Chicken egg white lysozyme variant I55T: 
Raman spectroscopic analysis and expression of random mutants.

by

SAEKO KANAGAWA

B.H.E., Ochanomizu University, Tokyo, JAPAN, 1996

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

MASTER OF SCIENCE

in

The Faculty of Graduate Studies

(Department of Food Science)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

March 2002

© Saeko Kanagawa, 2002
In presenting this thesis in partial fulfillment 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 Food Sciences

The University of British Columbia
Vancouver, Canada

Date March 19 102
ABSTRACT

Chicken egg white lysozyme is often used in protein studies because of its well-known properties. I55T variant chicken egg white lysozyme is an analogue of I56T amyloidogenic human lysozyme which causes hereditary non-neuropathic amyloidosis. The mechanism of alteration of physical properties by a single substitution is not well-understood.

Wild type and I55T lysozymes expressed in *Pichia pastoris* were compared using Raman spectroscopy to elucidate the conformational alteration caused by the single amino acid substitution. Wild type showed a similar spectrum to commercial lysozyme from chicken egg white; however, some structural changes were indicated. The I55T variant did not exhibit typical tryptophan signals, suggesting the exposure of tryptophan residue(s) to an aqueous environment. This Raman analysis suggests a structural change in I55T variant, which may be related to the aggregation property of I55T variant. It was suggested that *P. pastoris* expression system may affect protein folding.

Mutations were introduced to the I55T variant using Random Centroid Optimization for Genetics (RCG) to attempt to decrease the amyloidogenic nature of I55T variant. Nine sets of two site mutations were introduced to the I55T template and these nine variants were expressed in *P. pastoris* along with wild type and I55T lysozymes.

Two variants, K33V-C115F and P79I-G126K, were not secreted. The lack of secretion may indicate the protein structure was abnormal, which resulted in degradation of the protein by the yeast. Seven mutants exhibited comparable lysozyme expression levels to wild type but six displayed no measurable enzymatic activity. This may suggest the aggregation of the secreted protein and/or a conformation change around active sites which resulted in the loss of functionality. Effects of mutations were investigated using the Homology Similarity Analysis program; moderate changes in all propensities were found for T40E-A42K, the only mutant which exhibited enzymatic activity, and higher charge introduction in P79I-G126K.
Aggregation property may have been enhanced in K33V-C115F and P79I-G126K due to the increased β-strand propensity. Results indicate that the local conformational characteristics relative to the whole protein structure may be more important than the local sequence propensity. Data indicate the difficulty in altering the amyloidogenic property of I55T lysozyme.
# TABLE OF CONTENTS

ABSTRACT .................................................................................................................. ii  

TABLE OF CONTENTS ............................................................................................... iv  

LIST OF TABLES ......................................................................................................... vii  

LIST OF FIGURES ....................................................................................................... viii  

LIST OF APPENDICES ................................................................................................. xiii  

ACKNOWLEDGEMENT ................................................................................................. xiv  

## CHAPTER I  INTRODUCTION

1.1 Outline and objectives ......................................................................................... 1  

1.2 Literature Search

   1.2.1 Lysozyme .................................................................................................. 4  

   1.2.2 Amyloidogenic I55T chicken egg white lysozyme variant .................. 10  

   1.2.3 *Pichia pastoris* ......................................................................................... 13  

   1.2.4 Raman spectroscopy ............................................................................... 19  

   1.2.5 Random Centroid Optimization for Genetics (RCG) program .......... 22  

   1.2.6 Homology Similarity Analysis (HSA) ....................................................... 24  

## CHAPTER II  GENETIC MODIFICATION OF CHICKEN EGG WHITE LYSOZYME

2.1 Introduction .......................................................................................................... 26  

2.2 Materials and Methods

   2.2.1 Mutations by Random Centroid Optimization for Genetics (RCG) .... 26  

   2.2.2 Materials, bacterial strains and plasmids ....................................... 27  

   2.2.3 Primers for site directed mutagenesis .............................................. 31  

   2.2.4 Site directed mutagenesis ................................................................. 33
2.2.5 Subcloning

2.2.5.1 Insert DNA preparation 35
2.2.5.2 Vector DNA preparation 36
2.2.5.3 Ligation 36
2.2.5.4 Selection of transformants 38

