1-antichymotrypsin in chronic bronchitis and emphysema.* J Biol Chem, 1998. **273**(6): p. 3695-701.

160. Beauchamp, N.J., et al., *Antithrombins Wibble and Wobble (T85M/K): archetypal conformational diseases with in vivo latent-transition, thrombosis, and heparin activation.* Blood, 1998. **92**(8): p. 2696-706.

161. Wardell, M.R., et al., *Crystallization and preliminary X-ray diffraction analysis of two conformations of intact human antithrombin.* J Mol Biol, 1993. **234**(4): p. 1253-8.

162. Pearce, M.C., et al., *Identification and characterization of a misfolded monomeric serpin formed at physiological temperature.* J Mol Biol, 2010. **403**(3): p. 459-67.

163. Lomas, D.A. and R.W. Carrell, *Serpinopathies and the conformational dementias.* Nat Rev Genet, 2002. **3**(10): p. 759-68.

164. Lomas, D.A. and H. Parfrey, *Alpha1-antitrypsin deficiency. 4: Molecular pathophysiology.* Thorax, 2004. **59**(6): p. 529-35.

165. Bottomley, S.P., P.C. Hopkins, and J.C. Whisstock, *Alpha 1-antitrypsin polymerisation can occur by both loop A and C sheet mechanisms.* Biochem Biophys Res Commun, 1998. **251**(1): p. 1-5.
166. Koloczek, H., et al., *Serpin alpha 1proteinase inhibitor probed by intrinsic tryptophan fluorescence spectroscopy.* Protein Sci, 1996. **5**(11): p. 2226-35.
167. Yamasaki, M., et al., *Molecular basis of alpha1-antitrypsin deficiency revealed by the structure of a domain-swapped trimer.* EMBO Rep, 2011. **12**(10): p. 1011-7.
168. Marszal, E., D. Danino, and A. Shrake, *A novel mode of polymerization of alpha1-proteinase inhibitor.* J Biol Chem, 2003. **278**(22): p. 19611-8.
169. Dunstone, M.A., et al., *Cleaved antitrypsin polymers at atomic resolution.* Protein Sci, 2000. **9**(2): p. 417-20.
170. Huntington, J.A., et al., *A 2.6 A structure of a serpin polymer and implications for conformational disease.* J Mol Biol, 1999. **293**(3): p. 449-55.
171. Sharp, A.M., et al., *The active conformation of plasminogen activator inhibitor 1, a target for drugs to control fibrinolysis and cell adhesion.* Structure, 1999. **7**(2): p. 111-8.
172. Elliott, P.R., D. Bilton, and D.A. Lomas, *Lung polymers in Z alpha1-antitrypsin deficiency-related emphysema.* Am J Respir Cell Mol Biol, 1998. **18**(5): p. 670-4.
173. Mahadeva, R., et al., *Polymers of Z alpha1-antitrypsin co-localize with neutrophils in emphysematous alveoli and are chemotactic in vivo.* Am J Pathol, 2005. **166**(2): p. 377-86.
174. Mulgrew, A.T., et al., *Z alpha1-antitrypsin polymerizes in the lung and acts as a neutrophil chemoattractant.* Chest, 2004. **125**(5): p. 1952-7.
175. Zhou, A., et al., *Polymerization of plasminogen activator inhibitor-1.* J Biol Chem, 2001. **276**(12): p. 9115-22.
176. Bottomley, S.P., *The structural diversity in alpha1-antitrypsin misfolding.* EMBO Rep, 2011. **12**(10): p. 983-4.
177. Tsutsui, Y., et al., *The structural basis of serpin polymerization studied by hydrogen/deuterium exchange and mass spectrometry.* J Biol Chem, 2008. **283**(45): p. 30804-11.
178. Chang, W.S., et al., *Importance of the release of strand 1C to the polymerization mechanism of inhibitory serpins.* Protein Sci, 1997. **6**(1): p. 89-98.

179. Fitton, H.L., et al., *Mechanisms of antithrombin polymerisation and heparin activation probed by the insertion of synthetic reactive loop peptides.* Biol Chem, 1997. **378**(9): p. 1059-63.
180. Mahadeva, R., et al., *6-mer peptide selectively anneals to a pathogenic serpin conformation and blocks polymerization. Implications for the prevention of Z alpha(1)-antitrypsin-related cirrhosis.* J Biol Chem, 2002. **277**(9): p. 6771-4.
181. Schulze, A.J., et al., *Structural transition of alpha 1-antitrypsin by a peptide sequentially similar to beta-strand s4A.* Eur J Biochem, 1990. **194**(1): p. 51-6.
182. Zhou, A., et al., *How small peptides block and reverse serpin polymerisation.* J Mol Biol, 2004. **342**(3): p. 931-41.
183. Sivasothy, P., et al., *Pathogenic alpha 1-antitrypsin polymers are formed by reactive loop-beta-sheet A linkage.* J Biol Chem, 2000. **275**(43): p. 33663-8.
184. Yamasaki, M., et al., *Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization.* Nature, 2008. **455**(7217): p. 1255-8.
185. Miranda, E., et al., *A novel monoclonal antibody to characterize pathogenic polymers in liver disease associated with alpha1-antitrypsin deficiency.* Hepatology, 2010. **52**(3): p. 1078-88.
186. James, E.L., et al., *Probing the unfolding pathway of alpha1-antitrypsin.* J Biol Chem, 1999. **274**(14): p. 9482-8.
187. Pearce, M.C., H. Rubin, and S.P. Bottomley, *Conformational change and intermediates in the unfolding of alpha 1-antichymotrypsin.* J Biol Chem, 2000. **275**(37): p. 28513-8.
188. Tew, D.J. and S.P. Bottomley, *Probing the equilibrium denaturation of the serpin alpha(1)-antitrypsin with single tryptophan mutants; evidence for structure in the urea unfolded state.* J Mol Biol, 2001. **313**(5): p. 1161-9.
189. Villanueva, G.B. and N. Allen, *Demonstration of a two-domain structure of antithrombin III during its denaturation in guanidinium chloride.* J Biol Chem, 1983. **258**(18): p. 11010-3.
190. Wang, Z., J. Mottonen, and E.J. Goldsmith, *Kinetically controlled folding of the serpin plasminogen activator inhibitor 1.* Biochemistry, 1996. **35**(51): p. 16443-8.
191. Krishnan, B. and L.M. Gierasch, *Dynamic local unfolding in the serpin alpha-1 antitrypsin provides a mechanism for loop insertion and polymerization.* Nat Struct Mol Biol, 2011. **18**(2): p. 222-6.

192. Tsutsui, Y. and P.L. Wintrode, *Cooperative unfolding of a metastable serpin to a molten globule suggests a link between functional and folding energy landscapes.* J Mol Biol, 2007. **371**(1): p. 245-55.

193. Pande, V.S., et al., *Pathways for protein folding: is a new view needed?* Curr Opin Struct Biol, 1998. **8**(1): p. 68-79.

194. Fish, W.W., et al., *Denaturation behavior of antithrombin in guanidinium chloride. Irreversibility of unfolding caused by aggregation.* Biochemistry, 1985. **24**(6): p. 1510-7.

195. Koloczek, H., A. Guz, and P. Kaszycki, *Fluorescence-detected polymerization kinetics of human alpha 1-antitrypsin.* J Protein Chem, 1996. **15**(5): p. 447-54.

