ENZYMATIC PROPERTIES OF HEPATITIS C VIRUS NS3 SERINE PROTEASE AND BIO-ENGINEERING OF SERINE PROTEASE INHIBITORS (SERPINS) AGAINST THE NS3 PROTEASE AND ELASTASE

by

ADDY PO
B.Sc., University of British Columbia (Microbiology & Immunology), 2000

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

FACULTY OF GRADUATE STUDIES
(Department of Microbiology and Immunology)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April, 2003

© Addy Po, 2003
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Microbiology and Immunology

The University of British Columbia
Vancouver, Canada

Date March 28th, 2003
Abstract

Hepatitis C virus (HCV) has infected millions of people worldwide and emerged as a global health crisis. HCV NS3 serine protease domain (aa:1027-1218) of the HCV polyprotein, is required for the processing and maturation of the HCV viral non-structural (NS) proteins. As an essential protease for viral replication, HCV NS3 protease has been considered as a strategic target for anti-HCV drug development. This thesis investigates the designs and inhibitory properties of novel serine protease inhibitor (serpin) variants against the HCV NS3 protease. First, the substrate specificity of NS3 protease is evaluated for the purpose of generating protein-based inhibitors to target it. Data from kinetic experiments suggested that NS3 protease activity is greatly enhanced with the addition of its NS4A cofactor and that the three key residues in the substrate necessary for efficient recognition and cleavage by NS3 protease are cysteine in P1, serine/alanine in P'1 and aspartic/glutamic acids in P6 position. This information was used to bio-engineer 5 (serpins) variants based on α1-Antitrypsin (α1-AT) scaffold. Variants of α1-AT contained different mutations in the serpin Reactive Site Loop (RSL) involving the P1, P4, P6 & P’1 positions. The biological activities and the newly gained specificities of the α1-AT and its variants were tested mainly by performing SDS-stable complex formation studies and enzyme kinetic experiments directed against NS3 protease and elastase. Results show that NS3 protease was able to form a 72 kDa SDS-stable complex with α1-AT and 4 of its variants. Taken together, the ability of α1-AT and its variants to interact and form SDS-stable complexes with NS3 protease may represent an entry point for a novel class of protein-based HCV NS3 protease inhibitors.
Table of Contents

Abstract .......................................................................................................................... ii

Table of Contents .......................................................................................................... iii

List of Figures ................................................................................................................ vii

List of Abbreviations ..................................................................................................... ix

Acknowledgements ......................................................................................................... xiii

Chapter 1: Introduction .................................................................................................. 1
  1.1 Hepatitis C virus (HCV): Overview ................................................................. 1
    1.1.1 The infectious agent ....................................................................................... 1
    1.1.2 HCV genome organization and polyprotein processing ............................ 2
    1.1.3 Currently available therapeutics for HCV patients and new targets for drug development .......................................................... 4
  1.2 HCV NS3/NS4A protease ..................................................................................... 5
    1.2.1 Identification of the NS3 protease and Full-length NS3 protease/helicase ................................................................. 5
    1.2.2 The NS4A protease cofactor ........................................................................ 7
    1.2.3 Substrate specificity of HCV NS3/NS4A protease ........................................ 9
      1.2.3.1 IQFS and MCA substrate .................................................................... 10
  1.3 Strategies for inhibition of the NS3/NS4A protease complex ............................ 12
    1.3.1 HCV NS3 protease; an attractive target for anti-HCV therapy ................. 12
    1.3.2 Peptidyl Inhibitors of NS3 protease and their limitations ......................... 13
    1.3.3 Protein-based Inhibitor : The case of serine protease inhibitors (serpins) ................................................................. 14
    1.3.4 α1-antitrypsin (α1-AT) and bio-engineering of serpin reactive site loop (RSL) ..................................................................... 16
1.4 Thesis objectives ............................................................................................................. 20

Chapter 2: Materials and methods ..................................................................................... 21

2.1 Preparation of competent BL21 & DH5α E.coli cells ....................................................... 21

2.2 Transformation of E. coli strain BL21 with NS3 recombinant plasmid construct .......... 21

2.3 Expression of recombinant NS3 proteins ...................................................................... 22

2.4 Purification of NS3 protease using Fast Protein Liquid Chromatography (FPLC) ........ 23

2.5 Protein analysis with SDS PAGE and Western Blot ............................................................ 23

2.6 Development of HCV NS3 protease assays .................................................................... 25

2.7 Michaelis-Menten steady-state analysis of HCV synthetic peptide substrate .......... 26

2.8 α1-AT DNA isolation and sub-cloning ............................................................................ 27

2.9 Site-directed mutagenesis of α1-AT using Polymerase chain reactions (PCR) ............ 28

2.10 Expression and purification of α1-AT and its variants .................................................. 31

2.11 Pancreatic elastase assays ............................................................................................... 32

2.12 Quantification of proteins ............................................................................................... 32

2.13 Complex formation studies with α1-AT and its mutant variants against elastase, NS3 protease and FLNS3 .......................................................... 33

2.14 Inhibition assay of elastase and NS3 protease by α1-AT and all its variants ................. 33

2.15 Model structure of NS3 protease/NS4A complex, Helicase/NTPase and α1-AT/elastase complex structure ........................................................... 35

2.16 List of buffer solutions ................................................................................................. 36
Chapter 3: Expressions, purification and characterization of NS3 protease using internally quenched fluorogenic substrates (IQFS)......................... 37

3.1 Expression of recombinant NS3 protease in pLys BL21 E.coli.................... 37

3.2 Purification of recombinant NS3 protease from E.coli supernatant.............. 38

3.3 Optimization of assay conditions for detection of NS3 protease activity............................................................... 41

3.4 Probing the substrate specificity of NS3 protease using HCV IQFS.............. 43

3.5 Probing the substrate specificity of NS3 protease using GBV IQFS.............. 48

3.6 Summary................................................................................. 51

Chapter 4: Bio-engineering, expression & purification of α1-AT variants (mutants A-E).................................................................................. 53

4.1 Sub-cloning of α1-AT from pDS-56 vector to pET-21 vector......................... 53

4.2 Bio-engineering of α1-AT variants..................................................... 56

4.3 Purification of α1-AT and its variants.................................................. 60

4.4 Bio-engineering of KDEL mutant A & D............................................ 63

Chapter 5: Characterization of the ability of α1-AT and RSL variants to form complexes with and inhibit the proteolytic activity of elastase.............. 65

5.1 α1-AT, a natural inhibitor of neutrophil and pancreatic elastase................. 65

5.2 Complex formation between α1-AT, Mutant A & D with elastase.............. 65

5.3 Degradation of Mutant B, C & E by elastase........................................ 68

5.4 Inhibition assay of α1-AT and its variants towards elastase..................... 72

5.5 Titration and Progress curve analysis of α1-AT and its variants towards elastase................................................................. 75
Chapter 6: Characterization α₁-AT and its variants (Mut A-E) with complex formation studies and inhibition assays against HCV NS3 protease

6.1 Serpins as candidates for novel protein therapeutics against HCV NS3 protease

6.2 Complex formation between α₁-AT and its variants against NS3 protease

6.3 Studies investigating the effect of enzyme: inhibitor ratio and time of interaction upon complex formation between α₁-AT and variants and NS3 protease

6.4 Inhibition assay of α₁-AT and its variants towards NS3 protease

6.5 Complex formation between α₁-AT and its variants against FLNS3

6.6 Sub-cellular targeting of recombinant serpin-based inhibitors of HCV NS3 protease in human UNS 3/4A cells

6.7 Summary

Chapter 7: Conclusion

7.1 Current view of HCV NS3 serine protease inhibitors

7.2 α₁-Antitrypsin and its variants as potential inhibitors against HCV NS3 protease and future directions

7.3 Animal models and cellular systems for anti-HCV drugs development

7.4 Combination therapy for HCV infection

Chapter 8: References

Chapter 9: Appendix
List of Figures

Figure 1.1  HCV Genome/Polyprotein processing ............................................. 3
Figure 1.2  Ribbon representation of FLNS3 & NS3 protease ................................. 6
Figure 1.3  NS4A: HCV NS3 protease co-factor ............................................. 8
Figure 1.4  Serine protease inhibitor (Serpin) .................................................... 17
Figure 1.5  Conformational polymorphism of inhibitory serpins ......................... 18
Figure 3.1  Purification of bacterially expressed histidine-tagged NS3 protease ......... 39
Figure 3.2  NS3 protease activity titration with Pep4A21-34 ................................ 42
Figure 3.3  The intramolecularly quenched fluorogenic substrate (IQFS) are suitable for continuous monitoring of HCV NS3/4A protease complex activity .......... 44
Figure 3.4  Substrate specificity of NS3 protease ............................................ 46
Figure 3.5  NS3/Pep4A protease complex activities on peptide substrates corresponding to trans-cleavage sites ................................................................. 47
Figure 3.6  NS3/Pep4A protease complex specificities towards GBV-A and GBV-B IQFS ............................................................................................................. 49
Figure 3.7  Ribbon representation of the catalytic pocket of NS3/Pep4A protease Complex ........................................................................................................ 50
Figure 4.1  Subcloning of $\alpha_1$-AT from pDS-56 to pET-21 ........................................ 55
Figure 4.2  $\alpha_1$-AT and bio-engineered variants .............................................. 58
Figure 4.3  Ribbon representations of the backbone of the crystal structure of $\alpha_1$-AT wild-type and the reactive site loops of its variants mutant (A-E) .............. 59
Figure 4.4  Purification of bacterially expressed histidine-tagged $\alpha_1$-AT and its variants .............................................................................................................. 61
Figure 4.5  SDS-PAGE analysis of $\alpha_1$-AT and all its variants (Mut A-E) ............... 62
Figure 5.1  Serpin complex formation with elastase ............................................. 67
Figure 5.2  Branched pathway mechanism of serpins as suicide substrate inhibitors ... 68
Figure 5.3  Degradation of Mut B, C & E by elastase .................................................. 70
Figure 5.4  Residual activity of elastase with α₁-AT after different incubation time .... 73
Figure 5.5  Inhibition of elastase by α₁-AT, Mut A & D ............................................. 74
Figure 5.6  Inhibition constant of α₁-AT, Mut A and D vs elastase ........................... 76
Figure 5.7  Progress curve of α₁-AT, Mut A & D vs elastase .................................... 78
Figure 6.1  Complex formation between α₁-AT and its variants vs NS3 protease ...... 81
Figure 6.2  Complex formation between α₁-AT, Mut A, D & C with increasing [NS3] ........................................................... 84
Figure 6.3  Complex formation between α₁-AT, Mut A, D & C with NS3 protease at different time points ........................................... 85
Figure 6.4  Inhibition of NS3 protease by α₁-AT and its 5 variants ............................. 87
Figure 6.5  Complex formation between FLNS3 with α₁-AT and all its variants ......... 89
Figure 6.6  Co-localization of α₁-AT, Mut A-KDEL & Mut D-KDEL with NS3 in human cells ................................................................. 90
Appendix-1  Tight binding titration of pancreatic elastase by α₁-AT ......................... 110
Appendix-2  Restriction map of pET-21 (Novagen) ............................................... 111
Appendix-3  Restriction map of p-CMV shuttle vector (Stratagene) .......................... 112
Appendix-4  Clustal W alignment of α₁-AT and all its variants (amino-acids) ............ 113
Appendix-5  Clustal W alignment of α₁-AT and all its variants (nucleotides) ............. 115
Appendix-6  Clustal W alignment of α₁-AT-KDEL and Mut A-KDEL & Mut D-KDEL (amino-acids) .................................................. 119
Appendix-7  Clustal W alignment of α₁-AT-KDEL and Mut A-KDEL & Mut D-KDEL (nucleotides) .................................................. 120
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>Amino acid</td>
</tr>
<tr>
<td>Abz</td>
<td>Anthranillic Acid (O-Aminobenzoic acid)</td>
</tr>
<tr>
<td>Amp</td>
<td>Ampicillin</td>
</tr>
<tr>
<td>α₁-AT</td>
<td>α₁-antitrypsin</td>
</tr>
<tr>
<td>α₂AP</td>
<td>α₂ antiplasmin</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>BVDV</td>
<td>Bovine viral diarrhea virus</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>NS3</td>
<td>Non-Structural 3 protease (protease domain)</td>
</tr>
<tr>
<td>eDDnp</td>
<td>Ethylene Diamine 2,4, Dinitrophenyl</td>
</tr>
<tr>
<td>E</td>
<td>Enzyme</td>
</tr>
<tr>
<td>[E]</td>
<td>Enzyme concentration</td>
</tr>
<tr>
<td>EI*</td>
<td>Enzyme/Inhibitor complex</td>
</tr>
<tr>
<td>ECL</td>
<td>Electro-chemoluminescence</td>
</tr>
<tr>
<td>E.coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine tetra-acetic acid</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>FLNS3</td>
<td>Full length NS3 protease (protease/helicase domains)</td>
</tr>
<tr>
<td>FPLC</td>
<td>Fast Protein Liquid Chromatography</td>
</tr>
<tr>
<td>GBV-A</td>
<td>GB virus A</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>GBV-B</td>
<td>GB virus B</td>
</tr>
<tr>
<td>HAART</td>
<td>Highly Active Anti-Retroviral Therapy</td>
</tr>
<tr>
<td>HCMV</td>
<td>Human Cytomegalovirus virus</td>
</tr>
<tr>
<td>HCV</td>
<td>Hepatitis C virus</td>
</tr>
<tr>
<td>HEPES</td>
<td>(N-[2-hydroxyethyl] piperazine-N'-[2-ethane sulphonic acid])</td>
</tr>
<tr>
<td>HIV</td>
<td>Human Immunodeficiency virus</td>
</tr>
<tr>
<td>HTA</td>
<td>High Throughput Assay</td>
</tr>
<tr>
<td>I</td>
<td>Inhibitor</td>
</tr>
<tr>
<td>I°</td>
<td>Cleaved Inhibitor</td>
</tr>
<tr>
<td>IFN-α</td>
<td>Interferon α</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-thiogalactosidase</td>
</tr>
<tr>
<td>IQFS</td>
<td>Internally Quenched Fluorogenic Substrate</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal Ribosome Entry Sites</td>
</tr>
<tr>
<td>IV</td>
<td>Intravenous</td>
</tr>
<tr>
<td>Kan</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>kbp</td>
<td>Kilobase pair</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilo Dalton</td>
</tr>
<tr>
<td>K_{ass}</td>
<td>Association constant</td>
</tr>
<tr>
<td>k_{cat}</td>
<td>Catalytic constant</td>
</tr>
<tr>
<td>K_{d}</td>
<td>Dissociation constant</td>
</tr>
<tr>
<td>K_{i}</td>
<td>Inhibition constant</td>
</tr>
<tr>
<td>K_{m}</td>
<td>Affinity constant</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani medium</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Acronym</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>MCA</td>
<td>Methyl Coumarinamide Substrate</td>
</tr>
<tr>
<td>ml</td>
<td>milli litre</td>
</tr>
<tr>
<td>mm</td>
<td>milli metre</td>
</tr>
<tr>
<td>mM</td>
<td>milli molar</td>
</tr>
<tr>
<td>Mut A</td>
<td>Mutant A</td>
</tr>
<tr>
<td>Mut B</td>
<td>Mutant B</td>
</tr>
<tr>
<td>Mut C</td>
<td>Mutant C</td>
</tr>
<tr>
<td>Mut D</td>
<td>Mutant D</td>
</tr>
<tr>
<td>Mut E</td>
<td>Mutant E</td>
</tr>
<tr>
<td>µl</td>
<td>micro litre</td>
</tr>
<tr>
<td>µM</td>
<td>micro molar</td>
</tr>
<tr>
<td>nM</td>
<td>nano molar</td>
</tr>
<tr>
<td>(NS)</td>
<td>Non-Structural</td>
</tr>
<tr>
<td>NRTIs</td>
<td>Nucleoside reverse transcriptase inhibitors</td>
</tr>
<tr>
<td>NNRTIs</td>
<td>Non-nucleoside reverse transcriptase inhibitors</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>Pep4A</td>
<td>Peptide HCV NS4A cofactor</td>
</tr>
<tr>
<td>pH</td>
<td>Potential of hydrogen</td>
</tr>
<tr>
<td>pM</td>
<td>pico Molar</td>
</tr>
<tr>
<td>RdRp</td>
<td>RNA-dependent RNA polymerase</td>
</tr>
<tr>
<td>RFU</td>
<td>Relative Fluorescence Unit</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
</tbody>
</table>
RSL

SDS-PAGE

Serpin

$S_i$

[S]

TAE

TBS

Trigon-X-100

UV

$V_{max}$

[Vo]

Y(NO$_2$)

**Amino acids**

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>One Letter</th>
<th>Three Letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine (Ala)</td>
<td>A</td>
<td>Leucine (Leu)</td>
</tr>
<tr>
<td>Arginine (Arg)</td>
<td>R</td>
<td>Lysine (Lys)</td>
</tr>
<tr>
<td>Asparagines (Asn)</td>
<td>N</td>
<td>Methionine (Met)</td>
</tr>
<tr>
<td>Aspartate (Asp)</td>
<td>D</td>
<td>Phenylalanine (Phe)</td>
</tr>
<tr>
<td>Cysteine (Cys)</td>
<td>C</td>
<td>Proline (Pro)</td>
</tr>
<tr>
<td>Glutamate(Glu)</td>
<td>E</td>
<td>Serine (Ser)</td>
</tr>
<tr>
<td>Glutamine (Gln)</td>
<td>Q</td>
<td>Threonine (Thr)</td>
</tr>
<tr>
<td>Glycine(Gly)</td>
<td>G</td>
<td>Tryptophan (Trp)</td>
</tr>
<tr>
<td>Histidine (His)</td>
<td>H</td>
<td>Tyrosine (Tyr)</td>
</tr>
<tr>
<td>Isoleucine (Ile)</td>
<td>I</td>
<td>Valine (Val)</td>
</tr>
</tbody>
</table>

**Nucleotide Bases**

<table>
<thead>
<tr>
<th>Nucleotide Base</th>
<th>Letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenine</td>
<td>A</td>
</tr>
<tr>
<td>Cytosine</td>
<td>C</td>
</tr>
<tr>
<td>Guanine</td>
<td>G</td>
</tr>
<tr>
<td>Thymine</td>
<td>T</td>
</tr>
<tr>
<td>Uracil</td>
<td>U</td>
</tr>
</tbody>
</table>
Acknowledgments

I would first like to thank my supervisor Dr. François Jean and supervisory committee Dr. Robert Hancock and Dr. Jim Kronstad for their support throughout my graduate school program at UBC. Not only was their guidance instrumental in the completion of this work, but they also provided me with the autonomy to develop my abilities and confidence as a scientist. I would also like to thank Dr. Luiz Juliano, Dr. Nobuko Kakiuehi and Dr. Carl Hashimoto for the IQFS, NS3 protease construct and α1-AT-KDEL construct.

I am especially grateful to the anti-viral lab (Pamela, Morgan, Martin, Rebecca, Jen, Laura, and Jessie) for their knowledge and friendship. Special thanks to Pamela Hamill for her constant brilliant suggestions in every aspect of the lab, Martin Richer for aiding in technical advise and the cloning of several mutants and Morgan Martin for input with her unconventional “outside the box” ideas. I would also like to thank Helmut Kae for his technical support and friendship.

Financially, I would like to thank CIHR for two years of monetary support and to “Mom and Dad” for their constant supply of delicious food and a comfortable roof to stay in. During my time as a graduate student, many friends including my long time buddies have been supportive and injected much needed fun to balance my life. I would like to thank those friends once again for all their support and friendship over the years; and best of all, guys, it is finally time to join you all in the real world!!!!

Last, but certainly not least, I am especially grateful to my family and Sandy whose love, encouragement and unconditional support has been constant in my life. Your support has always been cherished.
Chapter 1: Introduction

1.1 Hepatitis C virus (HCV): Overview

1.1.1 The Infectious Agent

As recently as the late 1980s, few people other than physicians had heard of the Hepatitis C virus (HCV), a slowly progressing viral infection that over a couple of decades can lead to liver cancer (Di Bisceglie and Bacon, 1999). HCV is the causal agent for a largely chronic liver infection afflicting an estimated 170,000,000 people world-wide and is the primary cause of liver transplantation (Rosenberg, 2001). In general, HCV establishes persistent infections by inducing limited direct cell damage and/or by evading immunologic surveillance. However, the lack of serious direct cell injury allows the virus the opportunity for continued viral replication and establishment of chronic infection. About 20% of the infected individuals develop liver cirrhosis and hepato-cellular carcinoma after 20 years (Alter and Seef, 2000). HCV viral infection leads to cellular damage in vivo through two mechanisms; namely, direct cytopathicity, which is the result of the toxic actions of viral products on infected cells, and immune mediated injury, which is the result of cell lysis of viral-infected cells by either direct lymphocyte cytotoxicity, antibody-mediated damage or viral-induced autoimmunity (Gonzalez-Peralta et al., 1999). Methods of common transmission of HCV include: blood transfusion, hemodialysis, organ transplantation, needle sharing in IV drug users and tattooing (Caronia et al., 2001). Although many individuals are chronically infected, many more are newly infected carriers who are unaware of their infection, which could facilitate the spread of the virus (Fanning et al, 2000). Despite the
seriousness of the disease, so far there is no vaccine or effective antiviral therapy available (Lechmann and Liang, 2000).