196. Liu, T., P.A. Pemberton, and A.D. Robertson, *Three-state unfolding and self-association of maspin, a tumor-suppressing serpin.* J Biol Chem, 1999. **274**(42): p. 29628-32.

197. Crowther, D.C., et al., *Nucleation of alpha 1-antichymotrypsin polymerization.* Biochemistry, 2003. **42**(8): p. 2355-63.

198. Dafforn, T.R., et al., *A kinetic mechanism for the polymerization of alpha1-antitrypsin.* J Biol Chem, 1999. **274**(14): p. 9548-55.

199. Ekeowa, U.I., et al., *Defining the mechanism of polymerization in the serpinopathies.* Proc Natl Acad Sci U S A, 2010. **107**(40): p. 17146-51.

200. Kamimoto, T., et al., *Intracellular inclusions containing mutant alpha1-antitrypsin Z are propagated in the absence of autophagic activity.* J Biol Chem, 2006. **281**(7): p. 4467-76.

201. Teckman, J.H., et al., *Fasting in alpha1-antitrypsin deficient liver: constitutive [correction of consultative] activation of autophagy.* Am J Physiol Gastrointest Liver Physiol, 2002. **283**(5): p. G1156-65.

202. Perlmutter, D.H., *Liver injury in alpha1-antitrypsin deficiency: an aggregated protein induces mitochondrial injury.* J Clin Invest, 2002. **110**(11): p. 1579-83.

203. Fregonese, L. and J. Stolk, *Hereditary alpha-1-antitrypsin deficiency and its clinical consequences.* Orphanet J Rare Dis, 2008. **3**: p. 16.

204. Gadek, J.E., et al., *Antielastases of the human alveolar structures. Implications for the protease-antiprotease theory of emphysema.* J Clin Invest, 1981. **68**(4): p. 889-98.

205. Stoller, J.K. and L.S. Aboussouan, *Alpha1-antitrypsin deficiency.* Lancet, 2005. **365**(9478): p. 2225-36.

206. Ioachimescu, O.C. and J.K. Stoller, *A review of alpha-1 antitrypsin deficiency.* COPD, 2005. **2**(2): p. 263-75.

207. Cox, D.W., G.D. Billingsley, and J.W. Callahan, *Aggregation of plasma Z type alpha 1-antitrypsin suggests basic defect for the deficiency.* FEBS Lett, 1986. **205**(2): p. 255-60.

208. Le, A., et al., *Soluble aggregates of the human PiZ alpha 1-antitrypsin variant are degraded within the endoplasmic reticulum by a mechanism sensitive to inhibitors of protein synthesis.* J Biol Chem, 1992. **267**(2): p. 1072-80.

209. Janciauskiene, S., et al., *Detection of circulating and endothelial cell polymers of Z and wild type alpha 1-antitrypsin by a monoclonal antibody.* J Biol Chem, 2002. **277**(29): p. 26540-6.

210. Brantly, M., M. Courtney, and R.G. Crystal, *Repair of the secretion defect in the Z form of alpha 1-antitrypsin by addition of a second mutation.* Science, 1988. **242**(4886): p. 1700-2.

211. Sifers, R.N., C.P. Hardick, and S.L. Woo, *Disruption of the 290-342 salt bridge is not responsible for the secretory defect of the PiZ alpha 1-antitrypsin variant.* J Biol Chem, 1989. **264**(5): p. 2997-3001.

212. Foreman, R.C., *Disruption of the Lys-290--Glu-342 salt bridge in human alpha 1-antitrypsin does not prevent its synthesis and secretion.* FEBS Lett, 1987. **216**(1): p. 79-82.

213. Yu, M.H., K.N. Lee, and J. Kim, *The Z type variation of human alpha 1-antitrypsin causes a protein folding defect.* Nat Struct Biol, 1995. **2**(5): p. 363-7.

214. Pace, C.N., *Determination and analysis of urea and guanidine hydrochloride denaturation curves.* Methods Enzymol, 1986. **131**: p. 266-80.

215. Hanahan, D., *Studies on transformation of Escherichia coli with plasmids.* J Mol Biol, 1983. **166**(4): p. 557-80.

216. Wu, S. and G.J. Letchworth, *High efficiency transformation by electroporation of Pichia pastoris pretreated with lithium acetate and dithiothreitol.* Biotechniques, 2004. **36**(1): p. 152-4.

217. Stone, S.R. and J. Hofsteenge, *Kinetics of the inhibition of thrombin by hirudin.* Biochemistry, 1986. **25**(16): p. 4622-8.

218. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy*. 3rd ed2006, New York: Springer.

219. Lehrer, S.S., *Solute perturbation of protein fluorescence. The quenching of the tryptophyl fluorescence of model compounds and of lysozyme by iodide ion.* Biochemistry, 1971. **10**(17): p. 3254-63.

220. Cabrita, L.D., et al., *Enhancing the stability and solubility of TEV protease using in silico design.* Protein Sci, 2007. **16**(11): p. 2360-7.

221. Barrick, D. and R.L. Baldwin, *Three-state analysis of sperm whale apomyoglobin folding.* Biochemistry, 1993. **32**(14): p. 3790-6.

222. Bolen, D.W. and M.M. Santoro, *Unfolding free energy changes determined by the linear extrapolation method. 2. Incorporation of delta G degrees N-U values in a thermodynamic cycle.* Biochemistry, 1988. **27**(21): p. 8069-74.

223. Santoro, M.M. and D.W. Bolen, *Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants.* Biochemistry, 1988. **27**(21): p. 8063-8.

224. Stefani, M. and C.M. Dobson, *Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution.* J Mol Med (Berl), 2003. **81**(11): p. 678-99.

225. Gooptu, B. and D.A. Lomas, *Polymers and inflammation: disease mechanisms of the serpinopathies.* J Exp Med, 2008. **205**(7): p. 1529-34.

226. Whisstock, J.C. and S.P. Bottomley, *Molecular gymnastics: serpin structure, folding and misfolding.* Curr Opin Struct Biol, 2006. **16**(6): p. 761-8.

227. Chiti, F. and C.M. Dobson, *Amyloid formation by globular proteins under native conditions.* Nat Chem Biol, 2009. **5**(1): p. 15-22.

228. Chow, M.K., et al., *Polyglutamine expansion in ataxin-3 does not affect protein stability: implications for misfolding and disease.* J Biol Chem, 2004. **279**(46): p. 47643-51.

229. Ellisdon, A.M., B. Thomas, and S.P. Bottomley, *The two-stage pathway of ataxin-3 fibrillogenesis involves a polyglutamine-independent step.* J Biol Chem, 2006. **281**(25): p. 16888-96.

230. Liemann, S. and R. Glockshuber, *Influence of amino acid substitutions related to inherited human prion diseases on the thermodynamic stability of the cellular prion protein.* Biochemistry, 1999. **38**(11): p. 3258-67.

231. Swietnicki, W., et al., *Familial mutations and the thermodynamic stability of the recombinant human prion protein.* J Biol Chem, 1998. **273**(47): p. 31048-52.

232. Guo, Z. and D. Eisenberg, *The mechanism of the amyloidogenic conversion of T7 endonuclease I.* J Biol Chem, 2007. **282**(20): p. 14968-74.