### 1.1.2 HCV Genome Organization and Polyprotein Processing

HCV belongs to the Flaviviridae family, genus hepacivirus. It is an enveloped virus containing a single-stranded positive-sense RNA genome (9.6 kb) which encodes a single polyprotein of about 3100 amino acids (Bartenschlager., 1997). This polyprotein includes both structural and non-structural (NS) proteins which are cleaved into functional protein products by both cellular and viral proteases (Fig. 1.1 A). Structural proteins incorporated into mature virions are located at the N-terminus of the HCV polyprotein. The first product to be cleaved from the polyprotein is the highly basic core protein which forms the major constituent of the nucleocapsid (Yasui et al., 1998). E1 and E2 are highly glycosylated type I transmembrane proteins which have been shown to form a stable heterodimeric complex (Deleersnyder et al., 1997). Protein p7, located at the carboxy terminus of E2, is a highly hydrophobic polypeptide of unknown function. All HCV structural proteins are cleaved by an unknown host enzyme. The non-structural (NS) proteins are required for viral replication. They include NS2 metallo-protease, which catalyses the cleavage between NS2 and NS3 (Grakoui et al., 1993). NS3 is a bi-functional protein possessing in the amino-terminal ~190 residues, a chymotrypsin-like serine-type protease responsible for cleavage at the NS3/4A, NS4A/4B, NS4B/5A and NS5A/5B sites. The carboxyl-terminal of NS3 contains the NTPase/helicase activities essential for the translation and replication of the HCV genome (Bartenschlager et al., 1993). NS4A is an essential cofactor for the activity of NS3 protease and is required for efficient polyprotein processing (Failla et al., 1994). The function of
A Genome Proprotein Function

Nucleocapsid Core protein
Envelope protein
Envelope protein
Unknown
Metallo-protease
Serine Protease
Helicase
Protease co-factor
Unknown
Unknown
RNA-dependent
RNA polymerase

B Protease Helicase NS3 Domain
Protease NS3 protease – (191 aa)
Histidine-tag
Protease Helicase FL NS3 – (630 aa)

Figure 1.1. HCV Genome/Polyprotein processing. (A) HCV genome organization and virus encoded structural and non-structural proteins. Dotted arrows indicate junctions processed by host proteases, an orange arrow indicates the junction processed by the metallo-protease, solid black arrows indicate junctions processed by the NS3 protease. (B) Histidine-tagged NS3 variants: NS3 protease (191 aa) and FLNS3 (Full length NS3) (630 aa).
NS4B is so far unknown. NS5A is a highly phosphorylated protein the function of which is also unknown. NS5B has been identified as the RNA-dependent RNA polymerase (RdRp) required for viral genome replication (Al et al., 1998) (Fig. 1.1 A).

1.1.3 Currently Available Therapeutics for HCV Patients and New Targets for Drug Development

The development of effective HCV therapeutics has been seriously hampered by the lack of an efficient in vitro HCV replication system. Moreover, until the recent discovery of the mouse model which permits HCV replication (Mercer et al., 2001), chimpanzees had been the only animal shown to be permissive to HCV infection. At present, Interferon alpha (IFN-α) or pegylated interferon alone or in combination with oral ribavirin, are the only treatments currently approved for chronic Hepatitis C. Combination therapy (IFN + Ribavirin) is rapidly superseding interferon monotherapy because dual therapy is significantly more effective (Foster et al., 2001). The effect of IFN-α may include inhibition of HCV virion production by decreasing the synthesis of viral mRNA and proteins. Ribavirin is a nucleoside analogue which inhibits RNA-dependent RNA polymerase and decreases GTP availability in the cell to cause a reduction in viral protein synthesis (Klaus et al., 2000). However, even the combination therapy for chronic HCV infection is clearly unsatisfactory, as only a small proportion of patients (40%) respond to therapy and the side effects are considerable (Heathcote et al., 2000). Various other compounds have also been tried, including ursodeoxycholic acid, nonsteroidal anti-inflammatory agents, mycophenolate and amantadine. At present, none of these drugs have been shown to have significant advantages over the currently used therapy of IFN + Ribavirin (Craxi et al., 1999). The development of
vaccine against HCV also faces a variety of obstacles mainly due to the rapid mutation rate of HCV (Lechmann and Liang., 2000). The importance of NS3 protease in HCV replication (Major et al., 1999) has led to many studies attempting to understand its structure and function, with the ultimate goal being the design of novel anti-HCV therapeutics.

1.2 HCV NS3/NS4A Protease

1.2.1 Identification of the NS3 Protease and Full-length NS3 Protease/Helicase

There are two molecular forms of NS3 which have been widely studied so far: Truncated NS3 protease domain and Full length NS3 (FLNS3) (Fig 1.1 B). HCV NS3 protease consists of approximately 190 amino acids which contain the serine protease domain. Analysis of the X-ray crystal structures of the truncated NS3 protease revealed that the serine protease adopts a chymotrypsin-like fold and the NS3 protease active site is similar to the active site of the other chymotrypsin-like enzymes. The nucleophilic Ser-139, together with the general acid/base catalyst His-57 and Asp-81, form the NS3 catalytic triad (Fig 1.2 B) (Kim et al., 1996). FLNS3 consists of the serine protease domain and the helicase/NTPase domain, an enzyme that can unwind double-stranded regions of RNA or remove regions of secondary structure in an ATP-dependent reaction, allowing the RNA dependent RNA polymerase to copy the positive and negative strands (Shoji et al, 1999; Kim et al, 1998; Yan et al., 1998; Love et al., 1996). The NS3 helicase domain (440 amino acids) is about twice as large as the protease domain (190 amino acids) (Fig. 1.2 A) (Gallinary et al., 1999). Both the NS3 protease domain itself and the FLNS3 have been shown to be catalytically active in cleaving synthetic peptides mimicking the amino acid sequence of the junctions.
Figure 1.2. Ribbon representation of FLNS3 & NS3 protease. (A) (FLNS3)-NS3/Pep4A protease complex (yellow/purple) and Helicase/NTPase (cyan). (B) Ribbon representation of the backbone of the crystal structure of NS3/Pep4A protease complex; NS3 protease (yellow); NS4A-derived sequence (purple).
between the non-structural proteins of HCV (Fattori et al., 2000; Gallinari et al., 1998; Po et al., 2001). The role of NS3 in the processing and maturation of the non-structural junction NS3/4A, NS4A/NS4B, NS4B/NS5A and NS5A/NS5B of the viral polyprotein has been elucidated by transient transfection and cell-free translational studies (Bartenschlager et al., 1993; Grakoui et al., 1993; Lin et al., 1994). From these studies it emerged that the NS3/4A site is cleaved in cis and the rest of the junctions are cleaved in trans.

1.2.2 The NS4A Protease Cofactor

Although NS3 protease and FLNS3 show some intrinsic proteolytic activity in vitro, a second viral protein, NS4A, is an essential NS3 protease cofactor for efficient proteolytic processing of the HCV polyprotein (Failla et al., 1994; Bartenschlager et al., 1994; Lin et al., 1994; Tanji et al., 1995; Po et al., 2001). NS4A cofactor is a membrane-bound protein of 54 residues. Region I, residues 1-20, is a highly hydrophobic trans-membrane helix. Region II, residues 21-34, is also hydrophobic, but with a β-strand structure. Region III, residues 35-54, is more hydrophilic with an α-helical conformation (Fig 1.3. A) (Tomei et al., 1996). NS4A can exert its cofactor function in cis (i.e., when expressed as an NS3-NS4A precursor protein) as well as in trans (i.e., when expressed as a separate molecule to interact with NS3 to cleave other NS polyproteins). The interaction between the NS4A cofactor and NS3 causes the re-arrangement of the NS3 catalytic triad (His-57, Asp-81 and Ser-139) residues, which are important in forming the catalytic pocket of NS3, thereby enhancing the proteolytic activity of NS3 on all cleavage sites (Landro et al., 1997). In addition, the complex formation of the full length NS4A cofactor to NS3 also significantly increases the stability of NS3 in the cytoplasm of cultured mammalian cells as well as targets the NS3
Figure 1.3. NS4A: HCV NS3 protease co-factor (A) Representation of NS3 and full length NS4A cofactor complex attached to the ER of the cell. Predicted Structures I: Hydrophobic trans membrane α-helix, II: Hydrophobic β-sheet, III: Hydrophillic α-helix. Synthetic peptide cofactor only contain domain II with its amino acids sequence shown in red. (B) Ribbon representation of synthetic peptide cofactor docking in NS3.
protein to the membranes of the endoplasmic reticulum (ER) (Tanji et al., 1995). The production of the full length NS4A cofactor (54 amino acids) for the activation of NS3 in vitro is difficult as full length NS4A is highly hydrophobic and insoluble. However, it has been demonstrated that synthetic 17-mer peptides corresponding to the central domain of NS4A are as efficient as full length NS4A in activating the protease activity of NS3 in vitro. Consequently, such peptides have been utilized in vitro studies of NS3 protease activity (Fig. 1.3. B) (Landro et al., 1997; Lin et al., 1995; Shimizu et al., 1996; Po et al., 2001). Synthetic peptide cofactor has been found to bind to NS3 protease with 1:1 stoichiometry (Lin et al., 1995).

1.2.3 Substrate Specificity of HCV NS3/NS4A Protease

Substrate recognition by the chymotrypsin-like serine proteases involves the binding of the substrate on the active surface of the enzyme defined by a specific amino acid sequence. Elucidating the substrate specificity of the HCV NS3 protease is important for the development of high throughput assays for screening of potential HCV protease inhibitors and for the rational design of the HCV protease-specific inhibitors. Sequence comparison among isolated HCV strains has revealed structurally conserved residues flanking the cleavage sites between the NS3, NS4A, NS4B, NS5A and NS5B (Grakoui et al., 1993). Investigations of protease specificity have generally focused on the P1/S1 interaction (based on the nomenclature of Schecter and Berger (1967), where P1-P'1 denotes peptide residues on the acyl and leaving group side of the scissile bond, respectively. The adjacent peptide residues are numbered outward, while S1, S'1, etc. denote the corresponding enzyme binding sites), followed by consideration of the P2-Pn/S2-Sn interactions.
The primary specificity of a protease is defined by the side chain of the amino acid that precedes the scissile bond, i.e. the P1 position (De Francesco and Steinkuhler, 1999). In the case of NS3 protease, it was found that a Cys/Thr residue in P1, Asp/Glu at the P6 residue and Ser/Ala at the P'1 residue are required for the efficient cleavage by NS3 protease (Fig 3.4.B), with cleavage occurring after cysteine in all trans cleavage sites (i.e. NS4A/4B, NS4B/5A, and NS5A/5B) or after threonine in the intramolecular cleavage site between NS3 and NS4A. However, acidic residue in P6 is not a stringent requirement (Urbani et al., 1997). Preference for cysteine residues in the P1 positions of the NS3 substrates can be rationalized on the basis of the peculiar structure of the S1 pocket of the protease.

The substrate specificity of the NS3/Pep4A protease complex was studied in detail, using mainly purified recombinant enzyme and synthetic peptide substrate. For efficient activity on peptide substrates, the NS3/Pep4A protease complex requires at least a decamer peptide substrate spanning P6-P'4 (Steinkuhler et al., 1996). Although several research groups have studied the substrate specificity of NS3 protease, none have undertaken a comprehensive study comparing the substrate specificity of both the NS3 protease and the FLNS3 with the same set of substrates (NS4A/4B, NS4B/5A, NS5A/5B) and assay conditions using a high-throughput approach (HTA). Since the exact mechanism of HCV replication in vivo is still largely unknown, this comprehensive substrate specificity study with the same set of substrates is important because potential inhibitors of the HCV NS3 must be able to inhibit both the NS3 protease and the FLNS3.
1.2.3.1 IQFS and MCA Substrates

Most substrate specificity studies with HCV NS3/Pep4A protease complex were performed and determined with HPLC. Our study used internally quenched fluorogenic substrate (IQFS) and methyl coumarinamide substrate (MCA) containing a cysteine residue in P1 position in order to perform continuous monitoring of protease activity with high throughput assays (HTA). Screening of the substrate specificity of the purified recombinant NS3 protease. MCA substrates contain only 4-6 amino acids in total and in order to be fluorescently active, the junction between the final amino acid and the fluorophore has to be cleaved.

Internally quenched fluorogenic substrates (IQFS), which include 10-12 amino acids including the enzymatically cleaved bond, are characterized by the presence of a donor fluorescent group (Abz) positioned at the N-terminal end of the substrate and a quencher group (EDDnp) positioned at the C-terminal end (Yaron et al., 1979; Jean et al., 1995). The fluorescence energy of an intact substrate is initially quenched by an intra-molecular energy...
transfer between the donor and acceptor. However, upon cleavage of the substrate, the fluorescence energy is recovered, thereby allowing the increase in fluorescence over time to be measured as cleavage occurs (Yaron et al., 1979). Using a fluorescent spectrofluorometer equipped with a 96 well plate reader, it is therefore possible to measure the amount of increase in fluorescence in many different reactions at once. The results of such a kinetic study can then be analyzed further to obtain the $k_{cat}$, $V_{max}$ and $k_m$ values from the different reactions that will allow definition of the substrate specificity of the NS3 protease and FLNS3 in vitro.

1.3 Strategies for Inhibition of the NS3/NS4A Protease Complex

1.3.1 HCV NS3 Protease is an Attractive Target for Anti-HCV Therapy

There are four major types of protease enzymes (aspartic, serine, cysteine and metallo) which selectively catalyze the hydrolysis of polypeptide bonds (Leung et al., 2000). HCV NS3 protease, which belongs to the serine protease super family of chymotrypsin like enzymes, mediates the maturation cleavage of viral polyproteins. Virally encoded proteases are attractive targets for drug discovery since they are crucial in the life cycle of many viruses, including HCV (Bianchi and Pessi, 2002). Currently, there are a few designed potent and selective protease inhibitors that slow or halt disease progression. A well known example of a potent viral protease inhibitor is the aspartyl protease inhibitor of the human immunodeficiency virus (HIV). The HIV-1 protease has proven to be an attractive drug target due to its essential role in the replicative cycle of HIV (Campoy et al., 2000). Several low molecular weight inhibitors of HIV-1 protease including saquinavir, ritonavir, indinavir, nelfinavir and amprenavir are now used to treat HIV infections in humans (Leung et al.,
2000). These drugs are among the successful examples of drugs developed to target the proteolytic active site of the HIV-1 protease. The concept of inhibiting the viral protease as a means of stopping viral replication can potentially be applied to many other viruses, including HCV.

The most intensively studied and therefore best understood target for antiviral therapy against HCV is the NS3 protease. Previous studies have shown that the NS3 protease is responsible for most of the proteolytic maturation events within the non-structural portion of the viral polyproteins (Grakoui A et al., 1993). Studies carried out with HCV and other members of the Flaviviridae family support the hypothesis that inactivation of the homologous serine protease activity in these viruses leads to the production of non-infectious viral particles (Chambers et al., 1990; De Francesco and Steinkuhler., 1999; Major et al., 1999). Thus, the HCV NS3 protease is currently one of the major targets pursued for the discovery of novel anti-HCV drugs.

Currently, there are many peptide-based inhibitors directed against the NS3 protease in order to help elucidate the mechanism of substrate recognition by the NS3 and to act as the lead compound for the development of small molecule HCV NS3 protease inhibitors. Such peptides or derivatives thereof may be effective drugs as they may satisfy the criteria necessary for effective small molecule drugs such as; stability, high resistance to proteolytic degradation, good membrane permeability and bioavailability. These properties usually require the compounds to have a low molecular weight (< 1000 Da) (Leung et al., 2000).

1.3.2 Peptidyl Inhibitors of NS3 Protease and Their Limitations

Various peptidyl inhibitors including both product inhibitors and substrate inhibitors such as the α-ketoacid and non-cleavable decapeptides spanning P6-P’4 have been reported to target HCV NS3 protease (Ede et al., 2000; Narjes et al., 2002; Lahm et al., 2002; Llinas-Brunet et al.,
1998; Landro et al., 1997; Steinkuhler et al., 2001). The FLNS3 protease cannot efficiently bind small peptide inhibitors, i.e. P6-P1 residues, requiring at least a decapeptide spanning P6-P’4 residues distal to the scissile bond that contribute significantly to binding through hydrophobic and electrostatic interactions (Po A et al., 2001; Steinkuhler et al., 1996). Despite some success with peptidyl inhibitors, significant progress in using peptidyl inhibitors against NS3 protease and FLNS3 protease has been hampered by the very nature of the target: a protein whose binding site is highly charged, solvent-exposed and featureless, and whose minimal decapetidic substrate relies heavily on charge-charge and hydrogen bonding contacts along an unusually extended surface (Steinkuhler et al., 2001; Bisceglie et al., 1995; Komada et al., 1994; Kolykhalov et al., 1996).

It is becoming more evident how the viral polyprotein substrate can compensate for this lack of binding pockets and provide the interactions required for specificity of cleavage during polyprotein processing. Therefore, an alternative strategy to inhibit NS3 protease and FLNS3 is to use a polyprotein substrate that can be recognized by the NS3 protease, such as the use existing serine protease inhibitors (serpins) as an initial scaffold to inhibit the HCV NS3 protease (Carrell., 1986). With the recent finding that the HCV NS3 serine protease can interact with and form an SDS-stable complex with naturally occurring serpins (Drouet et al., 1999), modulation of NS3 enzymatic functions by bio-engineered serpin directed at the HCV NS3 could represent a new therapeutic approach to treating HCV infections.

1.3.3 Protein-Based Inhibitor: The Case of Serine Protease Inhibitors (Serpins)

Few protein-based inhibitors have been tested against the HCV NS3 protease so far. There is some evidence that inhibition of pathogen protease by host cell serpins in the virus-host cell relationship can occur in vitro i.e., C1 inhibitor (Ci Inh) and α-2 antiplasmin
(α2AP). Both of these belong to the serpin family and can interact with the HCV NS3 protease with proteolysis of the serpins and production of a higher molecular weight SDS-stable complex (Drouet et al., 1999). In addition, Eglin c, a 70 amino acid reversible potent inhibitor of several serine proteases has been mutated to specifically target the HCV NS3 protease and it has shown some success in inhibiting the NS3 protease (Martin et al., 1998). These observations open the way to new approaches to inhibit HCV replication.

Serine Protease Inhibitors (serpins) are a large family of proteins which have been identified from various sources such as viruses and mammals. The primary function of most members of the serpin family is to neutralize over-expressed serine protease activity (Travis et al., 1983). The irreversibility of protease inhibition achieved by the serpins has made them the principal inhibitors controlling both intra and extra-cellular proteolytic pathways. Serpins regulate such diverse physiological processes such as: blood coagulation, fibrinolysis, compartment activation, fetal development and inflammation (Silverman et al., 2001; Gettins et al., 1996; Stein et al., 1995).

The structure and mechanistic features of the serpins are typified by α1-antitrypsin (α1-AT) (Carrell et al., 1982). Serpins share a highly ordered structural architecture consisting of three β-sheets, nine α-helices and a reactive site loop (RSL) of ~20 amino acids which is solvent-exposed and mimics target protease recognition sites (Fig. 1.4) (Huber et al., 1989; Whisstock et al., 1998). Serpins act as pseudo-substrate and suicidal proteins by forming 1:1 complexes with their target proteases. The resulting irreversible SDS-heat stable complex of enzyme and inhibitor are then removed from circulation for subsequent breakdown (Wright et al., 1995). The mechanism of irreversible inhibition by serpin has three stages: 1) recognition of the serpin and target protease. 2) cleavage at the P1-P'1 bond
in the reactive site loop of the serpin that triggers the insertion of the cleaved reactive centre loop through the β-sheet of the serpin molecule, which accompanies translocation of the protease to the opposite pole. 3) formation of SDS-stable complex between the serpin and the protease, which also results in the disruption of the protease active site (Fig. 1.4) (Huntington et al., 2000). The molecular structure and physical properties of serpins also permit these proteins to adopt a number of variant conformations under various physiological conditions, including the native inhibitory forms and several inactive non-inhibitory forms. These include serpins that have already formed irreversible complexes with proteases, ligands or peptides, serpins that were recognized as substrates and cleaved but did not form complexes, or degraded serpins, oxidized serpins and polymerised serpins (Fig 1.5).

1.3.4 α1-antitrypsin (α1-AT) and Bio-engineering of the Serpin Reactive Site Loop (RSL)

α1-AT, which is an important component of the serine protease inhibitor (serpin) system in humans, has been studied since 1955. Although named α1-AT, its target protease is actually elastase (Carrell., 1986). α1-AT is synthesized primarily in the liver and is about ~426 amino acids in length (Lobermann., 1984). The reactive site loop (RSL) of the serpin is believed to resemble an ideal substrate for the target enzyme, since the substrate binds but is not hydrolysed and so does not leave the active site of the enzyme. Complex formation involves more than just interaction between the active site of the protease and the RSL of the serpin. However, this interaction is the basis for determining specificity and is therefore the basis of the direct potential for simple engineered changes in function (Travis and Salvesen., 1983). α1-AT is the first protein to have been modified by genetic engineering to display a
Figure 1.4. Serine protease inhibitor (Serpin). (A) When a protease reacts with an inhibitor, the inhibitor can either be cleaved while releasing the active protease, or the inhibitor can form a covalent SDS stable complex with the protease. When complex formation occurs, the protease loses its catalytic properties. (B) Ribbon representation of the elastase (E), \( \alpha_1 \)-AT (I), and \( \alpha_1 \)-AT /elastase complex (E/I*). RSL (yellow). (Huntington et al., 2000).
Figure 1.5. Conformational polymorphism of inhibitory serpins. (Janciauskiene, 2001).
novel inhibitory activity. α₁-AT P1 position, methionine 358, was changed to valine to make an α₁-AT variant that was resistant to oxidation (Rosenberg et al., 1984; Courtney et al., 1985). By making changes to the amino acid sequence in the RSL of the α₁-AT, it has been possible to target specific proteases thus resulting in an innovation of potential therapeutic value.

After the first oxidation-resistant bio-engineered serpin was produced, many more followed, including α₁AT-Portland (α₁-PDX), a bio-engineered serpin variant of α₁-AT highly selective for furin (Tsuji et al., 2002; Jean et al., 1998). α₁-PDX was bio-engineered specifically to target furin, an endogenous human serine protease that plays an important role in the proteolytic activation of proteins encoded by many pathogenic agents such as Pseudomonas exotoxin, HIV glycoproteins, Human cytomegalovirus (HCMV) glycoproteins and Influenza A viral hemaglutinin. The inactivation of furin suggests the broad applicability and potential of α₁-PDX as a therapeutic agent (Jean et al., 1998). The most recent example of a protein therapeutic for HCMV infections was demonstrated with α₁-PDX. The production of infectious HCMV is dramatically reduced by the exogenous addition of α₁-PDX (Jean et al., 2000).

A similar strategy could be applied to developing inhibitors of the HCV NS3 protease. α₁-AT has a methionine at the P1 position of its RSL and inhibits proteases with chymotrypsin-like, trypsin-like and elastase-like specificities (Whisstock et al., 2000). Since the HCV NS3 serine protease domain adopts a chymotrypsin-like fold (Bartenschlager et al., 1993), α₁-AT represents a good macromolecular protease inhibitor scaffold to initiate protein engineering studies.
1.4 Thesis objectives

The main objective of this research was to develop and evaluate bio-engineered serine protease inhibitors (serpins) directed at the HCV NS3 protease based on the results of substrate specificity studies. To achieve the main objective, several secondary specific aims had to be completed:

i) Purification of the active recombinant histidine-tagged HCV NS3 protease.

ii) Development and optimization of an enzymatic assay to monitor HCV NS3 proteolytic activity.

iii) Determination of the substrate specificity of the HCV NS3/Pep4A protease complex using IQFS ($K_m$, $V_{max}$, $V_{max}/K_m$).

iv) Bio-engineering of recombinant $\alpha_1$-AT variants to target the HCV NS3 protease.

v) Purification of the active recombinant histidine-tagged $\alpha_1$-AT and its variants.

vi) Assessment of the inhibitory activity of bio-engineered $\alpha_1$-AT variants against elastase.

vii) Assessment of the inhibitory activity of bio-engineered $\alpha_1$-AT variants against the HCV NS3/Pep4A protease complex.
Chapter 2: Materials and methods

2.1 Preparation of Competent BL21 & DH5α *E.coli* Cells

A single colony of DH5α or BL21 *E.coli* cells was used to inoculate 5ml of Luria Bertani (LB) medium and cells were grown overnight at 37°C in an orbital shaker. The next day, 1 ml of the overnight culture was diluted in 50ml LB and allowed to grow for 2 hours at 37°C until the optical density (A_{600}) of the culture reached 0.4, with shaking. Cells were chilled on ice for 20 minutes and then pelleted at 3000 rpm by centrifugation for 5 minutes at 4°C. The cells were resuspended in 25ml 0.1M CaCl\(_2\) and incubated on ice for 30 minutes. The cells were pelleted again at 3000 rpm by centrifugation for 5 minutes at 4°C and finally resuspended in 3.65 ml CaCl\(_2\) and 0.35 ml of glycerol. Cells were stored at 4°C for at least one hour before use in transformation reactions. The remaining competent cells were snap-frozen in -86°C freezer.

2.2 Transformation of *E. coli* Strain BL21 with the NS3 Recombinant Plasmid Construct.

The NS3 pMANS34NSH plasmid construct with a 6X-histidine tag attached to the minimal protease domain (1027-1218, NS3 protease) was kindly provided by our collaborator (Nobuko Kakiuchi) from Japan. Approximately 10ng of NS3 protease plasmid DNA was mixed with 200μl of competent *E.coli* BL21 and incubated on ice for 30 minutes. Competent cell/DNA mixtures were then subjected to “heat-shock” by incubating at 42°C for 45 seconds. 800 μl of LB was added immediately and samples were incubated at 37°C for 1 hour, with shaking. Cells were then pelleted at 13000 rpm in a microfuge for 30 seconds and
800μl of the supernatant discarded. The cell pellet was resuspended in the remaining 200μl of media and then spread onto LB agar plates containing with 100 μg/ml of ampicillin and 20 μg/ml Chloramphenicol.

2.3 Expression of Recombinant NS3 Proteins.

_E.coli_ strain BL21 bearing the NS3 protease expression plasmid was grown in 5 ml of (LB) containing 100 μg/ml of ampicillin and 20 μg/ml chloramphenicol overnight at 31°C. The overnight 5 ml culture was then used to inoculate 500 ml of LB (100 μg/ml of ampicillin) at 37°C for 4 hours until the optical density \((A_{600})\) of the culture reached 0.7. The culture was then supplemented with 700 μM IPTG (Calbiochem), and transferred to a 31°C incubator to grow for another 18 hours to induce expression of the NS3 protease. The induction was performed at 31°C to prevent the formation of inclusion bodies and to produce more soluble protein. Induced bacterial cells were then harvested by centrifugation in 500ml bottles in a GS3 rotor at 4000 rpm at 4°C for 10 minutes. The cell pellets were then frozen in a -86°C freezer and thawed on ice, and resuspended in 20 ml of 1X FPLC binding buffer (0.5M NaCl, 20mM sodium phosphate, pH 7.3, 20% glycerol, 0.1% triton-X and 1 capsule of EDTA free protease inhibitor (Roche)). Resuspended/thawed _E.coli_ BL21 cells were then lysed using a sonicator (Misonix) at output level 3 for 3 X 45 seconds intervals. The lysed _E.coli_ cells were then centrifuged at 12000 rpm in SS34 rotor at 4°C for 45 minutes to separate the soluble proteins from the insoluble proteins and cell membranes. The soluble proteins in the supernatant were then filtered using 0.45 μM filter (Millipore) alliquoted and stored at -86°C prior to FPLC purification.
2.4 Purification of NS3 Protease Using Fast Protein Liquid Chromatography (FPLC).

Protein purification of NS3 protease was done with an automated AKTA purifier FPLC system (Pharmacia). Frozen *E.coli* extract supernatant containing histidine tagged NS3 protease was thawed on ice and loaded into the FPLC superloop. It was then allowed to pass through a pre-packed nickel column (Pharmacia) at 1ml/minute (nickel column was pre-equilibrated with 15 ml of 1X binding buffer—0.5M NaCl, 50 mM sodium phosphate, pH 7.3). After the supernatant had been loaded into the column, proteins with weak interaction to the nickel were removed with 3 X 60 ml sequential washes using 1X washing buffer at pH 7.3, 6.0, 4.8 (0.5M NaCl, 50 mM sodium phosphate). NS3 protease was eluted with 20 ml of 1X elution buffer (0.5M NaCl, 50 mM sodium phosphate, pH 4.0). Eluted material was collected in 1 ml fractions and stored in aliquots at -86°C. Eluted fractions were then thawed, pooled and dialyzed against neutral pH buffer (50 mM Hepes, 150 mM NaCl, 10mM DTT, pH 7.5) for 3 X 1 hours using the slide-a-lyser dialysis cassette with a 3.5 kD size cut-off membrane (Millipore) to raise the pH of the purified protein solution before analyzing the purity using Sodium Dodecyl Sulfate Poly-Acrylamide Gel Electrophoresis (SDS-PAGE) analysis and Western blots.

2.5 Protein Analysis with SDS PAGE gels and Western Blot.

Protein induction levels were determined by comparison of pre-and post IPTG induced bacterial cell samples run on SDS-PAGE (12% acrylamide) and followed by analysis using anti-histidine specific antibody for a Western blot. The Bio-Rad Mini Protean II gel apparatus was used to prepare and run mini gels. A 6% stacking gel was prepared and
prior to loading, samples were suspended in reducing sample buffer and were heated to 100°C for 10 minutes. 10uL of protein sample was loaded to each well, and the gel was subjected to electrophoresis at 125V for 75 minutes. For SDS-PAGE analysis, protein were detected with Coomassie blue stain (Biorad) for 1 hour followed by de-staining for 1 hour. For Western blot analysis, proteins form the gels were transferred to 0.45 μM nitrocellulose membrane (Biorad). Protein transfer was done using the semi-dry transfer Trans-blot SD system (Biorad) at 10V for 15 minutes with transfer buffer (see buffer list section 2.16). Unoccupied binding sites within the dried nitrocellulose membrane were blocked with blocking buffer supplied by Qiagen for 18 hours. The membrane was then washed 3 X 10 minutes with TBS-T (TBS, 0.05% Tween), followed by 3 X 10 minutes with TBS before probing with (1 step) HRP-conjugated anti-histidine RGS primary antibody (Qiagen) (anti histidine mouse IgG antibody-1/2000 dilution in blocking buffer) for 1 hour on a shaker at room temperature. For two–step Western blot analysis requiring secondary antibody incubation e.g. anti FLNS3 monoclonal antibody (Novocastra Laboratories Inc), an additional 6 X 10 minutes washes with TBS-T were required between the primary antibody incubation and the secondary antibody incubation (1/10000 secondary antibody dilution in TBS-T) was carried out for 1 hour at room temperature using HRP-conjugated anti-mouse anti-body (Pharmacia). Finally, 8 washes with TBS-T for 10 minutes each were carried out before proceeding with Chemiluminescence visualization using ECL-HRP substrate (Pharmacia) followed by exposure to Hyperfilm (Pharmacia). Hyperfilm was developed using a (Kodak X-omat) x-ray processing machine.
2.6 Development of HCV NS3 Protease Assays.

To characterize the enzymatic activity of the purified NS3 protease, the ability to cleave synthetic peptides either with or without the NS4A cofactor was investigated. Commercially available MCA substrate (Ac-E-E-V-V-A-C-AMC) (Calbiochem) was used initially to screen for activity in the FPLC purified NS3 protease. The substrate specificity of NS3 protease was then investigated using synthetic IQFS and peptide NS4A cofactor (Ac-KKKGSVVIVGRIILSGR) designed in our laboratory but synthesized by our collaborator (Luiz Juliano) from Brazil. The structural representation of MCA, IQFS and the different HCV IQFS used are shown in Fig. 3.4. The cleavage of the IQFS occurs after the cysteine in P1 position. P1, P2, etc refers to the amino acids of the substrate prior to the cleavage site. P'1, P'2, etc refers to the amino acids of the substrate after the cleavage site as described in section 1.2.3. The continuous assays of protease activity were performed using a Gemini Spectromax XS 96 well plate spectrofluorometer. Enzymatic reactions were performed at 31°C. The excitation and emission spectra for MCA and (HCV or GBV) IQFS substrates were set at 370/460 nM and 320/420 nM, respectively. Each well contained 15μM NS4A cofactor, varying concentrations of (HCV or GBV) IQFS or MCA substrate, 20 μL of FPLC purified NS3 protease and the final volume in each reaction was adjusted to 100μl using the reaction buffer (50 mM Hapes, 150 mM NaCl, 0.1% triton-X, 10mM DTT, pH 7.5). (HCV or GBV) NS4A cofactor had to be pre-incubated with NS3 protease for 15 minutes at 31°C prior to any reaction involving NS3 protease to allow proper binding between NS3 protease and NS4A cofactor.
2.7 Michaelis-Menten Steady-State Analysis of HCV Synthetic Peptide Substrate.

When the purpose of the experiment is to determine the substrate specificity of an enzyme under specific conditions, $V_{\text{max}}$ and $K_{m}$, assays must be performed at various concentration of substrate $[S]$, including low concentrations. Otherwise the $K_{m}$ cannot be determined. At low $[S]$, changes in substrate concentration are linearly reflected in the initial rate of reaction, and the slope of this linear part of the reaction allow the sensitive determinations of initial velocity ($V_0$). The ($V_0$) of six different substrate concentrations (5, 10, 20, 50, 100, and 200 μM of each different IQFS) were determined and the kinetic measurements of NS3 protease cleavage efficiency to the different IQFS were calculated from the least-square fit of initial rates ($V_0$) as a function of the six different substrate concentrations with the help of Sigma plot 2002 software, assuming Michaelis-Menten kinetics. $V_{\text{max}}$ and $K_{m}$ value can be extrapolated by Sigma plot software from the Michaelis-Menten curve and the $V_{\text{max}}/K_{m}$ values were calculated for comparison of the relative cleavage efficiency. Kinetic reactions were done in triplicate for each of the 6 different substrate concentrations. Control with no substrate or NS3 protease alone were also done in triplicate.

\[
\text{Enzyme (E) + Substrate (S)} \xrightleftharpoons[k_2]{k_1} \text{Enzyme/Substrate (E/S)} \xrightarrow{k_3} \text{Enzyme (E) + Product (P)}
\]

Catalytic constant ($k_{\text{cat}}$) = $k_3$, is a catalytic rate constant that measures how much of the enzyme/substrate (E/S) complex, once formed, is converted to products.

Maximum velocity of enzyme ($V_{\text{max}}$) = $k_{\text{cat}}(E)$, is a rate of reaction that measures how fast a given amount of the enzyme can catalyze product formation at maximum speed.
Affinity constant \((K_m) = (k_2+k_3)/k_1\), is a concentration that measures the intrinsic property of an enzyme related to the binding constant for forming \((E/S)\) complex, which also corresponds to the concentration of substrate at which the rate of the reaction is half the maximum velocity \((V_{\text{max}})\).

Specificity constant \((k_{\text{cat}}/K_m)\) = defines the rate of the overall enzymatic reaction at which the substrate concentration is much lower than \(K_m\).

The ratio of \(V_{\text{max}}/K_m\) can be used to compare enzymes and measures the efficiency of the enzyme. Relative substrate specificity of NS3 protease towards the different IQFS junctions can be calculated by comparing \(V_{\text{max}}/K_m\) of NS3 protease against each of the different IQFS junctions because if the concentration of the enzyme [NS3 protease] is constant throughout the experiment, the relative substrate specificity of the NS3 protease to the different IQFS will still be comparable. The higher the \(V_{\text{max}}/K_m\) value, the higher the specificity of NS3 protease towards the substrate, which means faster turnover and more product formation.

2.8 \(\alpha_{1}\)-AT DNA Isolation and Sub-cloning.

\(\alpha_{1}\)-AT DNA was obtained from Dr. J.L.Christian (Portland, Oregan). \(\alpha_{1}\)-AT-KDEL was obtained from Dr. C. Hashimoto (New haven, Connecticut). \(\alpha_{1}\)-AT DNA was transformed into competent DH5-\(\alpha\) \(E.coli\) (Novagen) by the standard heat shock (42°C) transformation procedure. \(\alpha_{1}\)-AT plasmid DNA propagated in \(E.coli\) was pelleted by centrifugation at 13000 rpm (microfuged) for 30 seconds and the supernatant was discarded. Plasmid DNA was then isolated using the DNA mini-prep kit (Qiagen) according to manufacturer’s instruction. Plasmid DNA was eluted in a 30\(\mu\)l fraction and stored at -20°C.
The α₁-AT gene was removed from vector pDS by restriction digest using *Xho I* and *Bgl II*. In a typical restriction digestion, 1.0 μg of pDS-α₁-AT plasmid DNA and pET-21 vector DNA were incubated with 10 units of restriction enzymes *Xho I* and *Bgl II* (NEB) for 3 hours at 37°C. The resulting DNA fragments were resolved by electrophoresis using 0.8% agarose gels (70 volts for 90 minutes). pET 21 Vector (Novagen) was also prepared and purified in the same way as α₁-AT. Desired fragments were recovered from agarose gels using the QIA quick Gel Extraction Kit (Qiagen) according to manufacturer’s instruction.

Ligation of the α₁-AT gene into the pET-21 vector (Appendix II) was performed with 50-100 ng (0.015-0.03 pmol) of pET 21 vector and 0.2 pmol α₁-AT insert (50 ng of a 500 bp fragment) in a volume of 30 μl (insert to vector ligation ratios of approximately 10:1). 1μl of T4 DNA ligase (NEB) was added last and gently mixed by stirring with a pipet tip. Incubation generally lasted 16 hours at 4°C. All ligated products were transformed into competent DH5-α *E.coli* by the standard heat shock (42°C) transformation procedure and plated on LB containing 100 μg/ml of ampicillin (37°C O/N). Grown *E.coli* colonies were then randomly picked, grown overnight at 37°C, and the plasmid DNA was purified and analyzed for successful sub-cloning by restriction digest analysis. When successful ligation was confirmed, the α₁-AT/pET 21 DNA plasmid was transformed to competent BL21 *E.coli* for expression and protein purification.

2.9 Site-Directed Mutagenesis of α₁-AT Using Polymerase Chain Reactions (PCR).

All the α₁-AT variants including the α₁-AT-KDEL variants in pCMV shuttle vector (Appendix III) were mutated by PCR/site-directed mutagenesis using the QuickChange site-
directed mutagenesis kit protocol (Stratagene). Oligonucleotide primers were designed using the Vector NTI and Web Cutter 2.0 program according to the QuickChange manual. All oligonucleotides primers contained the desired mutation and annealled to the same sequence on opposite strands of the plasmid. They are between 25-45 bases in length, have at least 40% GC content and have a melting temperature of $78^\circ\text{C}$ or higher. All oligonucleotide primers were synthesized and purified by Alpha-DNA (Montreal).

Each PCR mixture (50μl total volume) contained 5μl of 10X reaction buffer (Stratagene), 1μl of (5-50 ng/μl) ds DNA template, 1.25μl (125 ng) of oligonucleotide primer #1, 1.25μl (125 ng) of oligonucleotide primer #2, 5μl (2.5 mM) of dNTP mix (Pharmacia), 35.5μl H$_2$O, and 1μl of PfuTurbo DNA polymerase (2.5 U/μl) (Stratagene).

**The Cycling parameters for the QuickChange Site-Directed Mutagenesis were as follows:**

<table>
<thead>
<tr>
<th>Segment</th>
<th>Cycles</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>$95^\circ\text{C}$</td>
<td>30 seconds</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>$95^\circ\text{C}$</td>
<td>30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$55^\circ\text{C}$</td>
<td>1 minutes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$68^\circ\text{C}$</td>
<td>2 minutes/ kB (template)</td>
</tr>
</tbody>
</table>

- The extension time for making all the $\alpha_1$-AT variants (~4.8 kB) was 10 minutes.
- The extension time for making all the KDEL $\alpha_1$-AT constructs (~6.8 kB) was 14 minutes.
- All reactions were allowed to cool down for at least 2 minutes at $4^\circ\text{C}$ before further processing.
List of Primers Used for the Site-Directed Mutagenesis of pET-21 α₁-AT Plasmid

Mut A-(use α₁-AT plasmid DNA as template) p1 (Cys)


Mut B-(use mut A plasmid DNA as template) p1 (Cys), p4 (Ser) & p6 (Asp)


Mut C-(Use mut A plasmid DNA as template) p1 (Cys) & p6 (Asp)


Mut D-(Use mut A plasmid DNA as template) p1 (Cys) & p4 (Ser)


Mut E-(Use mut B plasmid DNA as template) (ap 21,22)- p1 (Cys), p4 (Ser), p6 (Asp) & p’1 (Pro).


PCR reactions were then subjected to *Dpn I* digestion for 4 hours at 37°C to remove the original plasmid (methylated) DNA. Following the digestion, the DNA contained within each 50μl reaction was precipitated by adding 3M sodium acetate (5μl): 95% ethanol (100μl) and incubation at -86°C for 30 minutes. Reaction were then centrifuged at 13000 rpm for 15 mins at room temperature (Eppendorf) to pellet DNA. Excess ethanol was removed and precipitated mutant DNA was resuspended with 10μl TE buffer. 3μl of the DNA in TE buffer were then used to transform the mutant DNA into competent DH5-α *E.coli* using the standard (42°C) heat-shock method. Transformed DH5α cells were grown in overnight cultures and the plasmid DNA was isolated and sent for sequencing. All sequencing reactions were performed by the NAPS unit (University of British Columbia). After the mutations for all variants were confirmed, all mutants DNA were transformed to *E.coli* BL21 for expression.