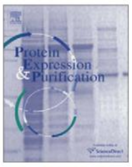
Appendix

Protein Expression and Purification 68 (2009) 226–232

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep



Expression, purification and characterization of recombinant Z α_1-Antitrypsin—The most common cause of α_1-Antitrypsin deficiency

Vita Levina, Weiwen Dai, Anja S. Knaupp, Dion Kaiserman, Mary C. Pearce, Lisa D. Cabrita, Phillip I. Bird, Stephen P. Bottomley *

Department of Biochemistry and Molecular Biology, Monash University, Clayton, Vic. 3800, Australia

ARTICLE INFO

Article history:
Received 27 May 2009
and in revised form 16 June 2009
Available online 23 June 2009

Keywords:
Serpin
Conformational disease
Proteinase inhibitor
Aggregation
Polymerization
Protein misfolding

ABSTRACT

α_1-Antitrypsin (α_1AT), the most abundant proteinase inhibitor circulating in the blood, protects extracellular matrix proteins of the lung against proteolytic destruction by neutrophil elastase. α_1AT deficiency predisposes patients to emphysema, juvenile cirrhosis and hepatocellular carcinoma. Over 90% of clinical cases of severe α_1AT deficiency are caused by the Z variant (E342K) of α_1AT. The presence of the Z mutation results in misfolding and polymerization of α_1AT. Due to its inherent propensity to polymerize there are no reported cases of recombinant Z α_1AT production. This has created a major impediment to studying the effect of the Z mutation on α_1AT. Here we report our attempts to produce recombinant Z α_1AT using both *Escherichia coli* and *Pichia pastoris* as host systems. Using a range of expression vectors in *E. coli* we were unable to produce soluble active Z α_1AT. Cytosolic expression of the Z α_1AT gene in *P. pastoris* was successful. Monomeric and active recombinant Z α_1AT was purified from the yeast cytosol using affinity chromatography and anion exchange chromatography. Biochemical analyses demonstrated that the recombinant Z α_1AT has identical properties to its native counterpart purified from plasma of patients homozygous for the Z allele. A recombinant source of pathological Z α_1AT will increase the chances of elucidating the mechanism of its polymerization and thus the development of therapeutic strategies.

© 2009 Elsevier Inc. All rights reserved.

Introduction

Recombinant protein production is a major challenge for many fields of study including biotechnology and biomedicine. Without the supply of correctly folded and active protein structural, biochemical and drug design projects are difficult to complete. Without doubt, one of the most challenging groups of targets for recombinant protein production are proteins involved in the conformational diseases [1,2]. The conformational diseases are a large class of progressive diseases such as Alzheimer disease, Huntington disease, and the serpinopathies in which proteins misfold and aggregate. Understanding how these proteins function, fold and misfold is critical for the development of therapeutic strategies.

The serpinopathies comprise diseases such as emphysema, liver cirrhosis and dementia, caused by the misfolding of serine proteinase inhibitors (serpins) [3,4]. α_1-Antitrypsin (α_1AT)[1] is the most abundant circulating proteinase inhibitor and the archetypal member of the serpin superfamily. It is expressed in hepatocytes and released into circulation where its primary function is to protect the lower respiratory tract from proteolytic destruction by elastase secreted by human neutrophils at sites of inflammation [5–8]. Most individuals have two M α_1AT alleles but approximately 1 in 2000 is homozygous for the Z allele [9,10]. The presence of the Z mutation results in the accumulation and retention of the polymerized α_1AT within the endoplasmic reticulum of hepatocytes [11,12]. The deposition of polymerized Z α_1AT within hepatocytes predisposes homozygotes to neonatal hepatitis, juvenile cirrhosis and hepatocellular carcinoma. In addition, the drastically decreased plasma levels of active α_1AT leads to a higher risk of developing early onset pulmonary emphysema [7].

There has always been a high demand for the recombinant form of α_1AT which would enable extensive studies of its mechanisms of proteinase inhibition, folding, misfolding and aggregation [13]. The first recombinant α_1AT (M allele) was purified from yeast in 1985 [14]. Since then, a number of expression systems have been utilized to successfully produce M α_1AT. Using the expression host *Escherichia coli*, a non-glycosylated, N-terminally truncated version of M α_1AT has been produced in inclusion bodies and successfully refolded [15,16]. Recently, soluble expression of full-length M α_1AT has been achieved using a different *E. coli* expression vector [17]. Glycosylated M α_1AT has also been produced from yeast (bakers and methylotrophic) [18] and plant systems [19–21]. However, even though many expression systems are now available for recombinant

---

* Corresponding author.
E-mail address: steve.bottomley@med.monash.edu.au (S.P. Bottomley).

[1] *Abbreviations used:* α_1AT, α_1-Antitrypsin; DTT, 1,4-dithiothreitol; EDTA, ethylenediaminetetraacetic acid; GST, glutathione-S-transferase; HNE, human neutrophil elastase; HRP, horseradish peroxidase; MBP, maltose binding protein; SI, stoichiometry of inhibition; TEV, tobacco etch virus; YPD, yeast peptone dextrose.

1046-5928/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2009.06.011

M α_1AT production, there are to our knowledge no published reports for the expression and purification of recombinant Z α_1AT.

In this current study we detail our ultimately successful attempts to produce recombinant Z α_1AT. Initially, we used *E. coli* as the expression host, however, despite extensive screening of expression space we were unable to successfully produce monomeric protein. We next attempted expression in the cytosol of the methylotrophic yeast *Pichia pastoris* and were able to purify monomeric and active Z α_1AT. Biochemical characterization of our recombinant Z α_1AT confirmed it possesses the same properties as Z α_1AT purified from patients [22–24]. The availability of recombinant Z α_1AT will be of major benefit for α_1AT deficiency research as it will allow large-scale production and thus the design and screening of therapeutic agents to commence, which otherwise would have been difficult with the limited supply of blood plasma from Z α_1AT patients.

Materials and methods

Materials

All media components were purchased from Merck and Sigma. Enzymes for DNA manipulations were from Promega. Oligonucleotides synthesized by Geneworks (Australia). Complete EDTA-free tablets from Roche were used for protease inhibition during purification. Bovine chymotrypsin and HNE were purchased from Merck and Sigma.

Cloning and mutagenesis of M and Z α_1AT

The M α_1AT gene was cloned into the pLIC-His and pLIC-MBP expression vectors using Ligation Independent Cloning as previously described [25]. The Z mutation was introduced into each expression vector using the Quickchange mutagenesis kit (Stratagene). All constructs were verified by sequencing prior to expression trials. The constructs, vectors and N-terminal fusion tags are listed in Table 1.

Small-scale expression screening of M and Z α_1AT

The expression plasmids were transformed into *E. coli* BL21 (DE3) and plated on 2xTY plates containing 0.1 mg/mL ampicillin. Single colonies were used to inoculate 1 mL cultures of Overnight Express Media (Novagen) with ampicillin (0.1 mg/mL) and incubated shaking (250 rpm) at 30 °C. After 18 h the cells were harvested by centrifugation (10,000*g* for 5 min at 4 °C). Solubility of the α_1AT fusion protein products was assessed using SDS–PAGE on samples from the supernatant and the whole cell lysate (Fig. 1A and B).