2.10 Expression and Purification of α₁-AT and Its Variants.

α₁-AT and its variants were expressed and purified using the protocols described for NS3 protease (section 2.3-2.4) with the following exception: The unbound proteins were removed with 3 sequential washes of 60 ml with buffer of decreasing pH value (0.5M NaCl, 50 mM sodium phosphate, pH 7.3, 5.8, and 5.2). α₁-AT was eluted with 20 ml of 1X elution buffer (0.5M NaCl, 50 mM sodium phosphate, pH 3.9). Eluted material was collected in different tubes and stored in aliquots at -86°C. Eluted fractions were also dialyzed against (50 mM Hepes, 150 mM NaCl, 10mM DTT, pH 7.5) for 3 X 1 hours using the slide-a-lyser dialysis cassette with a 10 kD size cut-off membrane (Pierce) to raise the pH of the purified
protein solution before analyzing the purity using SDS-PAGE analysis and Western blots (section 2.5).

2.11 Pancreatic Elastase Assays.

Purified $\alpha_1$-AT and its variants were tested for their inhibitory properties against elastase. Commercially available MCA substrate (MeoSuc-A-A-P-V-AMC) (Calbiochem) was used to detect the protease activity of elastase (Calbiochem). Continuous assays monitoring elastase activity and the inhibitory activities of $\alpha_1$-AT and variants were performed using a Gemini Spectromax XS 96 well plate spectrofluorometer. Enzymatic reactions were carried out at 31°C. The excitation and emission spectra for MCA were set at 370/460 nM. Each optimized elastase/inhibitor reaction contained the Hepes buffer (50 mM Hepes, 150 mM NaCl, 0.1% triton-X, 10mM DTT, pH 7.5), 9 nM of elastase, 50µM MCA substrate and 1.8 µM of FPLC purified $\alpha_1$-AT or its variants to make a total volume of 100µL in every reaction well. The concentration of purified $\alpha_1$-AT and its variants were determined using standard Bradford assays (Biorad).

2.12 Quantification of Proteins.

Protein concentrations were determined using the Bio-Rad protein assay kit (micro-assay), which is based on the Bradford dye-binding protein assay, according to the manufacturer’s instructions. A standard curve of protein concentration at absorbance 595 nm was produced, using BSA as the standard protein at the following concentrations (µg/ml); 50, 100, 200, 400, 800, 1000. The absorbance of the sample proteins was measured and protein concentrations were determined from the standard curve.
2.13 Complex Formation Studies with \( \alpha_1\)-AT and Its Mutant Variants against Elastase, NS3 Protease and FLNS3.

Elastase complex formation was performed with 1.8 \( \mu \text{M} \) of \( \alpha_1\)-AT and its variants using the elastase/inhibitor assay buffer (section 2.11) without the addition of MCA substrate. Some reactions were performed with varying concentrations of elastase or different complex incubation times (variations are detailed in the appropriate figure legends in the Result section). NS3 protease complex formation with \( \alpha_1\)-AT and its variants was performed using the NS3 protease enzymatic assay buffer, NS4A (section 2.6) with the addition of 1.3 \( \mu \text{M} \) of serpins. Some reactions were carried out with varying concentration of NS3 protease or different complex incubation times (variations are detailed in the figure legends of the Result section). FLNS3 complex formation with \( \alpha_1\)-AT and its variants was performed using the NS3 protease enzymatic assay buffer, NS4A (section 2.6) with the addition of 20\( \mu \text{l} \) of FPLC purified FLNS3 and 1.3 \( \mu \text{M} \) of serpins. The complex formation reaction was allowed to proceed for 4 hours. All complex formation reactions were analyzed using SDS-PAGE and Western blot (section 2.5). Anti histidine antibody was used to detect NS3 protease or elastase and \( \alpha_1\)-AT or all its variants complex formation. Anti FLNS3 antibody was used to detect complex formation between FLNS3 and \( \alpha_1\)-AT or all its variants.

2.14 Inhibition Assay of Elastase and NS3 Protease by \( \alpha_1\)-AT and All Its Variants.

Inhibition of elastase proteolytic activity by \( \alpha_1\)-AT and its variants was performed using elastase/inhibitor assay buffer (section 2.11). Titration assays of elastase activity were performed against \( \alpha_1\)-AT, Mut A and Mut D with the same buffer condition (section 2.11)
except with varying amount of serpins included (variations are detailed in the figure legends of the Result section). In titration assay, the serpins were incubated with elastase for 45 minutes prior to adding the MCA substrate and performing continuous kinetic assay for 3 hours. Progress curve assay of elastase with α₁-AT and its variants was also performed with buffer condition (section 2.11) except the amount of elastase used was 9 nM, and the amount of serpins used were also varied (variations are explained in the figure legends of the Result section). In the progress curve assay, elastase, serpins and substrate were added at the same time prior to performing continuous kinetic assay for 3 hours. Inhibition assay of NS3 protease by α₁-AT and its variants were performed using the optimized NS3 protease enzymatic assay buffer, (section 2.6) with the addition of 1.3 μM of serpins, however new stock of peptide NS4A cofactor was used for this experiment and was acquired from (Genome BC Proteomic Centre, U.Victoria, BC). NS3 protease and NS4A cofactor were pre-incubated for 15 minutes first, followed by further 4 hours incubation with the serpins at 31°C. Abz-E-D-V-V-C-C-S-M-S-Y-Q-Y(NO₂) IQFS substrate (Multiple Peptide Sequence, California) was added immediately before taking continuous kinetic measurements for 2 hours. All readings were taken in duplicate.

**Calculation of the Specific Inhibition Constant (Kᵢ) and Stochiometry of Inhibition (Sᵢ).**

The Specific Inhibition constant (Kᵢ) is a concentration that is used to describe how potent a specific inhibitor is against its target. Titration assays with different amount of inhibitors were performed to calculate the amount of residual protease activity. The various residual protease activities were then plotted against the concentration of inhibitor added.
From the resulting graph, the value of \((K_i)\) and \((S_i)\) can be obtained (for sample calculation see Appendix I).

2.15 Model Structure of NS3 Protease/NS4A Complex, Helicase/NTPase and \(\alpha_1\)-AT/Elastase Complex Structure.

All ribbon representation of the crystal structures was generated by using the Sybyl version 6.7 software (Tripos, St-Louis).
2.16 List of Buffer Solutions

Dialysis buffer-50 mM Hepes, 150 mM NaCl, 10mM DTT, pH 7.5

DNA loading buffer (6X)-4% sucrose, 0.25% bromphenol blue, 0.25% xylene cyanol

Enzymatic assay and complex reaction buffer-50 mM Hepes, 150 mM NaCl, 0.1% triton-X, 10mM DTT, pH 7.5

Gel destain-5% (v/v) methanol, 7% (v/v) acetic acid, 88% (v/v) water

Resolving gel buffer-0.74 M Tris-HCL (pH 8.0), 1% SDS

SDS-PAGE sample buffer (6X)-0.28 M Tris-CL, 30% (v/v) glycerol, 1% (w/v) SDS, 0.5 M DTT, 0.0012% (w/v) brompheno blue

SDS-PAGE resolving buffer (5X)-0.125 M Tris base, 0.96 M glycine, 0.5% (w/v) SDS

Stacking gel buffer-0.122 M Tris-HCL (pH 6.7), 0.1% SDS

Sodium phosphate FPLC washing buffer (dibasic)-0.5M NaCl, 50 mM sodium phosphate (dibasic), pH 7.3, 6.0, 5.8, 5.2, 4.8 adjusted using sodium phosphate (monobasic) buffer

Sodium phosphate FPLC elution buffer (monobasic)-0.5M NaCl, 50 mM sodium phosphate (monobasic), pH 3.9 adjusted using phosphoric acids

TBE (1X)-89 mM Tris base, 89 mM boric acid, 2 mM EDTA

TBS-100 mM Tris-Cl, pH7.5, 0.9% (w/v) NaCl

TBS-Tween-TBS with 0.05% Tween 20

TE buffer-10 mM Tris-Cl, 1 mM EDTA, pH 8.0

Western blot transfer buffer-39 mM Glycine, 48 mM Tris, 0.037% (w/v) SDS, 20% MeOH

Western blocking buffer (qiagen)-Qiagen blocking reagent buffer diluted 1/10 in distilled water (500 ml), 5 g of Qiagen blocking reagent, 0.1% Tween 20

Western blocking buffer (milk)-5% milk in TBS-T
Chapter 3: Expressions, Purification and Characterization of NS3 Protease Using Internally Quenched Fluorogenic Substrates (IQFS)

3.1 Expression of Recombinant NS3 Protease in pLys BL21 E.coli

The serine protease domain of HCV NS3 has previously been expressed in E.coli HB101 and purified by affinity chromatography using a nickel-agarose column (Qiagen) (Vishnuvardhan et al., 1997). However, this purification method was performed manually and hence was labor-intensive and time-consuming. This chapter presents work on the development of a novel expression and automated purification method for NS3 protease using Fast Protein Liquid Chromatography (FPLC). Two different strategies have been used successfully to elute histidine-tagged NS3 protease bound to a Ni$^{2+}$ column: 1) stepwise decrease in buffer pH value, 2) increasing imidazole concentration (Sali et al., 1998; Vishnuvardhan et al., 1997). In this study, the strategy used for the purification of the NS3 protease was a stepwise decrease in buffer pH values to elute the NS3 protease.

The transformation of pLysS BL21 E.coli with the NS3 protease plasmid construct was successful. pLysS BL21 E.coli was selected as the suitable host for the expression of the NS3 protease because the pLys plasmid encodes the T7 lysozyme that stabilizes the plasmid carrying the NS3 protease gene. BL21 E.coli strain is also protease-deficient. It was chosen in order to eliminate the degradation of the NS3 protease by the bacterial proteases during expression (Grodberg and Dunn., 1988). The expression of soluble NS3 protease was found to be temperature-dependent, because induction of protein expression above 31°C led to the production of insoluble NS3 protease. SDS-PAGE analysis of (+) IPTG and (-) IPTG
bacterial cell cultures after 18 hours of induction showed that a protein band of molecular weight of approximately 23 kDa was induced, and this was consistent with the molecular mass of the NS3 protease calculated from the primary sequence (data not shown). To further confirm the identity of the 23 kDa band, RGS-histidine-tag-specific mouse monoclonal antibody was used to probe Western blots of proteins from induced cell culture. Of the three different primary antibodies tried against the histidine tag of NS3 protease, The RGS mouse anti-histidine antibody seemed to be the best choice, because it specifically bound to the NS3 protease when it was induced in BL21 E.coli cells. The Western blot result shows that there was an intense band at the 23 kDa position only in the IPTG-induced culture, which was an indication that the 23 kDa protein band was the NS3 protease (data not shown).

3.2 Purification of Recombinant NS3 Protease from E.coli Supernatant

Prior to proceeding with FPLC purification of the NS3 protease, the enzymatic activity of the NS3 protease in the crude bacterial cell lysate was tested in an assay using MCA substrate (data not shown). Having established that the expressed NS3 protease was proteolitically active, FPLC purification of the NS3 protease from the crude supernatant was attempted. A pre-packed 1ml nickel-sepharose column (Pharmacia) suitable for affinity chromatography was used to purify histidine-tagged NS3 protease. Initial attempts at FPLC purification of the NS3 protease crude supernatant using the nickel column did not yield protein of sufficient purity for our assay. Hence, a variety of protocols including changing the pH of the washing and elution buffers, and changing the number and length of washes, were tested in order to optimize the purification conditions. This resulted in the production of much higher purity of NS3 protease and a clear elution peak (Fig. 3.1). The purity of the
Figure 3.1. Purification of bacterially expressed histidine-tagged NS3 protease. Representative chromatogram of NS3 protease purification using Ni\(^{2+}\) binding interaction-chromatography. NS3 protease was eluted at pH ~4.0. **Inset.** SDS/PAGE (lanes 1 and 2) and western blot (lanes 3 and 4). This procedure yielded recombinant NS3 protease (~23 kDa) that was essentially pure and intact as determined by coomassie blue staining (lane 2) and western blot analysis (lane 4). Histidine-tagged NS3 protease was detected by western blot using anti-histidine tag antibody.
eluted NS3 protease was 96%, as determined using an alpha imager spot densitometer. Protein elution was monitored using two wavelengths: 220 nm (adsorption of peptide bonds linking amino acids) and 280 nm (adsorption by aromatic amino acids, tyrosine and tryptophan) (Fig 3.1). In addition, the chromatogram also shows the pH value in the column throughout the purification. Following the binding of the crude bacterial cell lysate onto the Ni\(^{2+}\) column, thorough washes using buffers at decreasing pH values (7.3, 6.0, and 4.8) were carried out to remove any proteins bound non-specifically to the Ni\(^{2+}\) column. Histidine-tagged NS3 protease eluted at pH 4.0 because when the pH was lowered to pH 4.0, the protonation of the 6X-histidine segment disrupted its interaction with the Ni\(^{2+}\) and caused it to be repelled from the positively charged nickel ions and thus release the histidine-tagged NS3 protease (Gagnon., 1999). The main advantage of using FPLC to purify recombinant protein is that once the protocols are determined, the purification procedure can be made fully automated to yield consistent results. Moreover, the FPLC machine can be operated at 4°C to decrease protein degradation during the purification process.

To analyse the protein content of the different washing and elution peaks, SDS PAGE was performed. As expected, both the injection peak and the washing peaks eluted during the purification process contained many different proteins, including some NS3 proteases (data not shown). However, the peak eluted at pH 4.0 (collected in 20 X 1 ml fractions) contained ~96% pure NS3 protease as shown by the SDS-PAGE analysis (Inset of fig 3.1). The purified NS3 protease migrated as a single band with a molecular mass of ~23 kDa. The identity of this band was further confirmed by Western blot analysis with the RGS antibody directed against the histidine tag (Inset of Fig 3.1).
3.3 Optimization of Assay Conditions for Detection of NS3 Protease Activity

Knowing the important role that the NS4A cofactor has on the enzymatic activity of the HCV NS3 protease, it was important to ascertain the concentration of peptide NS4A cofactor necessary to induce optimal activity of the newly purified HCV NS3 protease. This process was part of the development procedure to find out the best buffer conditions and parameters to use in an enzyme kinetics experiment. All enzymatic assays were performed using the Gemini Spectramax XS 96-well plate spectro-fluorometer. Various experiments were designed to determine suitable buffer conditions for the NS3 protease assay. The parameters which were investigated included the buffer type, pH, temperature, effects of detergents and salt concentrations. The result of these experiments indicated that the optimal buffer to use was 50 mM Hepes, 150 mM NaCl, 0.1% Triton-X, 10mM DTT, pH 7.5. It was also found that 15μM of NS4A peptide cofactor was sufficient to induce the desired activity of the HCV NS3 protease under the conditions used (Fig 3.2.A). This result agrees with other published results (Zhang et al., 1997; Steinkuhler et al., 1996) which report using the same concentrations of peptide cofactor to perform NS3 protease assays. The amount of cofactor needed to activate the proteolytic activity of the NS3 protease to half of its maximum level or the affinity constant (K_{d0.5}) of peptide NS4A cofactor to the NS3 protease was about 4.5 μM.

In order to test whether the requirement of cofactor in the activation of the NS3 protease activity was specific to HCV NS4A, cofactors from closely-related Flaviviridae viruses: GBV-A, GBV-B, GBV-C and BVD-V all of which share ~ 30% amino acid sequence homology with the HCV NS4A cofactor were tested using the MCA substrate (Butkiewicz et al., 2000). Results showed that only the HCV NS4A cofactor gave rise to NS3 protease
Figure 3.2. NS3 protease activity titration with Pep4A<sub>21-34</sub>. (A) Cofactor assay was done with the standardized enzymatic buffer show that 15 μM NS4A cofactor is necessary and sufficient to activate optimal NS3 protease activity. NS4A cofactor was pre-incubated with 20 μl of NS3 protease for 15 minutes before performing the enzymatic assay with 100 μM IQFS-1 for 1 hour. (B) 50 μM of the peptide cofactors from related Flaviviridae virus (GBV A, B, C and BVD-V) failed to activate NS3 protease activity showing that NS3 protease activity is fully dependent and specific for its own HCV-NS4A cofactor. Enzymatic assay was done for 1 hour with 15 minutes prior incubation between NS3 protease and NS4A cofactors. All experiments were done in triplicate. Buffers compositions are described in Materials and Methods section 2.6.
proteolytic activity (Fig. 3.2.B). These results agree with previous findings which showed that NS3 protease activity is highly dependent on and specific for its own cofactor, the HCV-NS4A cofactor (Butkiewicz et al., 2000).

3.4 Probing the Substrate Specificity of NS3 Protease Using HCV IQFS

Substrate recognition by all serine proteases involves binding of the substrate onto the active site surface that contains catalytic triads. In order to design a potent protease inhibitor against the HCV NS3 protease, it is important to characterize the substrate specificity of the HCV NS3 protease. In order to address this question, the ability of NS3 protease to cleave IQFS substrates corresponding to the HCV polyprotein junctions NS4A/4B, NS4B/5A and NS5A/5B was tested. IQFS substrates which contain HCV NS junctions P6-P’4 residues are suitable for use in this experiment because they can be cleaved by both the NS3 protease and the FLNS3. In contrast the MCA substrate which contains only the P-side residues can only recognized and cleaved only by NS3 protease (Fig. 3.3). Since the goal of the Dr. Jean’s laboratory is to determine the substrate specificity of the NS3 protease and the FLNS3 using the same substrate, only IQFS substrate could be used to perform this experiment. Determining the substrate specificity of the NS3 protease and the FLNS3 also helped to establish which of the NS junctions was the most efficiently cleaved. NS3/4A IQFS junction was not included in this study because it was previously reported that this junction was not cleaved, probably supporting the notion that the site is recognized in an exclusively cis-intramolecular fashion (Steinkuhler et al., 1996).

In addition, a modified form of the NS5A/5B junction IQFS which was present in high quantity was initially used to test if the NS3 protease proteolytic activity could be
Figure 3.3. The intramolecularly quenched fluorogenic substrates (IQFSs) are suitable for continuous monitoring of HCV NS3/4A protease complex activity. (A) Structural representation of peptidyl MCA-1 and IQFS-1 (modified NS5A/5B). (B) Discrepancy is observed for NS3 variants in the efficiency of cleavage of MCA-1 and IQFS-1. Only IQFS can be used to measure the activity of FLNS3.
monitored using the IQFS substrate. The modified NS5A/5B IQFS has a slightly different sequence from that of the original NS5A/5B BK strain and was cleaved with the highest efficiency by the NS3 protease. Each IQFS substrate was used at varying concentrations (5,10,20,50,100,200 µM) in assay reactions containing constant concentrations of the NS3 protease. The slope of the initial velocity of each reaction was calculated to plot the Michaelis-Menten curves of initial velocity (V₀) against the IQFS concentration. Kₘ and Vₘₐₓ values were then obtained from the Michaelis-Menten plots for each different IQFS junction (Fig. 3.4). Kₘ is defined as the Michaelis constant corresponding to the concentration of substrate at which the rate of the reaction is half of the maximum velocity, (Vₘₐₓ). Kₘ and Vₘₐₓ values for each HCV (NS) junction IQFS are in Fig. 3.5. Vₘₐₓ/Kₘ values are an indication of the relative cleavage efficiency of the NS3 protease at each different junction, since the concentration of the NS3 protease in these kinetic experiments was constant (see Materials and Methods, section 2.6). The Vₘₐₓ/Kₘ value for the modified NS5A/5B junction was the highest (0.46), followed by the BK strain NS5A/5B (0.13), NS4A/4B (0.05) and NS4B/5A (0.02). Based on the values of the Vₘₐₓ/Kₘ calculated for the different HCV (NS) junctions, a putative order of HCV (NS) polyprotein cleavage was proposed. Owing to its high cleavage efficiency, NS5A/5B junction could be cleaved first, followed by the NS4A/4B and NS4B/5A junctions. This result agrees with previous findings (Steinkuhler, 1996). Whether this putative NS cleavage sequence also happens during HCV polyprotein maturation replication in vivo is yet to be determined. Therefore, whether the putative cleavage order is significant in HCV replication in vivo is still unknown. In addition, IQFS 3 & 4 seem to exhibit internal quenching effects at high IQFS substrate concentrations [200 µM]. These phenomena could possibly be the result of product inhibition.
Figure 3.4. Substrate specificity of NS3 protease. (A) Natural cleavage sites of the HCV NS polyprotein. (B) Quenched fluorogenic substrates. (C) Representative plots of the initial rate of cleavage (rate) versus substrate concentration [5, 10, 20, 50, 100, 200 μM], for the selected substrates, IQFS-1 (modified NS5A/5B), IQFS-2, IQFS-3 and IQFS-4, obtained with 20μl of NS3 protease with 15 μM of NS4A. Enzymatic assays were done in triplicates and data analyses were carried out as described in the Materials and Methods section 2.6 and 2.7.
Figure 3.5. NS3/Pep4A protease complex activities on peptide substrates corresponding to trans-cleavage sites. NS5A/5B (modified) junction was cleaved most efficiently by NS3/Pep4A protease complex followed by natural NS5A/5B, NS4A/4B and NS4B/5A junctions.
by the N-terminal cleavage products of substrate peptides corresponding to the NS4A/4B, NS4B/5A and NS5A/5B cleavage sites, as described in previous work (Steinkuhler et al., 1998). The NS3 protease also displays a high affinity for N-terminal cleavage products of the original substrate, especially with the N-terminal sequence of the original NS5A/5B junction. The modified NS5A/5B substrate did not show any sign of product inhibition, probably due to the nature of its N-terminal sequence. This phenomenon might partly explain the self-regulation and slow propagation of the HCV in infected individuals, as product inhibition by the N-terminal of NS5A/5B junction will inhibit the activity of NS3 protease and stops the processing of other non-structural junctions.