Large-scale expression and purification of M and Z MBP-α_1AT fusion proteins

Table 1
Escherichia coli expression constructs used to identify soluble expression of M and Z α_1AT.

	N-terminal tag	Expression	Purification	Solubility following TEV cleavage
M α_1AT	6xHis	S	Excellent	Excellent
	6xHis-MBP	S	Excellent	Excellent
Z α_1AT	6xHis	I	n/a	n/a
	6xHis-MBP	WS	Poor	No

I, insoluble; S, soluble; WS, weakly soluble.

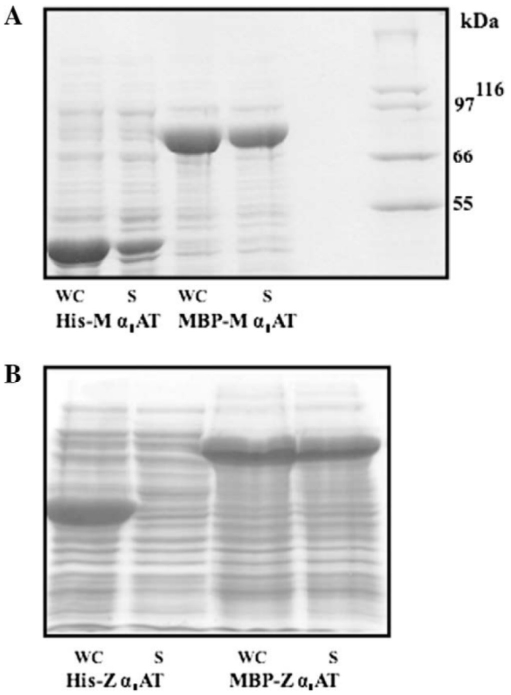


Fig. 1. SDS–PAGE analysis of M and Z α_1AT expression in *E. coli*. M (A) and Z (B) α_1AT and its fusion partner derivatives were expressed as described in Materials and methods. Proteins were analyzed using a 12% polyacrylamide gel and stained with Coomassie blue. Both whole cells (WC) and supernatants (S), following lysis and centrifugation, are shown.

Five-hundred milliliters of Overnight Express Media (Novagen) with ampicillin (0.1 mg/mL) was inoculated with a fresh colony of *E. coli* BL21 (DE3) cells carrying the pLIC-MBP-α_1AT plasmid and incubated shaking (250 rpm) at 30 °C. After 18 h the cells were harvested by centrifugation (10,000*g* for 5 min at 4 °C). For purification of M and Z MBP-α_1AT the cell paste (4 g of wet weight cells) was re-suspended in 30 mL lysis buffer (25 mM NaH_2PO_4, 0.5 M NaCl, 2 mM β-mercaptoethanol, 25 mM imidazole, pH 8.0). The cell suspension was lysed using sonication in the presence of 1 mg/mL lysozyme. The cell debris was removed by centrifugation (48,000*g* for 20 min at 4 °C). The supernatant was then filtered through 0.45 μm filter before loading onto a 1-mL HisTrap HP column (GE Healthcare) equilibrated with lysis buffer. The column was washed with lysis buffer containing 0.5% (v/v) Triton X-100 followed by lysis buffer with 150 mM NaCl. Subsequently, the protein was eluted with 25 mM NaH_2PO_4, 0.15 M NaCl, 2 mM β-mercaptoethanol, 300 mM imidazole, pH 8.0. Fractions containing either M or Z MBP-α_1AT as determined by SDS–PAGE were pooled.

The MBP-α_1AT fusion proteins (at about 2 mg/mL) were subjected to tobacco etch virus (TEV) protease cleavage to separate α_1AT from its fusion solubility tag (MBP). The cleavage was performed overnight at 4 °C in elution buffer containing 0.5 mM EDTA and 1 mM DTT and a final TEV concentration of 0.12 mg/mL [26]. Following incubation the protein sample was diluted with 25 mM NaH_2PO_4, 0.15 M NaCl, pH 8.0, to decrease the imidazole concentration to 25 mM and the solution was loaded onto a 5-mL HiTrap HP column (GE Healthcare) pre-equilibrated with five column volumes of 25 mM NaH_2PO_4, 0.5 M NaCl, 2 mM β-mercaptoethanol, 25 mM imidazole, pH 8.0, for "subtraction" his-tag purification. In this step, the uncleaved MBP-α_1AT, the cleaved MBP and the TEV protease (all three proteins are his-tagged) bind to the resin while the untagged α_1AT flows through. The flowthrough containing α_1AT was adjusted to 50 mM NaCl and loaded onto 5-mL HiTrap Q FF column (GE Healthcare) pre-equilibrated with 50 mM Tris, 50 mM NaCl, pH 8.0. α_1AT was eluted in a 50–500 mM NaCl gradient over eight column volumes. Fractions containing α_1AT were determined by SDS–PAGE and inhibitory activity.

Pichia pastoris expression vector construction

The genes for M and Z α_1AT were isolated from the *E. coli* expression vectors described above and ligated into the pHILD2 vector (Invitrogen) via an intermediate ligation step into the Blunt vector (Invitrogen). The DH5α *E. coli* cell line was used in all transformation reactions.

Yeast transformation and test expression

The *P. pastoris* strain SMD1163 (His4$^-$, pep4, PRB1) was transformed by electroporation of cells treated with lithium chloride. The pHILD2 vector carrying the α_1AT gene was linearized with SacI thus targeting the integration by a single crossover in the chromosomal AOX1 (alcohol oxidase) locus. Transformants were selected on RDB agar plates (Pichia Expression Kit, Invitrogen Life technologies). Plates were incubated at 30 °C for 72 h and the colonies were screened for α_1AT expression by Western blotting using an anti-6xHis antibody, on the cell lysates.

Expression and purification of α_1AT from P. pastoris

The clones with the highest M and Z α_1AT expression levels were used to inoculate 500 mL of yeast peptone dextrose (YPD) media in a 1 L baffled flask. Following incubation at 30 °C shaking (240 rpm) for 30 h the cells were centrifuged (4000 rpm for 15 min) or allowed to settle by gravity over night and then re-suspended in YPM. Cells were harvested after inducing them for 17 h (over night) to 72 h (3 days).

Cells were lysed using glass beads (0.5 mm) in lysis buffer (20 mM Tris, pH 8.0, 500 mM NaCl, 25 mM imidazole, 1 mM β-mercaptoethanol) containing EDTA-free protease inhibitor tablets (Roche). The supernatant was cleared by centrifugation (48,000*g* for 30 min at 4 °C) followed by filtration through a 0.45 μm filter, then loaded onto a 1-mL HisTrap HP column (GE Healthcare). The column was washed with lysis buffer containing 0.5% (v/v) Triton X-100 followed by lysis buffer containing 150 mM NaCl. The protein was then eluted with 20 mM Tris, pH 8.0, 150 mM NaCl, 500 mM imidazole, 1 mM β-mercaptoethanol. Fractions containing protein were combined, diluted 3-fold (to decrease the NaCl concentration to ~50 mM) and loaded onto a 1-mL HiTrap Q-Sepharose column (GE Healthcare). The column was washed with 50 mM Tris, pH 8.0, 75 mM NaCl and α_1AT was eluted with 50 mM Tris, pH 8.0 in a 75–300 mM NaCl gradient over 30 column volumes. Fractions containing active protein were pooled, concentrated and stored at −80 °C (M α_1AT) or 4 °C (Z α_1AT).