3.5 Probing the Substrate Specificity of NS3 Protease Using GBV IQFS

The substrate specificity of NS3 protease against GBV-A & GBV-B virus polyprotein IQFS substrate was also investigated. GB viruses A & B are closely related to HCV and cause acute hepatitis in tamarins (Sanguinus species), making them attractive surrogate viruses for in vivo testing of anti-HCV inhibitors in a small monkey model. It has been reported that the NS3 protease of GBV-B shares similar substrate specificities with its counterpart in HCV (Scarselli et al., 1997). However, results from our laboratory indicate this is not the case. Figure 3.6. (A) shows the amino acid sequences corresponding to the (NS) cleavage junctions of GBV-A & B, with the putative P1 and P6 residues highlighted in red. Of all the GBV IQFS substrates tested using the NS3 protease and the peptide NS4A cofactor, only the NS4A/4B junction of GBV-B was cleaved efficiently under the same assay conditions used to test the NS3 protease against the HCV IQFS substrates. The cleavage
**A  Internally Quenched Fluorogenic Substrate (IQFS): GBV-A and GBV-B**

**GBV-A**
- NS3-NS4A: `Abz-S-L-V-V-V-T-S-W-V-V-Q-EDDnp` not-cleaved
- NS5A-NS5B: `Abz-E-E-E-T-P-T-S-Y-S-Y-Q-EDDnp` not-cleaved

**GBV-B**
- NS4B-NS5A: `Abz-T-P-T-E-D-D-C-G-L-I-Q-EDDnp` not-cleaved
- NS5A-NS5B: `Abz-K-S-E-F-S-C-S-M-S-Y-Q-EDDnp` not-cleaved

**B**

Figure 3.6. NS3/Pep4A protease complex specificities towards GBV-A and GBV-B IQFS. (A) Lists of GBV-A & B IQFS from GBV-A & B non-structural pro-protein junctions that are tested against HCV NS3/Pep4A protease complex. (B) Efficiency of cleavage of internally quenched fluorogenic GBV-B (NS4A/4B) non-structural pro-protein related substrate by 20 μl of NS3 protease. Concentration of GBV-B IQFS used were: 5, 10, 20, 50, 100, 150 (μM). All enzyme kinetics assays were done for 1 hour in duplicates. Buffers compositions are described in Materials and Methods section 2.6.
Figure 3.7. Ribbon representation of the catalytic pocket of NS3/Pep4A protease complex. (A) Ribbon representation of the NS3/Pep4A protease. (B) Close up of ribbon representation of the binding pocket of the NS3/Pep4A protease catalytic site bound to P6-P’4 decapeptide inhibitors. Color code for NS3/4A: Yellow, hydrophobic, red, negative charge density, blue, positive charge density, magenta, NS4A residue; Color code for the peptide inhibitor: green, carbon, red, oxygen, blue, nitrogen, yellow, sulfur; Hydrogens omitted for clarity (Ingallinella., 2000).
efficiency of this junction ($V_{\text{max}}/K_m = 0.33$) (Fig. 3.6 B) was comparable to the cleavage efficiency of the HCV (NS) junctions.

The reason why only the NS4A/4B junction of the GBV-B was cleaved is probably due to the presence of cysteine residue in the P1 position, glutamic acid in the p6 position and alanine in the P'1 position. Having these three amino acids requirements in their respective positions was essential for the efficient cleavage of substrate by NS3 protease, as previously described in the Introduction. None of the other GBV substrates share these sequence requirements for the efficient cleavage by the NS3 protease. Although the NS5A/5B junction of the GBV-B has P1-cysteine and a P'1-serine, it was not cleaved efficiently. It has been suggested that although an acidic residue at P6 is not essential for cleavage by the NS3 protease, the overall negative charge in the P-region of the substrate might be important for improving binding and cleavage efficiency (De Francesco & Steinkuhler., 1999). Figure 3.7 (A&B) highlights the fact that the active site of the HCV NS3 protease is very hydrophobic, with some positive charge around it. Therefore, the presence of a positively charged lysine in P6 of the GBV NS5A/5B junction may have a big impact on the overall charge to the P-region resulting in the disruption of the enzyme/protease interaction. The P-region could be de-stabilized due to the repulsion caused by the introduction of a positively charged area around the S6 position of the NS3 protease induced by the arginine residue. (Fig 3.7 B) (Ingalinella et al., 2000).

3.6 Summary

This chapter describes a successful novel expression and purification protocol for recombinant histidine tagged HCV NS3 protease using a fully automated FPLC system. In
addition, enzymatic assay parameters and buffer requirements for characterizing the NS3 protease were determined and used for all subsequent enzymatic experiments involving the NS3 protease. A comparison of the results from the substrate specificity studies using the HCV NS3 protease and the HCV or GBV-A & B IQFS substrates indicates that the presence of cysteine in P1, aspartic/glutamic acid in P6 and serine/alanine in P'1 are required for the efficient cleavage of the substrates by the HCV NS3 protease under the condition used. The $V_{\text{max}}/K_{\text{m}}$ values for modified NS5A/5B junction were the highest (0.46), followed by BK strain NS5A/5B (0.13), NS4A/4B (0.05), and NS4B/5A (0.02). These results will be useful for the design of effective inhibitors against the HCV NS3 protease.
Chapter 4: Bio-engineering, Expression & Purification of $\alpha_1$-AT Variants (mutants A-E)

4.1 Sub-cloning of $\alpha_1$-AT from pDS-56 Vector to pET-21 Vector

$\alpha_1$-AT was chosen as the first serpin scaffold to initiate protein engineering studies with the aim of specifically targeting the HCV NS3 protease. It was selected due to its ability to recognize chymotrypsin-like serine proteases and because of previous success with using $\alpha_1$-AT variants as a scaffold for protein-based therapeutics (Jean et al., 1998). In addition, $\alpha_1$-AT is mainly produced in liver cells, the target cells of an HCV infection (Ray et al., 1977; Boskovic et al., 1998). Since HCV primarily infects the liver cells, it was of interest to investigate the possibility of wild-type $\alpha_1$-AT interacting with HCV NS3 protease and to find out whether the bio-engineered $\alpha_1$-AT could improve the interaction and specificity between the NS3 protease and serpins.

Prior to commencing sub-cloning of $\alpha_1$-AT, the pDS-56-$\alpha_1$-AT plasmid was sequenced to confirm that there were no mutations from the wild-type sequence. As the sequence obtained was in agreement with the Gen-Bank published sequence, the plasmid was transformed into *E.coli* BL21 for preliminary expression and purification of the protein. However, $\alpha_1$-AT was inducibly expressed at very low levels and there was not a satisfactory quantity of it produced for purification. In addition, the mutant variants of pDS-56 $\alpha_1$-AT (Mut A & B) were also expressed in such a low quantity that attempts to purify those variants yielded unacceptably low amounts of protein (data not shown).
To solve this problem, a different expression vector was used. The pET-21 vector was chosen because it is the most powerful system developed to date for the cloning and expression of recombinant proteins in *E.coli*. It utilizes the T7 RNA polymerase that is selectively induced and promotes very high protein expression levels. Almost all of the cell’s resources are converted to target gene expression; the desired product can comprise more than 50% of the total cell protein only a few hours after induction (Grodberg et al., 1988). The pET-21 vector was selected instead of other pET vectors because it contained the desired components and restriction sites (*Xho I* & *BLP I*) to allow the entire histidine tagged-α₁-AT gene to be sub-cloned from the pDS-56 vector to the new pET-21 vector (Appendix II).

The sub-cloning procedure was done according to the manufacturer’s instructions (Novagen), pDS-α₁-AT and pET-21 were propagated in DH5-α *E.coli* and purified with a plasmid mini-prep kit (Qiagen). Approximately 1 µg of the purified α₁-AT and pET-21 DNA was digested with *Xho I* & *BLP I* to drop out the α₁-AT insert and to prepare the pET-21 vector for ligation, respectively. After separation on an agarose gel, the α₁-AT insert and the pET-21 were recovered using the Qia-Quick Gel Extraction Kit (Qiagen). Ligation was then performed and ligated products were subjected to restriction digest using *Xho I* and *Blp I* to confirm successful sub-cloning. Agarose gel analysis showed that the sub-cloned pET-21 α₁-AT and the original pDS-56 α₁-AT both released α₁-AT insert (1320 bp) when digested with *Xho I* and *Blp I*. The released pDS-56 and pET-21 vectors were of correct size at 3413 bp and 5369 bp, respectively (Fig 4.1 A). These results confirmed that pDS-56 α₁-AT was successfully sub-cloned in the pET-21 vector.
Figure 4.1. Subcloning of \( \alpha_1\)-AT from pDS-56 to pET-21. (A) \( \alpha_1\)-AT was successfully sub-cloned from pDS-56 to pET-21 as indicated by the Xho I & Blp I digestion of pDS-56 \( \alpha_1\)-AT and pET-21 \( \alpha_1\)-AT analysed with DNA agarose gel. (B) SDS gel showing that pET-21 induced \( \alpha_1\)-AT is present in greater quantity than pDS-56 induced \( \alpha_1\)-AT. (C) Western blot of uninduced (\( \Phi \)) vs induced (I) pET-21 \( \alpha_1\)-AT probed with anti-histidine antibody.
Following the successful sub-cloning, pET-21 α₁-AT and pDS-56 α₁-AT expression plasmids transformed into BL21 E.coli were induced under the same conditions and culture samples from each pre-and post induction were analysed by SDS-PAGE to compare the level of expression of α₁-AT. As expected, α₁-AT (~49 KDa) was expressed at a much higher level from the pET-21 vector as compared to the pDS-56 vector (Fig 4.1 B). The identity of the expressed α₁-AT was confirmed by Western blot analysis of un-induced and induced whole cell lysate from BL21, containing pET-21-α₁-AT (Fig 4.1 C). The RGS anti-histidine antibody was specific in detection of the induced 6X-histidine tagged recombinant α₁-AT (~49 KDa). Following the successful expression of α₁-AT, site-directed mutagenesis of the α₁-AT gene to create reactive site loop sequence variants was performed to target HCV NS3 protease.

4.2 Bio-engineering of α₁-AT Variants

Using information regarding the substrate specificity of the NS3 protease from studies presented in Chapter 3 and others, it was concluded that P1, P6 & P’1 residues determined the cleavage efficiency of the NS3 protease. Previous studies also suggested that an alanine residue in P4 may decrease the cleavage efficiency of the NS3 protease (Zhang et al, 1997). Therefore, in order to better understand how these important residues within an Reactive site loop of a serpin acting as a pseudo-substrate for the NS3 protease would influence its inhibitory properties, 5 different mutants of α₁-AT were designed and constructed for testing as potential inhibitors of the HCV NS3 protease. The α₁-AT variants generated were:
1) Mutant A (Mut A) (P1, M→C)

2) Mutant B (Mut B) (P1, M→C), (P4, A→S), (P6, L→D)

3) Mutant C (Mut C) (P1, M→C), (P6, L→D)

4) Mutant D (Mut D) (P1, M→C), (P4, A→S)

5) Mutant E (Mut E) (P1, M→C), (P4, A→S), (P6, L→D), (P'1, S→P)

Mut A and B were created first, followed by Mut C, D & E. All α1-AT mutants were generated by the PCR protocols presented in Section 2.9 and sequenced to ensure that no additional mutations were incorporated into the mutants (NAPS unit, UBC).

Clustal W alignment of the nucleic acid and amino acid sequences of wild-type α1-AT and all the variants were performed to highlight the detailed changes made to each mutant (Appendix IV & V). All mutants are identical to the original wild-type α1-AT sequence except for their reactive site loop region (Fig. 4.2 A). A comparison of the reactive site loop residues P6-P'4 of the α1-AT variants are presented (Fig 4.2 B). Ribbon representations of the backbone of the crystal structure of α1-AT wild-type as well as the reactive site loops of its variants, Mut (A-E), are also presented (Fig 4.3). Mutated residues P6, P4, P1 & P'1 are shown in ball and stick representation and are solvent-exposed.

Mut A was the first to be designed, as it is widely known that P1 cysteine is an essential primary requirement for the recognition and cleavage by the HCV NS3 protease. We wanted to test whether the cysteine residue mutation alone in the RSL of α1-AT was sufficient to promote cleavage by the NS3 protease. The combination of RSL mutations introduced into Mut B was predicted to be appropriate to target the serpin to NS3 protease specifically. Mut C & D were generated later to investigate whether any of the P4 or P6 substitutions had a significant impact on improving the specificity of the mutant A.
Figure 4.2. α1-AT and bio-engineered variants. (A) Structural representation of α1-AT and its 5 different bio-engineered variants. All 6 constructs are identical from amino acid (1-384) and (395-426). The reactive site loop (RSL) of α1-AT consists of amino acid (385-394). (B) Detailed representation of the RSL to show the mutated amino acids within the RSL of each α1-AT variants. Red amino acids indicate that it has been mutated from the original α1-AT respective sequence.
Figure 4.3. Ribbon representations of the backbone of the crystal structure of the $\alpha_1$-AT wild-type and the reactive site loops of its variants mutant (A-E). Reactive site loops from P6-P’4 residues are shown in red. Mutated residues are shown in ball and stick representation.
towards the NS3 protease. Finally, Mut E was included to test if inhibition by serpin can still be achieved when the reactive site loop can no longer be cleaved by the NS3 protease. The proline in P'1 should inhibit cleavage, as P'1 mutation from serine to proline was previously proven to be recognizable but uncleavable by the NS3 protease (Paolo et al., 2000; Richer et al., unpublished results).

4.3 Purification of α1-AT and Its Variants

Plasmids encoding α1-AT and its variants were transformed and expressed in BL21 E.coli using the same protocols as the HCV NS3 protease. 1 L of bacterial cultures expressing α1-AT and the five variants were prepared for purification so that enough materials will be available to perform all related experiments from the same stock. However, some modification to the original protocol of purifying NS3 protease had to be introduced, since the washing and elution buffers used to elute pure NS3 protease did not result in the elution of high purity α1-AT (data not shown). Adjustments of the values of the pH step gradients (2nd wash-pH 5.8, 3rd wash-pH 5.2 and elution-pH 3.9) were found to be sufficient in order to yield α1-AT and its variants at ~96% purity (Fig. 4.4 & 4.5). Despite the success of purifying α1-AT and Mut A,B,D & E, purification of mutant C proved to be difficult. There were 3 lower molecular weight protein bands that were difficult to remove during the purification of mutant C (Fig. 4.5).

All purified α1-AT and variants were subjected to buffer exchange dialysis to avoid long-term protein instability in low pH environment (pH 3.9). Dialysis buffer consisted of 50mM Heps, 150 mM NaCL, 10mM DTT and pH 7.5. DTT was included to prevent the formation of disulfide bridges between serpins via the newly introduced cysteine residue at
Figure 4.4. Purification of bacterially expressed histidine-tagged α₁-AT and its variants. Representative chromatogram of α₁-AT purification using Ni²⁺ binding-interaction chromatography. Inset. SDS/PAGE (lanes 1 and 2) and western blot (lane 3). This procedure yielded recombinant α₁-AT and its variants (~49 kDa) that were essentially pure and intact, as determined by the Coomassie blue staining (lane 2) and western blot analysis (lane 3). Histidine-tagged α₁-AT were detected by western blot using anti-histidine antibody.
**Figure 4.5. SDS-PAGE analysis of α₁-AT and all variants (Mut A-E).** SDS-PAGE result shows that FPLC purified α₁-AT and its variants are ~96% pure except for mutant C ~77% pure as determined by the densitometer. Recombinant histidine-tagged α₁-AT and all variants are ~ 49 kda.
the P1 site. The 3 X 1 hours buffer exchange dialysis procedure using a 10kD cut off dialysis membrane was carried out according to manufacturer’s recommendation and was sufficient to raise the pH of the purified protein solution to pH 7.5. Dialyzed proteins were then aliquoted in 200 µl aliquot, snap-frozen and stored at -86°C.

4.4 Bio-engineering of KDEL Mutant A & D

The carboxy terminal tetra-peptide, KDEL, is found in many lumenal ER resident proteins. When attached onto various molecules, it results in their localization to the endoplasmic reticulum (ER), showing that it is a necessary sequence determinant for this process (Pelham, 1989). ER localization can be achieved in two ways: a) prevention of ER resident proteins from entering newly formed transport vesicles and b) retrieval of those ER residents that escape. Inclusion of a KDEL motif into the amino acid sequence of the desired protein is a useful tool in order to direct it to the ER (Teasdale and Jackson, 1996).

A mammalian expression plasmid containing C-terminally KDEL tagged α₁-AT sequence was constructed and cloned in the p-CMV vector by our collaborator at Yale University (Appendix III). I wished to express KDEL-tagged α₁-AT, because it has been shown that the non-structural proteins of HCV are associated with the ER membranes, confirming the hypothesis that the ER is the site of membrane-associated HCV RNA replication (Mottola et al, 2002).

In addition to the original α₁-AT-KDEL, I have also bio-engineered KDEL Mut A & D according to the PCR protocols described in Section 2.9. Clustal W alignment of all KDEL constructs and translated amino acid alignments are provided in (Appendix VI & VII). The KDEL constructs were sequenced to confirm that no additional mutations were
introduced. Preliminary experiments investigating the intracellular localization of \( \alpha_1 \)-AT-KDEL variants by immuno-fluorescents were performed by Dr. Pamela Hamill (Chapter 6).
Chapter 5: Characterization of the Ability of α_{1}-AT and RSL Variants to Form Complexes with and Inhibit the Proteolytic Activity of Elastase

5.1 α_{1}-AT, a Natural Inhibitor of Neutrophil and Pancreatic Elastase

In order to determine whether purified recombinant α_{1}-AT and its variants were biologically active, it was necessary to test their ability to inhibit pancreatic elastase, one of the natural protease targets of α_{1}-AT (Carrell, 1986). There have been various previous attempts to construct variants of α_{1}-AT with improved inhibitory properties against elastase (Kuisetti and Travis., 1996; Matheson et al., 1986). Hence, it was also of interest to determine whether the α_{1}-AT variants created in our laboratory were more effective inhibitors of elastase than the wild-type α_{1}-AT.

5.2 Complex Formation Between α_{1}-AT, Mutant A & D with Elastase

A common way to investigate the biological inhibitory function of serpins is to determine whether they can form SDS-stable complexes with their target protease (see Section 1.3.3). To perform complex formation experiments, suitable reaction buffers and reaction conditions had to be established. After several experiments to optimize the conditions required for α_{1}-AT/elastase complex formation, it was found that a buffer comprised of 50mM Hepes, 150mM NaCl, 0.1% triton-X, 10 mM DTT, pH 7.5 along with 9 nM of elastase and 1.8 μM of purified α_{1}-AT or its variants in a total volume of 100μl
reaction mixture was the optimal condition for the detection of complex formations between 
\(\alpha_1\)-AT and its variants with elastase. Complex formation experiments were conducted for 1 hour at 31°C.

\(\alpha_1\)-AT and its variants have the molecular mass of approximately 49 kDa, whereas pancreatic elastase has the molecular weight of about 21 kDa. Therefore, the molecular weight of the SDS-stable complex between \(\alpha_1\)-AT and its variants with pancreatic elastase should be \(\sim\) 70 kDa. Figure 5.1 A shows that \(\alpha_1\)-AT and all its variants were stable by themselves. When elastase was added, some of the \(\alpha_1\)-AT, Mut A and Mut D were cleaved (molecular weight of \(\sim\)45 kDa) and some formed SDS-heat stable complexes with elastase (molecular weight of \(\sim\)70 kDa). These results suggest that \(\alpha_1\)-AT, Mut A & Mut D produced in the laboratory still retained their biological functions and could represent effective elastase inhibitors.

Cleavage of \(\alpha_1\)-AT and other serpins can occur in 2 different circumstances: 1) \(\alpha_1\)-AT forms an irreversible inhibitor complex with the target protease, which subsequently gets degraded, or 2) \(\alpha_1\)-AT is cleaved by target or non-target protease, usually at the RSL region, without the formation of stable inhibitor complexes (Fig. 5.2) (Janciauskiene., 2001). Such cleavages generate a 4 kDa carboxyl-terminal fragment of \(\sim\) 36 residues, which remains non-covalently bound to the cleaved \(\alpha_1\)-AT. Therefore, upon the addition of denaturing agent (SDS) and heating at 95°C, this 4 kDa fragment is released and the resulting cleaved serpin has a lower molecular weight of \(\sim\)45 kDa as shown in the case of \(\alpha_1\)-AT, Mut A & D (Fig 5.1 A). In the case of \(\alpha_1\)-AT, Mut A & D, both pathways are probably occurring simultaneously, as shown by the formation of both SDS stable complexes and cleaved serpins.
Figure 5.1. Serpin complex formation with elastase. (A) 1 hour complex formation reaction of 1.8 µM of α₁-AT and all its variants vs 900 nM elastase. Only α₁-AT, Mut A, and D can form a complex with elastase. Mut B, C & E were all degraded during incubation reaction with elastase. (B) Increasing amount of complex formation was observed with the increasing [elastase] (nM) used vs 1.8 µM of α₁-AT, Mut A & Mut D (1 hour incubation). (E = elastase ; EI* = enzyme/inhibitor complex ; I = inhibitor ; I^c = cleaved inhibitor). Buffer compositions for these experiments are described in Materials & Methods sections 2.11 & 2.13. All western blots were probed with RGS anti-histidine mouse antibody.
Figure 5.2. Branched pathway mechanism of serpins as suicide substrate inhibitors. I represents the serpin and E represent the protease. EI represent protease/serpin intermediate, EI* represent protease/serpin complex and Ic represent cleaved serpin.