Characterization of the inhibitory properties of M and Z α_1AT

The stoichiometries of inhibition (SI) and the association rate constants (k_{ass}) against bovine chymotrypsin and human HNE were determined as previously described [27].

Rate of polymerization

The rate of M and Z α_1AT polymerization was determined using native PAGE analysis at 40 °C in 50 mM Tris, 100 mM NaCl, 5 mM β-mercaptoethanol, 1 mM EDTA, pH 7.4. Aliquots were taken at various time points and added to ice-cold PAGE loading buffer to quench the reaction. The reaction was then analyzed using 10% non-denaturing polyacrylamide gel electrophoresis. Concurrently, polymerization at 40 °C was also followed by monitoring the loss of inhibitory activity as previously described [28].

Rate of 6-mer peptide insertion into M and Z α_1AT

The rate of insertion of the 6-mer peptide FLEAIG into either M or Z α_1AT was assessed by following the formation of the binary complex as previously described [24].

Analytical ultracentrifugation

Sedimentation velocity experiments were performed using the XL-A analytical ultracentrifuge (Beckman/Coulter). Samples were prepared at concentrations of approximately 1 mg/mL in 90 mM NaCl, 50 mM Tris, pH 8.0, and analyzed at 20 °C. M and Z α_1AT were analyzed at a rotor speed of 40,000 rpm and 30,000 rpm (An-50 Ti), respectively, with radial scans measured at 280 nm at 8 min intervals using 0.002 cm increments. Data were fit to a continuous mass distribution model using SEDFIT with baseline absorbance set at zero and corrections were allowed for time and radial independent noise [29]. The molecular mass range was between 10,000 and 100,000 at a resolution of 150.

Results and discussion

Expression and purification of M and Z α_1AT from E. coli

We initially used a parallel screening strategy to identify constructs of M and Z α_1AT that produced soluble protein in *E. coli* [25]. The genes encoding for M and Z α_1AT were cloned into two different *E. coli* expression vectors (Table 1). Both vectors utilized a T7 promoter and possessed an N-terminal hexa-His-tag to aid purification. In the second vector the His-tag was followed by the solubility tag MBP. Small-scale expression screening showed that M and Z α_1AT were produced from all constructs (Fig. 1A and B). Soluble M α_1AT was produced using both vectors (Fig. 1A). The expression profile of the Z α_1AT was different with only Z MBP-α_1AT appearing in the soluble fraction (Fig. 1B). The expression profiles for M and Z α_1AT did not change by decreasing the expression temperature to 18 °C. Therefore, purification of M and Z MBP-α_1AT was attempted.

Expression of M MBP-α_1AT was carried out using 500 mL Overnight Express Media (Novagen). Following induction and lysis the protein was purified using immobilized metal affinity chromatography. The fusion protein was eluted with 300 mM imidazole and was typically ~90% pure as judged by SDS–PAGE. The fusion protein M MBP-α_1AT also displayed inhibitory activity against bovine chymotrypsin. Overnight incubation of the fusion protein with TEV protease resulted in complete cleavage of the fusion protein. After the separation of M α_1AT from His-tagged MBP and TEV, monomeric and active M α_1AT was finally purified via anion exchange chromatography. The total yield from 500 mL Overnight Express Media was 4 mg of M α_1AT. The same production method was applied to Z α_1AT. However, unlike M MBP-α_1AT only a small fraction, approximately 15–20%, of Z MBP-α_1AT bound to the immobilized metal affinity chromatography column with the majority of the fusion protein ending up in the flowthrough. Once eluted the fusion protein was subjected to TEV cleavage which resulted in complete cleavage of the fusion partners. However, unlike with M α_1AT, the TEV digested protein mixture showed no inhibitory activity against bovine chymotrypsin. Subsequent loading onto a immobilized metal affinity chromatography column to remove TEV and MBP and further purification via anion exchange chromatography resulted in fractions containing very low amounts of inactive Z α_1AT. A range of expression temperatures were screened, however, we never succeeded in producing monomeric and active Z α_1AT using *E. coli* as a host.

Therefore, it was decided to use yeast for the expression and purification of monomeric Z α_1AT. Since it is likely that Z α_1AT misfolds in the endoplasmic reticulum and induces the unfolded protein or ER stress response, which depends in part on quality surveillance involving glycosylation, we decided to remove its signal peptide and attempt expression in the cytosol. The cytoplasm also contains a complex set of folding chaperones, and as α_1AT has no disulfide bonds, and it is known that closely related clade B serpins can be produced at high level in yeast cytoplasm [30], we predicted that non-glycosylated M or Z α_1AT would be correctly folded if synthesized in the cytoplasm.

Expression and purification from P. pastoris

The genes encoding for His-tagged M or Z α_1AT were cloned into the pHILD2 vector for cytosolic expression in *P. pastoris*. After pHILD2/α_1AT DNA was linearized with SacI and transformed into the *P. pastoris* strain SMD1163 by electroporation. We subsequently selected 10 clones and confirmed that the α_1AT gene was correctly integrated into the AOX1 locus. We then used methanol to induce recombinant protein expression. The expression levels were detected by Western blot using an anti-6xHis antibody (data not shown). The transformant with the highest expression level was selected for large-scale production of the recombinant protein.

Following large-scale expression (typically 50 g of wet cell mass), the cells were lysed using glass beads (0.5 mm) in lysis buffer and subsequently the his-tagged M and Z α_1AT were purified using two column steps similar to the procedure described above for protein production in *E. coli*. The elution profiles from the Q-Sepharose column, were different for the two α_1AT variants (Fig. 2A). M α_1AT eluted in the main peak with a small amount of cleaved α_1AT eluting in a shoulder of the peak. The SI of the fractions containing intact protein was determined to be 1 indicating that the protein was fully active. The typical yield for M α_1AT was ~0.8 mg/10 g of cells. In contrast, Z α_1AT eluted from the anion exchange resin in two major peaks (Fig. 2B). The first smaller peak contained monomeric and active protein while the second peak consisted of aggregated, inactive Z α_1AT. The ratio of monomeric to aggregated Z α_1AT was typically ~1:2 and the typical yield of active protein was ~0.2 mg/10 g cells. Induction of Z α_1AT expression was also carried out at 21 °C in an attempt to minimize its aggregation, however, no increase in the yield of monomeric Z α_1AT was obtained.