In order to further investigate the complex formation between α1-AT, Mut A & D and elastase, experiments using varying amounts of elastase (90, 120, 180, 300 & 900 nM) reacting with the three serpins were carried out. In general, increasing elastase concentration led to greater amounts of complex formation. However, 180 nM of elastase was sufficient to produce a strong and visible complex between α1-AT, Mut A & D (Fig 5.2 B). In most cases, increasing the concentration of elastase to more than 180 nM did not increase the amount of complex formation. This could be due to the limited amount of active serpins that were present to form complexes with elastase. Nonetheless, results from these experiments showed that the amount of SDS-heat stable complex formation between α1-AT, Mut A & D was dependent upon the concentration of elastase, indicating that complex formation was specific.

5.3 Degradation of Mutant B, C & E by Elastase

Muts B, C & E, unlike α1-AT, Mut A & D, failed to form SDS-heat stable complexes with elastase. Not only did they fail to complex with elastase, they were also completely
degraded by elastase by the end of the 1 hour complex formation reaction, as indicated by the failure to detect any protein band of the right molecular weight (Fig 5.3). This phenomenon is very unusual and interesting. The main difference between α1-AT, Mut A & D versus Mut B, C & E was that the former three serpins do not contain any P6 substitution, whereas the latter three serpins all had their P6 residues mutated from leucine to aspartic acid (Section 4.2). There is some evidence that the presence of hydrophobic leucine in P6 is important to allow the conformational change required for a stable complex formation between α1-AT and its target protease (Warshel et al., 1991). Insertion of a bulkier and charged side chain such as arginine or aspartic acid in this case favors the substrate pathway in which the serpin is cleaved into its 45 kDa form and then released from the target protease still in its active form (Dufour et al., 2001). This might explain why Mut B, C & E were all further degraded by active elastase, while the 45 kDa cleaved forms of α1-AT, Mut A & D were not degraded further since all the elastase in the solution was presumably depleted due to complex formation.

Upon a more detailed inspection, it was also found that the degradation of Mut B, C & E into their 45 kDa form without complex formation was time-dependent, as after 45 minutes of incubation between these serpins with elastase, all three were completely degraded by elastase (Fig. 5.3). The mechanism of degradation of the three serpins may be different, as Mut B was degraded more readily (complete degradation after 30 minutes) than Mut E and C. The reasons for the degradation of serpins in general are still largely unknown, but it can be speculated that since elastase is a non-specific enzyme able to recognize and cleave different amino acid sequences of different substrates, the excess active elastase in the solution probably managed to cleave newly exposed sites following the cleavage of the P1-
Figure 5.3. Degradation of Mut B, C & E by elastase. 1.8 μM of Mut B, C & E were degraded by 900 nM of elastase as time progresses. All of Mut B, C & E were completely degraded by elastase after 45 minutes. (E = elastase; EI* = enzyme/inhibitor complex; I = inhibitor; I° = cleaved inhibitor). Buffer compositions for these complex experiments are described in Materials and Methods section 2.11 & 2.13.
P'1 in the reactive site loop in a cascade of sequential cleavages. These sequential cleavages eventually lead to a complete degradation of the serpins.

Studies of the functional and conformational polymorphism of inhibitory serpins show that under certain physiological conditions, serpins can also undergo conformational change due to mutation, chemical modification or interaction with other molecular species as described in Figure 1.5 (Janciauskiene., 2001). Furthermore, one of the most serious and persistent concerns about affinity chromatography purification of serpins is the potential for product denaturation and conformational change. Elution of the bound protein is often achieved by passing low pH buffer over the column (pH 2.5-4.0). Detailed studies of protein conformation under these conditions have documented permanent conformational changes as the result of such exposure (Gagnon., 1999). An increase tendency toward aggregation or polymerization and, less commonly, proteolysis of the product are also the result of eluting the proteins at low pH (Gagnon., 1999). Further evidence of α₁-AT polymerization was presented, in December of 2002 (Devlin et al., 2002). Devlin et al. showed that the native serpin architecture is extremely sensitive to mutation and environmental factors. The polymerization of α₁-AT at pH 4.0 occurred with initial fast rate, reversible conformational change that resulted in partial loss of secondary structure, followed by subsequent slow rate that is irreversible resulting in the stabilization of the dimmers and subsequent polymer extension. Therefore, it is very likely that a large percentage of the α₁-AT and variants purified from the pH 3.9 elution had distorted conformation and formed polymers, which made them inactive and susceptible to protease degradation by the excess active elastase in the solution. This hypothesis is supported by the fact that most of the elastase at low concentration (~ 900 nM) was able to form SDS-stable complexes with enough active α₁-AT.
Thus, inactive α₁-AT was not degraded or degraded very slowly. However, when high concentrations of elastase (~9 μM) were added to α₁-AT, inactive α₁-AT in the solution, including any small amounts of elastase/α₁-AT complexes that may have been formed, were completely degraded by the excess elastase (data not shown). Further experiments including running a low pH, non-denaturing gel can be performed to investigate the polymerization of α₁-AT at pH 3.9.

5.4 Inhibition Assay of α₁-AT and Its Variants Towards Elastase

1.3 μM of α₁-AT was pre-incubated with 9 nM of elastase for a range of times prior to the commencement of enzymatic assays as described in Section 2.14. The results of this investigation are presented in Figure 5.4. They clearly show that the pre-incubation time allowed for the interaction between α₁-AT and elastase is crucial in determining the residual elastase activity during the enzymatic assay. As the pre-incubation time between α₁-AT and elastase increased, the residual activity of elastase decreased. Forty five minutes of pre-incubation time between α₁-AT and elastase in this case was sufficient to fully inhibit the activity of elastase.

Assays assessing the inhibition of elastase by α₁-AT and variants were performed using 1.3 μM of each serpin and pre-incubating it with 9 nM of elastase for 45 minutes in the buffer described in the materials and methods, section 2.11. The inhibition assay was performed in duplicate and was allowed to proceed for 3 hours at 31°C, with continuous readings of elastase activity taken every 2 minutes. Results showed that the control of elastase alone after 45 minutes of pre-incubation was still active, with residual activity of 100% relative cleavage efficiency against which the activity of the other α₁-AT variants was
Figure 5.4. Residual activity of elastase with $\alpha_1$-AT after different incubation time. The length of incubation time that was allowed between 9 nM of elastase and 1.3 $\mu$M of $\alpha_1$-AT was directly affecting the residual activity of elastase in vitro. 45 minutes of incubation time between elastase and $\alpha_1$-AT was sufficient to completely inhibit elastase activity in vitro ($t_{1/2} = 3$ min). Enzymatic assay was performed for 3 hours and in duplicate. Enzymatic assays conditions and buffer requirements are described in the Materials and Methods sections 2.11 & 2.14.
Figure 5.5. Inhibition of elastase by α₁-AT, Mut A & D. Pre-incubation time between elastase and serpins were 45 minutes and enzymatic assay was done for 3 hours. The 9 nM elastase control is active and the addition of 1.3 μM of α₁-AT, Mut A & D completely inhibit elastase activity, whereas 1.3 μM of Mut B, C and E did not inhibit elastase activity. Experiments were done in duplicate. Buffer compositions and enzymatic assays conditions for this experiment are described in the Materials and Methods sections 2.11 & 2.14.
compared. $\alpha_1$-AT, Mut A & D completely inhibited elastase activity as expected. However, Mut B,C & E were not able to inhibit the activity of elastase (Fig 5.5). These results agree with the results from complex formation studies and allow for the conclusion that only serpins that can form SDS-heat stable complexes with elastase can disrupt the catalytic active site of elastase, rendering it inactive.

5.5 Titration and Progress Curve Analysis of $\alpha_1$-AT and Its Variants Towards Elastase

Titration experiments for $\alpha_1$-AT, Mut A & D against elastase were carried out in the same way as the inhibition assay, except for the fact that varying amounts of serpin were used to pre-incubate with elastase. The purpose of this experiment was to determine the Stochiometry of Inhibition (SI) and Specific Inhibition Constant ($K_i$) of each of the serpins that inhibit elastase. The amount of each serpin used was (nM): 0, 1.2, 2.5, 4.9, 6.2, 9.3, 12.3, 18.5, 30.9, 61.7, 123.4. An overlay of $\alpha_1$-AT, Mut A and D titration inhibition curves showed that the inhibition constant ($K_i$) of the three serpins vs elastase were identical (Fig 5.6 A). The ($K_i$) values of $\alpha_1$-AT, Mut A and Mut D against elastase were all 53 pM (Fig 5.6 B,C,D). These ($K_i$) values are all in the pM range, meaning that $\alpha_1$-AT, Mut A & D are all excellent inhibitors of elastase. The Stochiometry of Inhibition (Si) values for $\alpha_1$-AT, Mut A & D against elastase were 2. This means that approximately 2 mol of serpin are required to inactivate 1 mol of elastase, indicating that equal amounts of these serpins went through the complex formation inhibitory and substrate pathways (Figure 5.6 B,C,D).

The difference between the progress curves analysis and the titration curve analysis of $\alpha_1$-AT, Mut A & D vs elastase was the lack of pre-incubation time in the progress curve.
Figure 5.6. Inhibition constant of $\alpha_1$-AT, Mut A and D vs Elastase. Inhibition curves obtained from the titration experiments of $\alpha_1$-AT, Mut A and D vs 9 nM of elastase. The amount of each serpin used were (nM): 0, 1.2, 2.5, 4.9, 6.2, 9.3, 12.3, 18.5, 30.9, 61.7, 123.4. Enzymatic assays were done for 3 hours in triplicates with 45 minutes of pre-incubation between elastase and serpins. Buffers composition was described in Materials and Methods section 2.11 & 2.14. (A) Over-lay of $\alpha_1$-AT, Mut A and D inhibition curves shows that the ($K_i$) and (SI) of the three serpin vs elastase are nearly identical. (B) Inhibition curve of $\alpha_1$-AT vs elastase. $K_i = 53$ pM. (C) Inhibition curve of Mut A vs elastase. $K_i = 53$ pM. (D) Inhibition curve of Mut D vs elastase. $K_i = 53$ pM. The (SI) for all three serpins are 2.
analysis. The purpose of performing enzymatic assays without pre-incubation was to find out whether serpins can compete and inactivate elastase in the presence of competitive elastase substrates. This study is important because in vivo it is essential for an effective inhibitor to be able to out-compete the natural substrate of the protease. Figure 5.7 shows that increasing the amount of α1-AT, Mut A & D (nM) decreased the activity of 9 nM of elastase to different extents. Higher amount of serpins inhibit the activity of elastase to a greater extent. Enzymatic assays were done for 3 hours. The inhibition of elastase by α1-AT, Mut A & D obeyed slow-binding inhibition kinetics, as indicated by the biphasic plots, where maximal inhibition was achieved more rapidly with increasing concentrations of α1-AT, Mut A & D. The biphasic plots also showed that the serpins and proteases require time to interact and form complexes which deactivate the protease (illustrated by the linear part of the curve). Once complexes were formed, the process was irreversible and, as a result, all protease had been inactivated (illustrated by the asymptotic component of each curve) (Fig. 5.7).
Figure 5.7. Progress curve of $\alpha_1$-AT, Mut A & D vs elastase. Increasing amount of $\alpha_1$-AT, Mut A & D decreased 9 nM of elastase activity. The curves also show tight-binding inhibition pattern of serpin/enzyme inhibition. Enzymatic assays were done for 3 hours with no incubation time between elastase and serpins. Experiments were done in duplicates and buffer compositions are described in the Materials and Methods section 2.11. & 2.14.
Chapter 6: Characterization α₁-AT and Its Variants (Mut A-E) with Complex Formation Studies and Inhibition Assays Against HCV NS3 Protease

6.1 Serpins as Candidates for Novel Protein Therapeutics Against HCV NS3 Protease

As described in the introduction, there has been evidence of HCV NS3 protease interaction with host cell serpins in vitro. Cl inhibitor and α-2 antiplasmin, both of which belong to the serpin family, can interact with HCV NS3 protease, resulting in proteolysis of the serpins and production of higher molecular weight SDS-stable complexes (Drouet et al., 1999). These observations suggest that the use of serpins engineered to recognize and inhibit the NS3 protease may be a promising new strategy for the development of HCV therapeutics.

6.2 Complex Formation Between α₁-AT and Its Variants Against NS3 Protease

After characterizing the inhibitory properties of the α₁-AT and variant serpins toward elastase (Chapter 5), the ability of these serpins to inhibit and/or form complexes with NS3 protease was assessed next. The optimization of conditions necessary to conduct complex formation experiments between α₁-AT and variants against NS3 protease was necessary before performing in vitro complex experiments. It was determined that 50mM Hepes, 150mM NaCl, 0.1% triton-X, 10mM DTT, pH 7.5, 15µM NS4A cofactor, 20µl of NS3 protease and 1.3 µM of purified α₁-AT or its variants in a total volume of 100µl reaction mixture were the optimal conditions under which to detect the formation of complexes
between $\alpha_1$-AT and its variants with the NS3 protease. It was also established that 15 minutes of pre-incubation between the NS4A cofactor and the NS3 protease at 31°C was necessary to activate the NS3 protease prior to performing complex experiment with serpins for a period of 4 hours at 31°C. $\alpha_1$-AT and all its variants were mixed with the NS3 protease as described above, and the complex formation reactions were subjected to SDS-PAGE and Western blot analyses. Results show that the NS3 protease which has a molecular weight of ~23 kDa can interact with $\alpha_1$-AT and all its variants (~49 kDa) except for Mut E to form SDS-heat stable complexes of molecular weight ~72 kDa (Fig 6.1 A & B). Interestingly, the $\alpha_1$-AT that was able to react with the NS3 protease, all formed SDS-stable complexes with NS3 protease and did not go through the substrate pathway as described earlier in Section 5.2. It was expected that $\alpha_1$-AT would not be cleaved as a substrate by the NS3 protease due to its strict cleavage site specificity requiring a cysteine in P1. However, the fact that $\alpha_1$-AT can still form SDS-stable complexes with the NS3 protease without its reactive site loop (RSL) being cleaved first is rather unique. It is widely accepted that in order to form a covalent SDS-stable complex between protease and serpins, the cleavage of the RSL is necessary for the migration of the proteinase to the opposite pole of the serpin (Wilczynska et al., 1995; Janciauskiene., 2001). This observation leads one to believe that the general scaffold, especially the reactive site loop region of $\alpha_1$-AT, had a suitable structure to allow its recognition by the NS3 protease and subsequent complex formation. This supports our initial reasoning behind using $\alpha_1$-AT as a starting scaffold (section 1.3.4) upon which to design specific NS3 inhibitors. The phenomenon of complex formation between a protease and serpin without RSL cleavage is a new concept and the molecular mechanisms underlying this complex formation are yet to be elucidated.
Figure 6.1. Complex formation between $\alpha_1$-AT and its variants vs NS3 protease. (A & B) All of the 1.3 $\mu$M of serpins can form SDS-heat stable complexes with 20$\mu$l of NS3 protease except for mutant E. Complex experiments were done for 4 hours with 15 minutes of pre-incubation between NS3 protease & NS4A cofactor. (EI* = enzyme/inhibitor complex ; I = inhibitor ; Ic = cleaved inhibitor). Buffer compositions for these experiments are described in Materials & Methods sections 2.6 & 2.13.
Mut A, B, C & D were all proteolytically cleaved by the NS3 protease because some portion of the serpins went through the substrate pathway. These observations agree with our prediction that the introduction of cysteine in the RSL P1 position of α1-AT would make it a pseudo-substrate for the NS3 protease. We also found that although the four serpins can form SDS-stable complexes with the NS3 protease, Mut D formed most complexes (Fig 6.1 A & B). We still do not fully understand why Mut D, which has P1 and P4 mutations, was able to form the most complexes with NS3 protease. There are three possible hypotheses to explain these results: 1) The pooled stock of Mut D simply contained more biologically active serpin that can form SDS-stable complexes with the NS3 protease as compared to the other stock of serpins. 2) The P1, P4 mutation in Mut D were better recognized by the NS3 protease than the P1 mutation in Mut A, thus leading to more complex formations. However, the addition of P6 mutation in Mut B & C improved substrate recognition to such a great extent that most of the Mut B & C that interacted with the NS3 protease went through the substrate pathway instead of the complex inhibition pathway. Therefore, this observation leads to the question of how efficiently a serpin should be recognized by the NS3 to achieve the most complex formation. 3) Mut D can form more complexes with NS3 protease than Mut A due to the same reason as above. However, the addition of the P6 mutation in Mut B & C could have disrupted the efficient conformational change as described in Section 5.3. The P6 substitution in the RSL of α1-AT should prevent Mut B & C from forming complexes with the NS3 protease (Dufour et al., 2001). Therefore, these results lead one to question whether the complex formation between all the α1-AT serpin mutants and the NS3 protease follow the traditional pathway requiring RSL cleavage and insertion or follow the possible alternative as seen in α1-AT/NS3 protease complex pathway without RSL cleavage. Small amounts of
complex formation can also be seen between the serpins and the NS3 protease without the addition of NS4A cofactor, due to the intrinsic activity of the NS3 protease alone.

Mut E was not proteolytically cleaved by the NS3 protease. This is because the proline substitution has been shown to be uncleavable when introduced in the P’1 site of the NS3 protease peptide inhibitor (Paolo et al., 2000). Although a proline in P’1 is a potent peptide inhibitor against the NS3 protease in vitro, it is not effective against the FLNS3 (Richer et al., unpublished result). Moreover, when the substitution was made in the RSL of α1-AT, it failed to form a complex with the NS3 protease (Fig 6.1 B), probably due to the major distortion of the α1-AT RSL from the introduction of proline residue. Finally, none of the serpins were completely degraded by excess NS3 protease because, unlike elastase which can cleave a broad range of substrates, the NS3 protease specifically cleaves only after cysteine residue.

6.3 Studies Investigating the Effect of Enzyme: Inhibitor Ratio and Time of Interaction Upon Complex Formation Between α1-AT and Variants and NS3 Protease

Figure 6.2 shows the effect of varying the amount of NS3 protease on the amount of complex formed in a reaction using 1.3 μM of either α1-AT, Mut A, D & C respectively. The increased amount of SDS-stable complexes formed with the increased amount of the NS3 protease suggests that the complex formation was ratio-dependent. 20μl of NS3 protease usually resulted with the highest amount of complex formation. Figure 6.3 shows that increasing the incubation time between α1-AT, Mut A, D & C and the NS3 protease also increased the amount of complex formation. Additional incubation times of 6 hours and 18
Figure 6.2. Complex formation between α₁-AT, Mut A, D, & C with increasing [NS3]. SDS-stable complex formation between α₁-AT, Mut A, D, and C with NS3 protease increases as [NS3] increases. Complex experiments were done for 4 hours. All complex formation reactions were done with 15 minutes of pre-incubation between NS3 protease & NS4A cofactor. (EI* = enzyme/inhibitor complex; I = inhibitor). Buffer compositions for these experiments are described in Materials & Methods sections 2.6 & 2.13.
Figure 6.3. Complex formation between α₁-AT. Mut A, D, & C with NS3 protease at different time points. Complex formation between α₁-AT. Mut A, D, & C with NS3 protease increases as the amount of time allowed for the complex reactions increase. All complex formation reactions were done with 15 minutes of pre-incubation between NS3 protease & NS4A cofactor. (EI* = enzyme/inhibitor complex ; I = inhibitor; Ic = cleaved inhibitor). Buffer compositions for these experiments are described in Materials & Methods sections 2.6 & 2.13.
hours were also tried, but did not significantly increase the amount of complexes formed (data not shown). Therefore, 4 hours incubation time between the serpins and the NS3 protease was optimal. This result supports the idea that most of the serpins in the pooled samples were inactive since after 4 hours of incubation, most active serpins had already formed complexes with the NS3 protease and there were still plenty of unreacted serpins and the NS3 protease left in the solution. None of the remaining serpins could contribute to the formation of more complexes after 4 hours. However, there was still active NS3 protease left because as described in the next section, when similar concentrations of serpins and NS3 protease were tested in vitro for inhibition assays, NS3 protease activity was not inhibited.

6.4 Inhibition Assay of α₁-AT and Its Variants Towards NS3 Protease

Although five of the serpins (α₁-AT, Mut A, D, B & C), could form SDS-stable complexes with the NS3 protease, surprisingly they were not able to inhibit the NS3 protease activity in vitro completely (Fig. 6.4). In the control reaction in which the NS3 protease and the NS4A cofactor were pre-incubated together for 4 hours, the NS3 protease was still active. α₁-AT, Mut A,B,C slightly decrease the NS3 protease activity by approximately 20%. Mut D, however, decreased the NS3 protease activity by ~50%, while Mut E did not decrease the NS3 protease activity. These results are in agreement with those from the complex experiment in which Mut D formed the most complexes with the NS3 protease compared to the other mutants, while Mut E was not able to form any detectable complexes with the NS3 protease. The reason for the incomplete inhibition of the NS3 protease could be due to the low amount of active serpins present in the pooled samples as discussed before in Chapter 5.

Therefore, only a small proportion of active serpin was capable of complex formation with
Figure 6.4. Inhibition of NS3 protease by α₁-AT and its 5 variants. The 10 μl of NS3 protease control was active (~450 RFU = 100% relative cleavage efficiency). 1.3 μM of α₁-AT and all its variants cannot completely inhibit NS3 protease activity, although Mut D reduced NS3 protease activity level by a factor 1/2. Inhibition assay were done for 2 hours with 15 minutes of pre-incubation between NS3 protease & NS4A cofactor. Enzymatic assays conditions and buffer compositions for these experiments are described in Materials & Methods sections 2.6 & 2.14.
the NS3 protease, and subsequently able to inhibit NS3 protease activity. The inhibition assays using low concentrations of NS3 protease were challenging to perform, because the sensitivity of detection of fluorescence from the IQFS was low. Hence, lower enzyme activity resulting in lower rates of IQFS cleavage was difficult to detect. One of the possible solutions to this problem could be producing higher amounts of biologically active serpins via a different method of purification and re-evaluating the ability to inhibit NS3 protease activity. In addition, Mut A, B & C also tend to go through the substrate pathway more than through the complex pathway (Fig 6.2 & 6.3). An alternative way to deal with the low concentration of active serpins is to genomically express the recombinant α1-AT-KDEL construct in human cell cultures, so that the serpins are over-expressed in cells without being subjected to processing procedures such as purification and dialysis, which might inactivate some of the serpin.

6.5 Complex Formation Between α1-AT and Its Variants Against FLNS3

Preliminary studies were performed to investigate if the α1-AT and its variant serpins can form complexes with recombinant FLNS. The reaction conditions and buffer requirements for performing these FLNS3 complex reactions were identical to those with the NS3 protease (Section 6.2), except 20μl of FLNS3 was used instead of 20μl of NS3 protease. Western blot results from these experiments showed that none of the serpins can form detectable SDS-stable complexes with FLNS3 (Fig. 6.5 A & B). However, Mut A, B, C & D were still recognized by the FLNS3 as a substrate and cleaved. These results suggest that perhaps the overall scaffold of the FLNS3 containing the helicase domain does not permit
Figure 6.5. Complex formation between FLNS3 with $\alpha_1$-AT and all its variants. (A) 1.3 μM of $\alpha_1$-AT, Mut A and D are cleaved by 20 μl of FLNS3, however no complex is seen at ~116 kDa. FLNS3 were also processed with the addition of NS4A cofactor. (B) 1.3 μM of Mut B and C were cleaved by 20 μl of FLNS3 but Mut E is not cleaved due to the proline in P1 position. Complex experiments were done for 4 hours. The lack of complex formation could be attributed to many different factors as discussed in the discussion sections. FLNS3 was detected with anti FLNS3 anti-body. Serpins are detected with anti-histidine antibody. Complex experiments were done for 4 hours with 15 minutes of pre-incubation between FLNS3 & NS4A cofactor. (FLNS3 = Full length NS3; FLNS3* = FLNS3 cleaved ; I = inhibitor ; Ic = cleaved inhibitor). Buffer compositions for these experiments are described in Materials & Methods sections 2.6 & 2.13.
Figure 6.6. Co-localization of \( \alpha_1 \)-AT-KDEL, Mut A-KDEL & Mut D-KDEL with NS3 in human cells. (A-B) Subcellular localization of transiently expressed \( \alpha_1 \)-AT-KDEL motif fusion, Mut A-KDEL, and Mut D-KDEL in UNS3-4A cells, which also express NS3-4A protease/cofactor. Double immuno staining was performed to detect \( \alpha_1 \)-AT-KDEL and variants using a FLAG-epitope tag-specific antibody together with NS3 using NS3 polyclonal antibody. (C) Overlay of the double immuno staining shows that both the serpins-KDEL and HCV NS3 were expressed in the same compartment in the cell, which is presumably the ER. (Work done by Dr. Pamela Hamill)
complex formations to occur. The other reason for the absence of complex formation could be that the amount of active FLNS3 was too low to detect any complex formations. A similar observation has been made with the NS3 protease (data not shown). Note that FLNS3 was also degraded with time in the presence of the NS4A cofactor after 4 hours of incubation without the serpins. The internal cleavage of FLNS3 only in the presence of the NS4A cofactor has been previously reported (Yang et al., 2000).

6.6 Sub-cellular Targeting of Recombinant Serpin-based Inhibitors of HCV NS3 Protease in Human UNS 3/4A Cells

Dr. Pamela Hamill performed transfection of mammalian expression plasmids encoding α₁-AT-KDEL, Mut A-KDEL and Mut D-KDEL in human cells (UNS 3-4A) which inducibly expressed the HCV NS3/4A protease complex with the control of tetracycline. Figure 6.6 A & B shows that novel variants of the serpin α₁-AT fused to a C-terminal KDEL ER retention motif (α₁-AT-KDEL) were transiently expressed in UNS 3-4A which had been induced to expressed the viral NS3-4A protease in tetracycline-regulated manner. The results also showed that HCV NS3-4A, α₁-AT-KDEL and variants were apparently all co-localized in the ER compartment of the cells (to be confirmed by using ER KDEL-receptor marker). The overlay of α₁-AT-KDEL and variants with the NS3-4A protease expressed in the cells suggests that the co-localization of the serpin/protease in same subcellular-compartment might improve the probability of inhibition of the HCV protease by bio-engineered serpins (Fig. 6.6 C).
6.7 Summary

Chapters 4, 5 and 6 present results of successful bio-engineering of α₁-AT variants and α₁-AT-KDEL variants. The purification of these serpins was largely successful, with some of the serpins retaining their biological activity as shown by the ability to form complexes and be cleaved by elastase and the NS3 protease. Complex formation studies between α₁-AT and variants with elastase and the NS3 protease showed that only α₁-AT, Mut A & D can form SDS-stable complexes with elastase, whereas all the serpins except for Mut E can form a complexes with the NS3 protease. Inhibition studies showed that α₁-AT, Mut A & D can completely inhibit elastase activity in *vitro*. However, none of the serpins can completely inhibit NS3 protease activity in *vitro*. Despite the lack of marked inhibition of the NS3 protease by the novel α₁-AT variants and until the problems of inactive serpins proposed in the discussions are addressed by performing purification of α₁-AT and variants using an increasing imidazole gradient, the possibility that α₁-AT and its variants might be able to inhibit HCV NS3 protease activity in *vitro* and in *vivo* is still viable and commands further investigation.
Chapter 7: Conclusion

7.1 Current View of HCV NS3 Serine Protease Inhibitors

In light of the success of protease inhibitors in controlling HIV infections, many research groups in both industry and academia have chosen to embark on programs aimed at identifying potent and selective inhibitors of the NS3 protease as candidate anti-HCV drugs. In the decade that has elapsed since the discovery of HCV, marked progress has been made towards understanding the novel molecular structure and functions of the HCV NS3 protease.

Protease inhibitor discovery began with natural product screenings and substrate-derived analogue-based drug-design, soon progressing to the incorporation of mechanism-based drug design strategies such as using serpin. Today, the design of protease inhibitors involves a powerful combination of all of these traditional drug discovery approaches and it is supplemented by computer-assisted structure-based inhibitor design using 3-dimensional structures of proteases determined by X-ray crystallography, and rapid robotic assay methods to quickly screen various potential inhibitors. These techniques are currently being refined and expanded to accelerate the discovery process for new protease inhibitors. Coupled with advances in molecular and cellular biology, protein chemistry, microbiology, structural biology and molecular pharmacology, there can be little doubt that the next several decades will see the development of a new class of protease inhibitors.

To date, there have been reports of peptide-based inhibitors of the NS3 protease and many more research groups are probably working on improving the potency of these inhibitors (Landro et al., 1997, Paolo et al., 2000). However, the remarkably shallow and solvent-exposed substrate binding region which has resulted in weak interactions that are
distributed along an extended contact surface between the proteases and the inhibitors, makes
the design of potent small molecule inhibitors a quite challenging task.

7.2 α1-Antitrypsin and Its Variants as Potential Inhibitors Against HCV NS3
Protease and Future Directions

Since the first report of interactions of HCV NS3 protease by Cl INH and α2-
antiplasmin, that can form SDS-stable complex with NS3 protease (Drouet et al., 1999), there
have been no other reports on serpins that can interact with and inhibit NS3 protease. Data in
this thesis showed the ability of α1-AT and its variants Mut A, B, C & D to interact and form
SDS-stable complexes with the NS3 protease. These results provide further evidence of the
possibility of using serpin as an alternative way of inhibiting the HCV NS3 serine protease.
Some of the advantages of using serpin and α1-AT or its variants as NS3 protease inhibitors
include: 1) the availability of extended contact surface needed for proper interaction with the
NS3 protease that was absent when using small peptide based inhibitors. 2) α1-AT is a
naturally occurring human serpin that is produced in the liver and there have been no reports
of diseases linked to the over-expression of α1-AT. Therefore, α1-AT and its variants should
be better tolerated by the immune system and should have limited side effects when applied
as therapeutics in humans. 3) The introduction of α1-AT variants, Mut B (P1,P4 & P6) or
Mut C (P1 & P6) lost their specificity towards elastase, but were still able to form SDS-stable
complex with NS3 protease. This phenomenon reinforces the possibility of using Mut B and
Mut C as good candidates for the design of specific inhibitors of the NS3 protease. 4) α1-AT
and its variants can be used to target the ER of the cells where HCV replication is believed to
occur, thereby improving the chance of interaction between the serpins and the NS3 protease
(Wu., 2001). 5) Serpin is a suicidal inhibitor because it forms an irreversible interaction with the target. It is therefore effective in removing the protease that is bound to it permanently. 6) SP6, a Drosphila serpin that has been mutated to contain cysteine in P1, serine in P'1 and glutamic acid in P6 to target the NS3 protease and the FLNS3, has been shown not only to form SDS-stable complexes with the NS3 protease and the FLNS3, but it also inhibits the protease activity in vitro (Richer et al, in preparation).

Despite the listed advantages of using serpin as an inhibitor of the NS3 protease, there are various stages of research that still need to be performed further to support the hypothesis that serpin can be used as one of the potential therapeutic against the HCV NS3. One of the experiments that is currently being performed in Dr. François Jean’s laboratory explores cell culture systems of transfecting various serpins into cultured human osteosarcoma cells (UNS3-4A) which express the HCV NS3 protease and the NS4A cofactor. These intracellularly expressed recombinant serpins should reduce the problems of dealing with possible inactivation of serpins caused by expression and purification from bacterial cells. Future experiments should also include experiments in Huh-7 cell lines that support the replication of HCV replicon or even using animal models such as chimeric mice both of which will be described later.

7.3 Animal Models and Cellular System for Anti-HCV Drugs Development

Although both peptide-based and protein-based inhibitors of the HCV NS3 protease have been identified, there has been no cell-culture assay available for the propagation of HCV in the laboratory which limits the ability to test whether those protease inhibitors can actually inhibit HCV replication in vivo. For unknown reasons, the propagation of HCV in
cell cultures including human hepatocytes has been very challenging, because available system all suffer from low reproducibility and efficiency (Bartenschlager., 2001; Kato., 2000).

However, the recent development of the replicon system opens up an encouraging possibility for drug discovery (Bartenschlager., 2002). Huh-7 cells were originally isolated from the hepatoma tissue of a man suffering from a well differentiated hepatocellular carcinoma (Nakabayashi et al., 1982). Huh-7 is currently the only efficient cell that allows the proper propagation of the HCV replicon. This replicon system will probably be important for the development of HCV-specific drugs. The advantage of using the HCV replicon system for drug development over screening assays carried out with purified viral protease is that not all compounds that were successful when tested in vitro were effective in blocking viral replication in vivo (Young., 2001). Moreover, HCV proteins form a higher-order multiprotein replicase complex, binding sites for a compound might be accessible in vitro when working with an isolated enzyme but not within an infected cells, where this protein is incorporated into a macromolecular complex (replicase). Therefore, drugs identified by in vitro screening will not necessarily be active at blocking RNA replication in the cell (Bartenschlager., 2002).

Although HCV replicons can be propagated in Huh-7 cells, the production of full viral particles is still not possible. In addition, until recently only two HCV isolates of genotype 1b could be propagated in the cell culture (Lohmann et al., 1999; Guo et al., 2001; Kishine et al., 2002). The understanding of viral pathogenesis and the development of potential drug candidates have been hindered by slow progress in the development of an animal model that can support replicating viruses (Schinazi et al., 1999). Chimpanzees,
which are both expensive and not readily available, have been the only non-human animal model for reproducible propagation of HCV over the past decade, until the recent development of the chimeric mouse model that was suitable for studying the human hepatitis C virus in vivo (Mercer et al., 2001). By transplanting normal human hepatocytes into SCID mice carrying the plasminogen activator transgene (Alb-uPA), mice with chimeric human liver were generated. These mice allow the replication of HCV viral proteins and the release of infectious viral particles through three generations of mice. These mice represent the first murine model suitable for studying HCV in vivo. Work is currently in progress to test adenovirus recombinant α1-AT and SP6 variants in these murine model (in collaboration with Dr. Norman Kneteman, University of Alberta).

7.4 Combination Therapy for HCV Infection

HIV protease inhibitors have proven to be valuable therapeutics in combination with nucleoside reverse transcriptase inhibitors (NRTIs) and non-nucleoside reverse transcriptase inhibitors (NNRTIs) (a drug combination known as “highly active antiretroviral therapy” (HAART)) in the treatment of HIV infections (De Clercq., 2002). Likewise, a successful therapy for HCV will probably have to involve a combination of drugs targeting different stages of viral replication cycles to reduce the chance of developing viral resistance. Possible targets for anti-HCV drug development in the future may include the inhibition of viral entry, the inhibition of IRES, the inhibition of capping, the inhibition NS3/4A serine protease, the inhibition of ribonucleic acid (RNA) helicase and the inhibition of RNA-dependent RNA polymerase. In addition, antisense oligonucleotides or ribozymes may also become part of the cocktail inhibitors to minimize HCV replication. Finally, immunotherapies to enhance
HCV-specific immune responses are also attractive strategies to control HCV infections and to prevent chronic liver disease (Cornberg et al., 2001; Leyssen et al., 2000; Dymock et al., 2000).

In conclusion, the intense search for inhibitors against various targets of HCV replication with the help of rapidly advancing technologies and the development of new cell lines and animal models to test these compounds will probably result in the discovery of potent cocktail inhibitors of HCV replication in the near future.
Chapter 8: References


Appendix I. Tight binding titration of pancreatic elastase by α₁-AT. Elastase (9 nM) was incubated with increasing amount of α₁-AT for 3 hours at 31°C. MeoSuc-A-A-P-V-AMC (50 μM) was added to determine residual activity of elastase. Data analysis revealed a value of $[E_0] = 9.0 \text{ nM}$, $K_i = 53 \text{ pM}$ and $SI = 2$. The α₁-AT concentration is in (nM), $E_0$ is determined by the distance indicated in the graph converted to (nM) accordingly. D and d can be obtained from directly measuring the distance as indicated in the graph. All other $K_i$ are estimated with the same method.
pET-21(+) (Cat. No. 69770-3) is a transcription vector designed for expression from bacterial translation signals carried within a cloned insert. It therefore lacks the ribosome binding site and ATG start codon present on the pET translation vectors. A C-terminal His-tag sequence is available. Unique sites are shown on the circle map. Note that the sequence is numbered by the pBR322 convention, so the T7 expression region is reversed on the circular map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The T7 origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. Therefore, single-stranded sequencing should be performed using the T7 terminator primer (Cat. No. 88337-3).

Appendix II. Restriction map of pET-21 (Novagen). (A) pET-21 vector is ~5.4kbp with ampicillin resistance and multiple cloning sites. (B) pET-21 vector utilize the T7 promoter and include the desired restriction sites, Xho I and Bpu1102 I in the multiple cloning sites. Bpu1102 I and Bsp I are Isoschizomer.
KDEL constructs of α₁-AT, Mut A & Mut D in pShuttle-CMV vector

Appendix III. Restriction map of p-CMV shuttle vector (Stratagene). α₁-AT-KDEL was made and cloned into the p-CMV shuttle vector by our collaborator. Mut A and D-KDEL were made using quick change mutagenesis kit in the lab. Site-directed mutagenesis procedure is described in the Materials and Method section.
Appendix IV

CLUSTALW Result (A1AT, Mutant A,B,C,D & E)

Sequence type explicitly set to Protein
Sequence format is Pearson
Sequence 1: MutantB 400 aa
Sequence 2: MutantE 400 aa
Sequence 3: MutantD 400 aa
Sequence 4: MutantC 400 aa
Sequence 5: A1AT 400 aa
Sequence 6: MutantA 400 aa

Start of Pairwise alignments

MutantB HQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGFQELLHT
MutantE HQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGFQELLHT
MutantD HQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGFQELLHT
MutantC HQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGFQELLHT
A1AT HQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGFQELLHT
MutantA HQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGFQELLHT

MutantB LNQPDSQLQTLTNGNLFLSEGKLKVLDFLEDVKKLYHSEAFTVNGDTEEAKKQINDYVE
MutantE LNQPDSQLQTLTNGNLFLSEGKLKVLDFLEDVKKLYHSEAFTVNGDTEEAKKQINDYVE
MutantD LNQPDSQLQTLTNGNLFLSEGKLKVLDFLEDVKKLYHSEAFTVNGDTEEAKKQINDYVE
MutantC LNQPDSQLQTLTNGNLFLSEGKLKVLDFLEDVKKLYHSEAFTVNGDTEEAKKQINDYVE
A1AT LNQPDSQLQTLTNGNLFLSEGKLKVLDFLEDVKKLYHSEAFTVNGDTEEAKKQINDYVE
MutantA LNQPDSQLQTLTNGNLFLSEGKLKVLDFLEDVKKLYHSEAFTVNGDTEEAKKQINDYVE

MutantB KGTQGKIVDLVKEKLDRTDVFALVNYIFFKGRWPEFVEVKTDEEDFHVQDDQTVTVKPMMK
MutantE KGTQGKIVDLVKEKLDRTDVFALVNYIFFKGRWPEFVEVKTDEEDFHVQDDQTVTVKPMMK
MutantD KGTQGKIVDLVKEKLDRTDVFALVNYIFFKGRWPEFVEVKTDEEDFHVQDDQTVTVKPMMK
MutantC KGTQGKIVDLVKEKLDRTDVFALVNYIFFKGRWPEFVEVKTDEEDFHVQDDQTVTVKPMMK
A1AT KGTQGKIVDLVKEKLDRTDVFALVNYIFFKGRWPEFVEVKTDEEDFHVQDDQTVTVKPMMK
MutantA KGTQGKIVDLVKEKLDRTDVFALVNYIFFKGRWPEFVEVKTDEEDFHVQDDQTVTVKPMMK

************************************************************
Appendix V

CLUSTALW Result (α1AT, Mutant A,B,C,D & E)

GenomeNet CLUSTALW Server (Kyoto Center) on Thu Oct 24 03:36:41 JST 2002

CLUSTAL W (1.81) Multiple Sequence Alignments

Sequence type explicitly set to DNA
Sequence format is Pearson
Sequence 1: α1AT 1201 bp
Sequence 2: MutantA 1201 bp
Sequence 3: MutantB 1201 bp
Sequence 4: MutantE 1201 bp
Sequence 5: MutantD 1201 bp
Sequence 6: MutantC 1201 bp
Start of Pairwise alignments
Aligning...