Characterization of Z α_1AT

The structural, inhibitory and polymerization characteristics of recombinant Z α_1AT were analyzed to determine whether it has the same properties as the pathological Z variant purified from patients with α_1AT deficiency. Due to the polymerization-prone nature of Z α_1AT, we first assessed the aggregation status of the purified recombinant proteins using size exclusion chromatography (Superdex 200) and analytical ultracentrifugation (AUC) (Fig. 3). Size exclusion chromatography showed that both proteins eluted in a single peak corresponding to a MW of 45 kDa, suggesting that both recombinant proteins, M and Z α_1AT were folded to a monomeric form with the same hydrodynamic volume. The sedimentation velocity experiment using AUC showed that both proteins, M and Z α_1AT have the same rate of sedimentation in the centrifugal field suggesting that

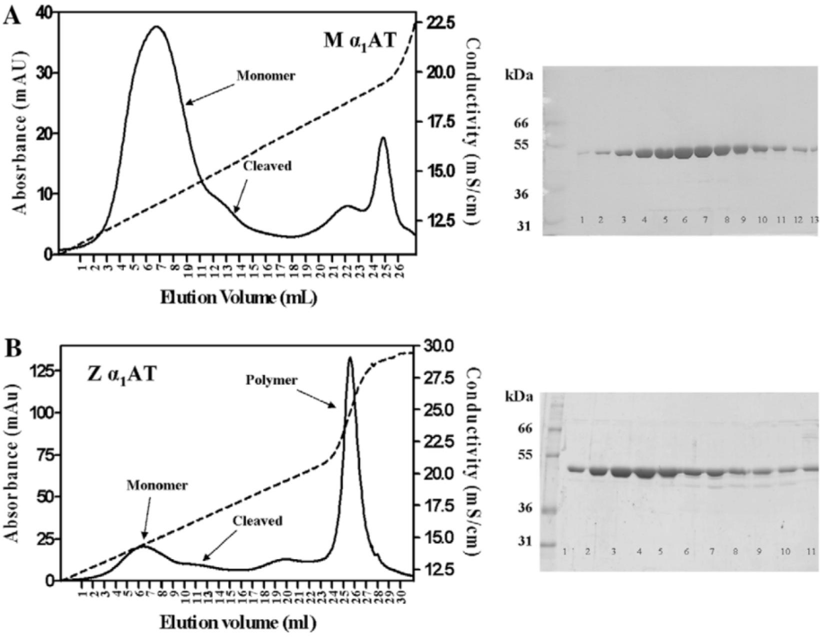


Fig. 2. Purification of M (A) and Z (B) α_1AT. Following expression and initial purification using nickel affinity chromatography, the soluble material was loaded onto a Q-Sepharose column and the protein eluted using a NaCl gradient. The chromatograms and corresponding SDS–PAGE analysis (lanes correspond to fractions across the monomer peak) for both M and Z α_1AT are shown.

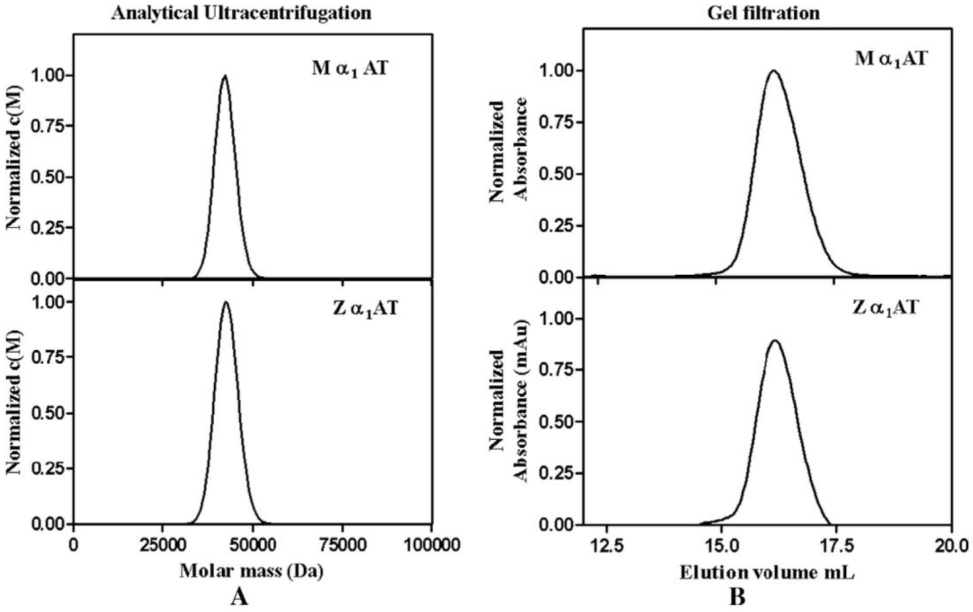


Fig. 3. Aggregation status of recombinant M and Z α_1AT. (A) Analytical ultracentrifugation and (B) size exclusion chromatography (Superdex 200) were used to confirm the monomeric nature and similar masses of both recombinant proteins.

both macromolecules are of the same mass (populated around 45 kDa) and shape [31]. Thus, both techniques confirmed that our two recombinant α_1AT variants have the same hydrodynamic radius corresponding to the monomeric size of α_1AT.

Differential association rates of the FLEAIG peptide with M and Z α_1AT

Due to unknown conformational or/and dynamic differences between Z and M α_1AT, Z α_1AT binds the 6-mer peptide, FLEAIG, much more rapidly than M α_1AT [24]. Our recombinant Z α_1AT bound the 6-mer peptide faster than M α_1AT (Fig. 4) suggesting that our recombinant Z α_1AT has similar structural properties as those previously reported on plasma Z α_1AT.

Inhibitory properties of Z α_1AT

The inhibitory properties of M and Z α_1AT were tested against chymotrypsin and human neutrophil elastase (HNE) (Table 2). α_1AT inhibits serine proteinases via an irreversible suicide substrate mechanism in which the effectiveness of inhibition depends upon the partitioning between the inhibitory and the substrate pathway. This ratio is described by the stoichiometry of inhibition (SI). Our data showed that the SI of Z α_1AT is higher against both proteinases (chymotrypsin and HNE) compared to M α_1AT with ~40% of Z α_1AT following the substrate pathway. Thus, our data agree well with findings from previous studies with Z α_1AT purified from blood plasma [22,23]. The differences in SI values between M and Z α_1AT found by our group are similar to the data reported by Ogushi et al. [22] (30% difference in SI value between M and Z α_1AT for elastase inhibition) but is higher than that reported by Lomas et al. [23] (5% difference in SI value between M and Z α_1AT for chymotrypsin inhibition). The apparent association rate constant, k_{assapp}, for both proteinases was between 2- and 5-fold slower for Z α_1AT than for M α_1AT, which agrees well with the published difference in k_{assapp} for the plasma derived α_1AT variants [22,23]. Further analysis of the inhibition kinetics showed that the association rate constant (k_{ass}), which describes the formation of the complex before the branching into the inhibitory and substrate pathways, is 5-fold faster for M compared to Z α_1AT (Table 2). This suggests that the Z mutation affects the rate limiting formation of the protease–inhibitor complex, thus supporting the findings from the 6-mer peptide binding study [24] that Z α_1AT exists in a non-native conformation.

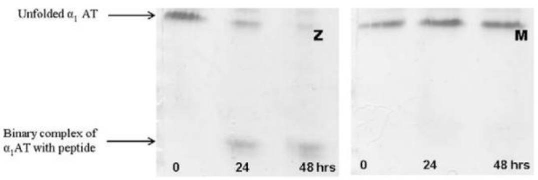


Fig. 4. Binary complex formation of recombinant Z and M α_1AT with the 6-mer peptide FLEAIG. Both recombinant α_1AT variants were incubated with the 6-mer peptide FLEAIG (100× molar excess). Aliquots were taken at either 24 or 48 h and analyzed using 7.5% (w/v) acrylamide 8 M urea gels as described in Materials and methods. The binary complex is stable to unfolding by urea and therefore migrates faster in the electric field.