α1AT
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantA
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantB
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantE
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantD
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantC
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC

α1AT
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantA
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantB
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantE
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantD
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantC
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC

α1AT
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantA
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantB
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantE
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantD
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantC
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC

α1AT
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantA
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantB
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantE
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantD
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantC
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC

α1AT
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantA
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantB
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantE
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantD
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantC
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC

α1AT
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantA
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantB
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantE
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantD
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantC
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC

α1AT
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantA
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantB
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantE
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantD
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantC
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC

α1AT
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantA
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantB
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantE
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantD
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC
MutantC
ATGCATCACATCACATCACATCACATCATGATACATCCCACCATGATCAGGATCACCCAACC

115
<table>
<thead>
<tr>
<th>Mutant</th>
<th>ATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTGAATTTC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutantA</td>
<td>ATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTGAATTTC</td>
</tr>
<tr>
<td>MutantB</td>
<td>ATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTGAATTTC</td>
</tr>
<tr>
<td>MutantE</td>
<td>ATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTGAATTTC</td>
</tr>
<tr>
<td>MutantD</td>
<td>ATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTGAATTTC</td>
</tr>
<tr>
<td>MutantC</td>
<td>ATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTGAATTTC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutant</th>
<th>AACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTCCATACC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutantA</td>
<td>AACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTCCATACC</td>
</tr>
<tr>
<td>MutantB</td>
<td>AACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTCCATACC</td>
</tr>
<tr>
<td>MutantE</td>
<td>AACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTCCATACC</td>
</tr>
<tr>
<td>MutantD</td>
<td>AACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTCCATACC</td>
</tr>
<tr>
<td>MutantC</td>
<td>AACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACTCCTCCATACC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutant</th>
<th>CTCAACCAGCCAGCAAGCTCCACGTGACCCAGCGAATTGGCTGTTCCTCCACGAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutantA</td>
<td>CTCAACCAGCCAGCAAGCTCCACGTGACCCAGCGAATTGGCTGTTCCTCCACGAG</td>
</tr>
<tr>
<td>MutantB</td>
<td>CTCAACCAGCCAGCAAGCTCCACGTGACCCAGCGAATTGGCTGTTCCTCCACGAG</td>
</tr>
<tr>
<td>MutantE</td>
<td>CTCAACCAGCCAGCAAGCTCCACGTGACCCAGCGAATTGGCTGTTCCTCCACGAG</td>
</tr>
<tr>
<td>MutantD</td>
<td>CTCAACCAGCCAGCAAGCTCCACGTGACCCAGCGAATTGGCTGTTCCTCCACGAG</td>
</tr>
<tr>
<td>MutantC</td>
<td>CTCAACCAGCCAGCAAGCTCCACGTGACCCAGCGAATTGGCTGTTCCTCCACGAG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutant</th>
<th>GGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCAGAAGCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutantA</td>
<td>GGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCAGAAGCC</td>
</tr>
<tr>
<td>MutantB</td>
<td>GGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCAGAAGCC</td>
</tr>
<tr>
<td>MutantE</td>
<td>GGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCAGAAGCC</td>
</tr>
<tr>
<td>MutantD</td>
<td>GGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCAGAAGCC</td>
</tr>
<tr>
<td>MutantC</td>
<td>GGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCACTCAGAAGCC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutant</th>
<th>TTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTACGTGGAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutantA</td>
<td>TTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTACGTGGAG</td>
</tr>
<tr>
<td>MutantB</td>
<td>TTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTACGTGGAG</td>
</tr>
<tr>
<td>MutantE</td>
<td>TTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTACGTGGAG</td>
</tr>
<tr>
<td>MutantD</td>
<td>TTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTACGTGGAG</td>
</tr>
<tr>
<td>MutantC</td>
<td>TTCACTGTCAACTTCGGGGACACCGAAGAGGCCAAGAAACAGATCAACGATTACGTGGAG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutant</th>
<th>AAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutantA</td>
<td>AAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTTT</td>
</tr>
<tr>
<td>MutantB</td>
<td>AAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTTT</td>
</tr>
<tr>
<td>MutantE</td>
<td>AAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTTT</td>
</tr>
<tr>
<td>MutantD</td>
<td>AAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTTT</td>
</tr>
<tr>
<td>MutantC</td>
<td>AAGGGTACTCAAGGGAAAATTGTGGATTTGGTCAAGGAGCTTGACAGAGACACAGTTTTT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutant</th>
<th>GCTCTGGTGAATTACATCTTCTTTAAGGCAAATGGAGGAGACCCTTTGGAAGAGCAGAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MutantA</td>
<td>GCTCTGGTGAATTACATCTTCTTTAAGGCAAATGGAGGAGACCCTTTGGAAGAGCAGAC</td>
</tr>
<tr>
<td>MutantB</td>
<td>GCTCTGGTGAATTACATCTTCTTTAAGGCAAATGGAGGAGACCCTTTGGAAGAGCAGAC</td>
</tr>
<tr>
<td>MutantE</td>
<td>GCTCTGGTGAATTACATCTTCTTTAAGGCAAATGGAGGAGACCCTTTGGAAGAGCAGAC</td>
</tr>
<tr>
<td>MutantD</td>
<td>GCTCTGGTGAATTACATCTTCTTTAAGGCAAATGGAGGAGACCCTTTGGAAGAGCAGAC</td>
</tr>
<tr>
<td>MutantC</td>
<td>GCTCTGGTGAATTACATCTTCTTTAAGGCAAATGGAGGAGACCCTTTGGAAGAGCAGAC</td>
</tr>
<tr>
<td>Mutant</td>
<td>Sequence</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>α1AT</td>
<td>ACCGAGGAAGAGGACTTCCACGTGACCAGGTGACCCCGTGAAGGTGCTATGATGAAAG</td>
</tr>
<tr>
<td>Mutant A</td>
<td>ACCGAGGAAGAGGACTTCCACGTGACCAGGTGACCAGGTGCTATGATGAAAG</td>
</tr>
<tr>
<td>Mutant B</td>
<td>ACCGAGGAAGAGGACTTCCACGTGACCAGGTGACCAGGTGCTATGATGAAAG</td>
</tr>
<tr>
<td>Mutant E</td>
<td>ACCGAGGAAGAGGACTTCCACGTGACCAGGTGACCAGGTGCTATGATGAAAG</td>
</tr>
<tr>
<td>Mutant D</td>
<td>ACCGAGGAAGAGGACTTCCACGTGACCAGGTGACCAGGTGCTATGATGAAAG</td>
</tr>
<tr>
<td>Mutant C</td>
<td>ACCGAGGAAGAGGACTTCCACGTGACCAGGTGACCAGGTGCTATGATGAAAG</td>
</tr>
<tr>
<td>α1AT</td>
<td>CGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTGCTGATG</td>
</tr>
<tr>
<td>Mutant A</td>
<td>CGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTGCTGATG</td>
</tr>
<tr>
<td>Mutant B</td>
<td>CGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTGCTGATG</td>
</tr>
<tr>
<td>Mutant E</td>
<td>CGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTGCTGATG</td>
</tr>
<tr>
<td>Mutant D</td>
<td>CGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTGCTGATG</td>
</tr>
<tr>
<td>Mutant C</td>
<td>CGTTTAGGCATGTTTAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTGCTGATG</td>
</tr>
<tr>
<td>α1AT</td>
<td>AAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>Mutant A</td>
<td>AAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>Mutant B</td>
<td>AAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>Mutant E</td>
<td>AAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>Mutant D</td>
<td>AAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>Mutant C</td>
<td>AAATACCTGGGCAATGCCACCGCCATCTTCTTCCTGCCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>α1AT</td>
<td>CTGGAAAATGAACCTCACCCGGGACATTCCTTCCTGCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>Mutant A</td>
<td>CTGGAAAATGAACCTCACCCGGGACATTCCTTCCTGCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>Mutant B</td>
<td>CTGGAAAATGAACCTCACCCGGGACATTCCTTCCTGCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>Mutant E</td>
<td>CTGGAAAATGAACCTCACCCGGGACATTCCTTCCTGCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>Mutant D</td>
<td>CTGGAAAATGAACCTCACCCGGGACATTCCTTCCTGCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>Mutant C</td>
<td>CTGGAAAATGAACCTCACCCGGGACATTCCTTCCTGCTGATGAGGGGAAACTACAGCAC</td>
</tr>
<tr>
<td>α1AT</td>
<td>TCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAGAGCATC</td>
</tr>
<tr>
<td>Mutant A</td>
<td>TCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAGAGCATC</td>
</tr>
<tr>
<td>Mutant B</td>
<td>TCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAGAGCATC</td>
</tr>
<tr>
<td>Mutant E</td>
<td>TCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAGAGCATC</td>
</tr>
<tr>
<td>Mutant D</td>
<td>TCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAGAGCATC</td>
</tr>
<tr>
<td>Mutant C</td>
<td>TCTGCCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAGAGCATC</td>
</tr>
<tr>
<td>α1AT</td>
<td>CTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGGGTCACA</td>
</tr>
<tr>
<td>Mutant A</td>
<td>CTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGGGTCACA</td>
</tr>
<tr>
<td>Mutant B</td>
<td>CTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGGGTCACA</td>
</tr>
<tr>
<td>Mutant E</td>
<td>CTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGGGTCACA</td>
</tr>
<tr>
<td>Mutant D</td>
<td>CTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGGGTCACA</td>
</tr>
<tr>
<td>Mutant C</td>
<td>CTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGGGTCACA</td>
</tr>
<tr>
<td>α1AT</td>
<td>GAGGAGGCACCCCTGGAAGCTCTCCAAAGGCGGTGCAATAAGGCTGCTGACCATCAGTAG</td>
</tr>
<tr>
<td>Mutant A</td>
<td>GAGGAGGCACCCCTGGAAGCTCTCCAAAGGCGGTGCAATAAGGCTGCTGACCATCAGTAG</td>
</tr>
<tr>
<td>Mutant B</td>
<td>GAGGAGGCACCCCTGGAAGCTCTCCAAAGGCGGTGCAATAAGGCTGCTGACCATCAGTAG</td>
</tr>
<tr>
<td>Mutant E</td>
<td>GAGGAGGCACCCCTGGAAGCTCTCCAAAGGCGGTGCAATAAGGCTGCTGACCATCAGTAG</td>
</tr>
<tr>
<td>Mutant D</td>
<td>GAGGAGGCACCCCTGGAAGCTCTCCAAAGGCGGTGCAATAAGGCTGCTGACCATCAGTAG</td>
</tr>
<tr>
<td>Mutant C</td>
<td>GAGGAGGCACCCCTGGAAGCTCTCCAAAGGCGGTGCAATAAGGCTGCTGACCATCAGTAG</td>
</tr>
</tbody>
</table>
α1AT    AAAGGACTGGAAGCTGCAGGCCATGTTTTTAGGCCATACCCATGTCTATCCCCCCC
MutantA AAAGGACTGGAAGCTGCAGGCCATGTTTTTAGGCCATACCCATGTCTATCCCCCCT
MutantB AAAGGACTGGAAGCTGCAGGCCATGTTTGACGAGTCCATACCCATGTCTATCCCCCCT
MutantE AAAGGACTGGAAGCTGCAGGCCATGTTTGACGAGTCCATACCCATGTCTATCCCCCCT
MutantD AAAGGACTGGAAGCTGCAGGCCATGTTTTTAGGCCATACCCATGTCTATCCCCCCT
MutantC AAAGGACTGGAAGCTGCAGGCCATGTTTTTAGGCCATACCCATGTCTATCCCCCCT

α1AT    GAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAGTCTCCC
MutantA GAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAGTCTCCC
MutantB GAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAGTCTCCC
MutantE GAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAGTCTCCC
MutantD GAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAGTCTCCC
MutantC GAGGTCAAGTTCAACAAACCCTTTGTCTTCTTAATGATTGAACAAAATACCAAGTCTCCC

α1AT    C
MutantA C
MutantB C
MutantE C
MutantD C
MutantC C
*
Appendix VI

CLUSTALW Result (KDEL-α1AT, Mutant A & D)

GenomeNet CLUSTALW Server (Kyoto Center) on Wed Oct 23 06:13:36 JST 2002

CLUSTAL W (1.81) Multiple Sequence Alignments
Sequence type explicitly set to Protein
Sequence 1: MutantA-KDEL 429 aa
Sequence 2: MutantD-KDEL 429 aa
Sequence 3: α1AT-KDEL 429 aa
Start of Pairwise alignments
Aligning...

Flag-tag

MutantA MPSSVSWGILLAGLCLVLPSLAEDPQGDAAQYKWDADKSTTDSSHDQDHPFTENKTP
MutantD MPSSVSWGILLAGLCLVLPSLAEDPQGDAAQYKWDADKSTTDSSHDQDHPFTENKTP
α1AT MPSSVSWGILLAGLCLVLPSLAEDP0AQQYKWDADKSTTDSSHDQDHPFTENKTP

MutantA NLAEFAFSLYRQLAHQSNSTNIFFSPVSIAFAMSLGTGKADTHDEILEGLNFLNTEIP
MutantD NLAEFAFSLYRQLAHQSNSTNIFFSPVSIAFAMSLGTGKADTHDEILEGLNFLNTEIP
α1AT NLAEFAFSLYRQLAHQSNSTNIFFSPVSIAFAMSLGTGKADTHDEILEGLNFLNTEIP

MutantA EAQIHEGFQELLRTLNQPDSQLQTGNGFLSEGKLVLQKLDVEKLYHSEAFTVNF
MutantD EAQIHEGFQELLRTLNQPDSQLQTGNGFLSEGKLVLQKLDVEKLYHSEAFTVNF
α1AT EAQIHEGFQELLRTLNQPDSQLQTGNGFLSEGKLVLQKLDVEKLYHSEAFTVNF

MutantA DTEEAKKQINDYVEKGTQGKIVDLVKEDLDTVFALVYIFKGKWERPFEVKDTEEDF
MutantD DTEEAKKQINDYVEKGTQGKIVDLVKEDLDTVFALVYIFKGKWERPFEVKDTEEDF
α1AT DTEEAKKQINDYVEKGTQGKIVDLVKEDLDTVFALVYIFKGKWERPFEVKDTEEDF

MutantA HVDQVTTVKPQMRKLGMFNIQHCKKLSSWLLMKYLG3NATAIFLPEDKGQLHELEN
MutantD HVDQVTTVKPQMRKLGMFNIQHCKKLSSWLLMKYLG3NATAIFLPEDKGQLHELEN
α1AT HVDQVTTVKPQMRKLGMFNIQHCKKLSSWLLMKYLG3NATAIFLPEDKGQLHELEN

MutantA HDIITKFLHNEDDSASLHLPKLSITGTYDLKSVLQGQLGTVKFSNGADLSGVTEA
MutantD HDIITKFLHNEDDSASLHLPKLSITGTYDLKSVLQGQLGTVKFSNGADLSGVTEA
α1AT HDIITKFLHNEDDSASLHLPKLSITGTYDLKSVLQGQLGTVKFSNGADLSGVTEA

RSL (P6-P4)

KDEL-motif

MutantA VNPTQKDEL
MutantD VNPTQKDEL
α1AT VNPTQKDEL
Appendix VII

CLUSTALW Result (KDEL α1AT, Mutant A & D)

GenomeNet CLUSTALW Server (Kyoto Center) on Thu Oct 24 03:48:21 JST 2002

<table>
<thead>
<tr>
<th>Sequence 1: α1AT-KDEL</th>
<th>1298 bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence 2: MutantD-KDEL</td>
<td>1298 bp</td>
</tr>
<tr>
<td>Sequence 3: MutantA-KDEL</td>
<td>1298 bp</td>
</tr>
</tbody>
</table>

Start of Pairwise alignments

Aligning...

α1AT  ATGCCGTCTTCTCTGCTGTGGGCTCCTCACGTCTTCTGTGGCAGGCCTGTGCTGCCTGGTCCCT

MutantD  ATGCCGTCTTCTCTGCTGTGGGCTCCTCACGTCTTCTGTGGCAGGCCTGTGCTGCCTGGTCCCT

MutantA  ATGCCGTCTTCTCTGCTGTGGGCTCCTCACGTCTTCTGTGGCAGGCCTGTGCTGCCTGGTCCCT

************************************************

α1AT  GTCTCCCTGGCTGAGGATCCCCAGGGAGATGCTGCCCAGGACTACAAAGACGACGACGAC

MutantD  GTCTCCCTGGCTGAGGATCCCCAGGGAGATGCTGCCCAGGACTACAAAGACGACGACGAC

MutantA  GTCTCCCTGGCTGAGGATCCCCAGGGAGATGCTGCCCAGGACTACAAAGACGACGACGAC

************************************************************

α1AT  AAAAAGACAGATACAT CCCACCAT GAT CAG GAT CAC C CAAC C T T CAACAAGAT CAC C C C C

MutantD  AAAAAGACAGATACAT C C CAC CAT GAT CAGGAT CAC C CAAC C T T CAACAAGAT CAC C C C C

MutantA  AAAAAGACAGATACAT CCCACCAT GAT CAG GAT CAC C CAAC C T T CAACAAGAT CAC C C C C

************************************************************

α1AT  AACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAGCTGGCACACCAGTCCAACAGC

MutantD  AACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAGCTGGCACACCAGTCCAACAGC

MutantA  AACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAGCTGGCACACCAGTCCAACAGC

************************************************************

α1AT  AATATCTTCTTCTCTCCCCAGTGAGCATCGCTACAGCCTTTGCAATGCTCTCCCTGGGGACC

MutantD  AATATCTTCTTCTCTCCCCAGTGAGCATCGCTACAGCCTTTGCAATGCTCTCCCTGGGGACC

MutantA  AATATCTTCTTCTCTCCCCAGTGAGCATCGCTACAGCCTTTGCAATGCTCTCCCTGGGGACC

************************************************************

α1AT  AAGGCTGACACTCACGATGAAATCCTGGAGGGCCTGAATTTCAACCTCACGGAGATTCCG

MutantD  AAGGCTGACACTCACGATGAAATCCTGGAGGGCCTGAATTTCAACCTCACGGAGATTCCG

MutantA  AAGGCTGACACTCACGATGAAATCCTGGAGGGCCTGAATTTCAACCTCACGGAGATTCCG

************************************************************

α1AT  GAGGCTCAGATCCATCAGGAGATGAGCTACGATCCCTTGGATATCTCACCTTCTCGGGGACC

MutantD  GAGGCTCAGATCCATCAGGAGATGAGCTACGATCCCTTGGATATCTCACCTTCTCGGGGACC

MutantA  GAGGCTCAGATCCATCAGGAGATGAGCTACGATCCCTTGGATATCTCACCTTCTCGGGGACC

************************************************************

α1AT  AAGGTGACACTACAGGAGGAGAGAGGACTACGATCCCTTGGATATCTCACCTTCTCGGGGACC

MutantD  AAGGTGACACTACAGGAGGAGAGAGGACTACGATCCCTTGGATATCTCACCTTCTCGGGGACC

MutantA  AAGGTGACACTACAGGAGGAGAGAGGACTACGATCCCTTGGATATCTCACCTTCTCGGGGACC

************************************************************

120
**α1AT**

**Mutant D**

**Mutant A**

---

**α1AT**

**Mutant D**

**Mutant A**

---

**α1AT**

**Mutant D**

**Mutant A**

---

**α1AT**

**Mutant D**

**Mutant A**

---

**α1AT**

**Mutant D**

**Mutant A**

---

**α1AT**

**Mutant D**

**Mutant A**

---

**α1AT**

**Mutant D**

**Mutant A**

---

**α1AT**

**Mutant D**

**Mutant A**

---

**α1AT**

**Mutant D**

**Mutant A**

---
CTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATCGACGAGAAAGGGACTGAAGCTGCT
CTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATCGACGAGAAAGGGACTGAAGCTGCT
CTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATCGACGAGAAAGGGACTGAAGCTGCT
GGGGCCATGTTTTTAGAGGCCATACCCATGTCTATCCCCCCCGAGGTCAAGTTCAACAAA
GGGGCCATGTTTTTAGAGTCCATACCCTGCTCTATCCCCCCTGAGGTCAAGTTCAACAAA
GGGGCCATGTTTTTAGAGGCCATACCCATGTCTATCCCCCCTGAGGTCAAGTTCAACAAA
CCCTTTGTCTTCTTTAATGATTGAACAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTG
CCCTTTGTCTTCTTTAATGATTGAACAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTG
CCCTTTGTCTTCTTTAATGATTGAACAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTG
GTGAATCCCACCCAAAAAGACGAGCTCTGAAGCTTCTA
GTGAATCCCACCCAAAAAGACGAGCTCTGAAGCTTCTA
GTGAATCCCACCCAAAAAGACGAGCTCTGAAGCTTCTA
CCCTTTGTCTTCTTTAATGATTGAACAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTG
CCCTTTGTCTTCTTTAATGATTGAACAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTG
CCCTTTGTCTTCTTTAATGATTGAACAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTG
CTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATCGACGAGAAAGGGACTGAAGCTGCT
CTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATCGACGAGAAAGGGACTGAAGCTGCT
CTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATCGACGAGAAAGGGACTGAAGCTGCT
GGGGCCATGTTTTTAGAGGCCATACCCATGTCTATCCCCCCCGAGGTCAAGTTCAACAAA
GGGGCCATGTTTTTAGAGTCCATACCCTGCTCTATCCCCCCTGAGGTCAAGTTCAACAAA
GGGGCCATGTTTTTAGAGGCCATACCCATGTCTATCCCCCCTGAGGTCAAGTTCAACAAA
CCCTTTGTCTTCTTTAATGATTGAACAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTG
CCCTTTGTCTTCTTTAATGATTGAACAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTG
CCCTTTGTCTTCTTTAATGATTGAACAAAATACCAAGTCTCCCCTCTTCATGGGAAAAGTG
GTGAATCCCACCCAAAAAGACGAGCTCTGAAGCTTCTA
GTGAATCCCACCCAAAAAGACGAGCTCTGAAGCTTCTA
GTGAATCCCACCCAAAAAGACGAGCTCTGAAGCTTCTA
**************************************************
**************************************************
**************************************************
122