There are differing reports regarding the stability of the Z α_1AT proteinase complex. Two previous studies [22,23] showed that unlike M, Z α_1AT purified from blood plasma forms unstable complexes with proteases as assessed by SDS–PAGE. However, our results show that both recombinant M and Z α_1AT form stable complexes with both HNE and chymotrypsin. As evident from Fig. 5, there is a band at ~66 kDa due to the α_1AT protease complex for both M and Z α_1AT. Thus, our results match previous findings which also showed the formation of the complex of *in vitro* translated Z α_1AT with elastase by SDS–PAGE [32]. Therefore, it is likely that the differences in complex stability of plasma derived Z α_1AT with proteases are due to glycosylation.

Appendix

Table 2
Inhibitory properties of M and Z α_1AT.

	SI[a]		k_{assapp}[b] ($M^{-1}\ s^{-1}$) ($\times 10^6\ M^{-1}\ s^{-1}$)		k_{ass}[c] ($M^{-1}\ s^{-1}$) ($\times 10^6\ M^{-1}\ s^{-1}$)	
	Chymotrypsin	HNE	Chymotrypsin	HNE	Chymotrypsin	HNE
WT	1.0 ± 0.1	1.1 ± 0.1	0.87 ± 0.08	6.3 ± 0.4	0.96 ± 0.08	6.9 ± 0.4
Z	1.6 ± 0.1	1.8 ± 0.2	0.43 ± 0.01	1.3 ± 0.2	0.68 ± 0.02	2.3 ± 0.4

[a] SI was measured as described in Materials and methods. Average values ± SE from at least five experiments are reported.
[b] Calculated using the total concentration of α_1AT that forms an inhibitory complex with the proteinase as described in Materials and methods.
[c] Calculated using the fractional concentration of α_1AT that forms an inhibitory complex with the proteinase as described in Materials and methods.

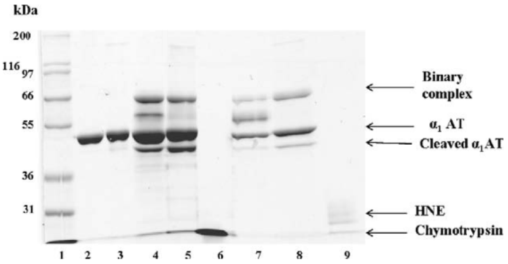


Fig. 5. SDS–PAGE analysis of complex formation between M and Z α_1AT and chymotrypsin and HNE. M and Z α_1AT were mixed with either chymotrypsin or HNE in a molar ratio of 2:1 and allowed to incubate for 30 min at 37 °C, following which the complexes were analyzed by 12% SDS–PAGE. Lane1: MW markers, lane 2: M α_1AT, lane 3: Z α_1AT, lane 4: M α_1AT and chymotrypsin, lane 5: Z α_1AT and chymotrypsin, lane 6: chymotrypsin, lane 7: M α_1AT and HNE, lane 8: Z α_1AT and HNE, lane 9: HNE. The positions of complexed, intact and cleaved α_1AT as well as chymotrypsin and HNE alone are indicated.

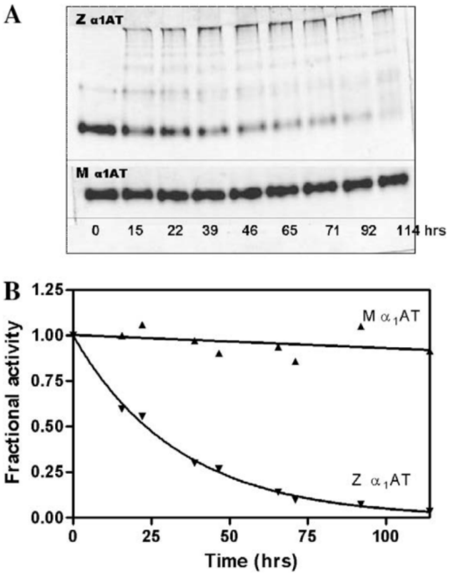


Fig. 6. Polymerization of M and Z α_1AT. (A) M and Z α_1AT (10 μM) were incubated at 40 °C. Samples were removed at the times indicated and rapidly added to ice-cold non-denaturing loading buffer to quench the reaction before electrophoresis. (B) Loss of inhibitory activity against chymotrypsin for M α_1AT (▲) and Z α_1AT (▼). M and Z α_1AT (10 μM) were incubated at 40 °C in a water bath. At various times portions of the protein were removed and placed on ice to prevent further polymerization. The aliquots were subsequently assayed for inhibitory activity against chymotrypsin. The result is representative of five independent experiments.

Characterization of Z α_1AT polymerization

Polymerization of M and Z α_1AT at 40 °C was assessed by native PAGE and loss of inhibitory activity (Fig. 6). These data clearly shows that Z α_1AT polymerizes at a much faster rate than M α_1AT which did not polymerize to a significant extent under the conditions tested. The rate of polymerization, as monitored by loss of inhibitory activity, was determined using a single exponential decay function, which yielded a polymerization rate constant of 0.054 h^{-1} for Z α_1AT. *In vitro* polymerization studies of our recombinant Z α_1AT matched the results of previous investigations on the aggregation properties of glycosylated Z α_1AT purified from blood plasma [23,33] which showed that Z α_1AT polymerizes at physiological temperatures in contrast to M α_1AT which polymerizes at higher, non-physiological temperatures.

Thus, the detailed analysis of the biochemical properties of our recombinant, non-glycosylated, Z α_1AT showed that intracellular cytosolic expression in *P. pastoris* yielded monomeric protein (AUC and size exclusion data) with the same fold (6-mer peptide binding studies) and characteristics (polymerization and inhibitory activity studies) as glycosylated Z α_1AT secreted by human hepatocytes suggesting that yeast cytoplasm possesses the necessary folding machinery to achieve folding of conformationally labile proteins.

Conclusion

Here we have described the method of expression and purification of recombinant monomeric Z α_1AT, the most common pathological variant of human α_1AT. The only source of Z α_1AT known to date is blood plasma of individuals homozygous for the Z α_1AT allele. The main difficulty in producing recombinant Z α_1AT is its high propensity to polymerize, thus yielding only polymeric species if purified from recombinant sources. Our method of expression and purification resulted in the monomeric active form of Z α_1AT. Detailed inhibitory and polymerization characterization confirmed that our recombinant Z α_1AT has the same properties as the one purified from human plasma of patients. An inexhaustible source of the recombinant monomeric Z α_1AT will allow intensified investigations into therapy design and drug screening for the α_1AT deficiency condition.

Acknowledgments

The authors would like to thank members of the lab for useful discussions. This work was funded in part by a Program grant from the NH&MRC. S.P.B. is a NH&MRC Senior Research Fellow. L.D.C. is a NH&MRC CJ Martin Fellow.

References

[1] R.W. Carrell, B. Gooptu, Conformational changes and disease—serpins, prions and Alzheimer's, Curr. Opin. Struct. Biol. 8 (1998) 799–809.
[2] M.K. Chow, D.A. Lomas, S.P. Bottomley, Promiscuous beta-strand interactions and the conformational diseases, Curr. Med. Chem. 11 (2004) 491–499.

[3] B. Gooptu, D.A. Lomas, Polymers and inflammation: disease mechanisms of the serpinopathies, J. Exp. Med. 205 (2008) 1529–1534.

[4] D.A. Lomas, Molecular mousetraps, alpha1-antitrypsin deficiency and the serpinopathies, Clin. Med. 5 (2005) 249–257.

[5] L. Fregonese, J. Stolk, Hereditary alpha-1-antitrypsin deficiency and its clinical consequences, Orphanet J. Rare Dis. 3 (2008) 16.

[6] A.S. Knaupp, S.P. Bottomley, Serpin polymerization and its role in disease—the molecular basis of alpha(1)-antitrypsin deficiency, IUBMB Life 61 (2009) 1–5.

[7] J.K. Stoller, L.S. Aboussouan, Alpha1-antitrypsin deficiency, Lancet 365 (2005) 2225–2236.

[8] J.E. Gadek, G.A. Fells, R.L. Zimmerman, S.I. Rennard, R.G. Crystal, Antielastases of the human alveolar structures. Implications for the protease–antiprotease theory of emphysema, J. Clin. Invest. 68 (1981) 889–898.

[9] S. Gershagen, S. Janciauskiene, ELISA for specific detection of PiZ-related alpha(1)-antitrypsin deficiency, Clin. Chem. 50 (2004) 2407–2410.

[10] D.A. Lomas, B. Parker, Francis lectureship. Antitrypsin deficiency, the serpinopathies, and chronic obstructive pulmonary disease, Proc. Am. Thorac. Soc. 3 (2006) 499–501.

[11] M.J. Dycaico, S.G. Grant, K. Felts, W.S. Nichols, S.A. Geller, J.H. Hager, A.J. Pollard, S.W. Kohler, H.P. Short, F.R. Jirik, et al., Neonatal hepatitis induced by alpha 1-antitrypsin: a transgenic mouse model, Science 242 (1988) 1409–1412.

[12] J.H. Teckman, D. Qu, D.H. Perlmutter, Molecular pathogenesis of liver disease in alpha1-antitrypsin deficiency, Hepatology 24 (1996) 1504–1516.

[13] L.D. Cabrita, S.P. Bottomley, How do proteins avoid becoming too stable? Biophysical studies into metastable proteins, Eur. Biophys. J. 33 (2004) 83–88.

[14] J. Travis, M. Owen, P. George, R. Carrell, S. Rosenberg, R.A. Hallewell, P.J. Barr, Isolation and properties of recombinant DNA produced variants of human alpha 1-proteinase inhibitor, J. Biol. Chem. 260 (1985) 4384–4389.

[15] S.P. Bottomley, S.R. Stone, Protein engineering of chimeric serpins: an investigation into effects of the serpin scaffold and reactive centre loop length, Protein Eng. 11 (1998) 1243–1247.

[16] P.I. Bird, S.C. Pak, D.M. Worrall, S.P. Bottomley, Production of recombinant serpins in *Escherichia coli*, Methods 32 (2004) 169–176.

[17] D. Gilis, H.R. McLennan, Y. Dehouck, L.D. Cabrita, M. Rooman, S.P. Bottomley, In vitro and in silico design of alpha1-antitrypsin mutants with different conformational stabilities, J. Mol. Biol. 325 (2003) 581–589.

[18] H.A. Kang, J.H. Sohn, E.S. Choi, B.H. Chung, M.H. Yu, S.K. Rhee, Glycosylation of human alpha 1-antitrypsin in *Saccharomyces cerevisiae* and methylotrophic yeasts, Yeast 14 (1998) 371–381.

[19] T.K. Huang, M.A. Plesha, B.W. Falk, A.M. Dandekar, K.A. McDonald, Bioreactor strategies for improving production yield and functionality of a recombinant protein in transgenic tobacco cell cultures, Biotechnol. Bioeng. 102 (2009) 508–520.

[20] K.A. McDonald, L.M. Hong, D.M. Trombly, Q. Xie, A.P. Jackman, Production of human alpha-1-antitrypsin from transgenic rice cell culture in a membrane bioreactor, Biotechnol. Prog. 21 (2005) 728–734.

[21] M. Terashima, Y. Murai, M. Kawamura, S. Nakanishi, T. Stoltz, L. Chen, W. Drohan, R.L. Rodriguez, S. Katoh, Production of functional human alpha 1-antitrypsin by plant cell culture, Appl. Microbiol. Biotechnol. 52 (1999) 516–523.

[22] F. Ogushi, G.A. Fells, R.C. Hubbard, S.D. Straus, R.G. Crystal, Z-type alpha 1-antitrypsin is less competent than M1-type alpha 1-antitrypsin as an inhibitor of neutrophil elastase, J. Clin. Invest. 80 (1987) 1366–1374.

[23] D.A. Lomas, D.L. Evans, S.R. Stone, W.S. Chang, R.W. Carrell, Effect of the Z mutation on the physical and inhibitory properties of alpha 1-antitrypsin, Biochemistry 32 (1993) 500–508.

[24] R. Mahadeva, T.R. Dafforn, R.W. Carrell, D.A. Lomas, 6-Mer peptide selectively anneals to a pathogenic serpin conformation and blocks polymerization. Implications for the prevention of Z alpha(1)-antitrypsin-related cirrhosis, J. Biol. Chem. 277 (2002) 6771–6774.

[25] L.D. Cabrita, W. Dai, S.P. Bottomley, A family of *E. coli* expression vectors for laboratory scale and high throughput soluble protein production, BMC Biotechnol. 6 (2006) 12.

[26] L.D. Cabrita, D. Gilis, A.L. Robertson, Y. Dehouck, M. Rooman, S.P. Bottomley, Enhancing the stability and solubility of TEV protease using in silico design, Protein Sci. 16 (2007) 2360–2367.

[27] B.F. Le Bonniec, E.R. Guinto, S.R. Stone, Identification of thrombin residues that modulate its interactions with antithrombin III and alpha 1-antitrypsin, Biochemistry 34 (1995) 12241–12248.

[28] G.L. Devlin, H. Parfrey, D.J. Tew, D.A. Lomas, S.P. Bottomley, Prevention of polymerization of M and Z alpha1-antitrypsin (alpha1-AT) with trimethylamine N-oxide. Implications for the treatment of alpha1-at deficiency, Am. J. Respir. Cell Mol. Biol. 24 (2001) 727–732.

[29] P. Schuck, P. Rossmanith, Determination of the sedimentation coefficient distribution by least-squares boundary modeling, Biopolymers 54 (2000) 328–341.

[30] L.N. Benning, J.C. Whisstock, J. Sun, P.I. Bird, S.P. Bottomley, The human serpin proteinase inhibitor-9 self-associates at physiological temperatures, Protein Sci. 13 (2004) 1859–1864.

[31] J.C. Hansen, J.L. Cole, Analytical ultracentrifugation as a contemporary biomolecular research tool, J. Biomol. Technol. 10 (1999) 163.

[32] M.H. Yu, K.N. Lee, J. Kim, The Z type variation of human alpha 1-antitrypsin causes a protein folding defect, Nat. Struct. Biol. 2 (1995) 363–367.

[33] D.A. Lomas, D.L. Evans, J.T. Finch, R.W. Carrell, The mechanism of Z alpha 1-antitrypsin accumulation in the liver, Nature 357 (1992) 605–607.