2.2.6 Transformation of gene into P. pastoris

2.2.6.1 Direct PCR screening method 41
2.2.6.2 Sequence of integrated gene in Pichia genome 41

2.3 Results and Discussion 43

CHAPTER III EXPRESSION OF WILD TYPE AND I55T LYSOZYMES IN PICHIA PASTORIS AND RAMAN SPECTROSCOPIC ANALYSIS

3.1 Introduction 47

3.2 Materials and Methods

3.2.1 Expression media 47
3.2.2 Protein separation and purification 48
3.2.3 Protein detection 48
3.2.4 Lysozyme enzymatic activity assay 49
3.2.5 Small-scale expression of wild type and I55T lysozymes 50
3.2.6 Small-scale expression of wild type and I55T lysozyme with Vectrase-P™ 50
3.2.7 Large-scale expression of wild type lysozyme and purification 52
3.2.8 Large-scale expression of I55T lysozyme and purification 53
3.2.9 Raman spectroscopy 53
3.3 Results and Discussion

3.3.1 Enzymatic activity of wild type and I55T lysozymes 54
3.3.2 Expression of wild type and I55T lysozyme with Vectrase-P™ 55
3.3.3 Expression and purification of wild type lysozyme 58
3.3.4 Expression and purification of I55T lysozyme 63
3.3.5 Raman spectroscopic analysis 70

CHAPTER IV EXPRESSION OF NINE LYSOZYME MUTANTS SUGGESTED BY RCG

4.1 Introduction 77
4.2 Materials and Methods
   4.2.1 Protein expression and detection 77
   4.2.2 Homology Similarity Analysis (HSA) program 78
4.3 Results and Discussion 78
   4.3.1 Homology Similarity Analysis (HSA) 85
   4.3.2 Summary 99

CHAPTER V CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK 102

REFERENCES 107

APPENDICES 114
LIST OF TABLES

Table 1  Sequence and $T_m$ of primers used for site directed mutagenesis. 32

Table 2  Tentative assignment of some major peaks of Raman spectrum of chicken egg white lysozyme (Li-Chan and Nakai, 1991). 73

Table 3  The values of helix, strand, and turn propensities, hydrophobicity, hydrogen bond forming ability, charge and bulkiness used in the HSA program for each amino acid. 80

Table 4  Summary of mutation site, original amino acid residue, amino acid residues observed among all animal lysozymes c, substitution amino acid residue, expression level, and enzymatic activity level of nine mutants. 84

Table 5  Differences in average propensity values of sequence segments flanking site 55 in I55 chicken egg white lysozyme mutants comparing to wild type (wild type-mutant). (dTm is a difference in the melting temperature from that of I55T) 87

Table 6  Similarity constants of sequence segments flanking site 55 in I55 chicken egg white lysozyme mutants with wild type as a reference. (dTm is a difference in the melting temperature from that of I55T) 88

Table 7  Differences in average propensity values of sequence segments flanking site 56 in I56 human lysozyme mutants comparing to wild type (wild type-mutant). (dTm is a difference in the melting temperature from that of I55T) 90

Table 8  Similarity constants of sequence segments flanking site 56 in I56 human lysozyme mutants with wild type as a reference. (dTm is a difference in the melting temperature from that of I55T) 91

Table 9  Differences in average propensity values of 5 residue sequence segments flanking each mutation in 9 double mutants comparing to wild type (wild type-mutant). (Bold letters indicate the values which exhibit the largest and second largest variations in each factor) 95

Table 10  Similarity constants of 5 residue sequence segments flanking each mutation in 9 double mutants with wild type as a reference. (Bold letters indicate the values which exhibit the smallest and second smallest similarity constants in each factor) 96
LIST OF FIGURES

Figure 1  Schematic views of the native structure of chicken egg white lysozyme. (A) a lysozyme molecule: helix A (4-15), helix B (24-36), helix C (88-99), helix D (108-115) and 310 helix (120-125) form the α-domain; triple-stranded antiparallel β-sheet (41-60), small β-sheet (1-2 and 39-40), single turn 310 helix (79-84) and the large loop (61-78) for the β-domain (Protein Explorer 1HEL); (B) a lysozyme molecule co-crystallized with tri-N-acetyl-chitotriose at pH 4.7 shown with catalytic residues, Glu-35 and Asp-52, and Ile-55 (Protein Explorer 1LZB).

Figure 2  Reaction mechanism of lysozyme (Imoto, 1996). Glu-35, which is located in a hydrophobic environment, participates in catalysis in protonated form and Asp-52, which is located in a hydrophilic environment, in dissociated form. Glu-35 donates its proton to the oxygen linking DE sugars. Then the bond between this oxygen and C1 on the D sugar is cleaved. A carbonium ion is formed on the D sugar and this ion is stabilized by the formation of an oxocarbonium ion. The negative charge on Asp-52 stabilizes the formation of a positive charge on the D sugar.

Figure 3  Gene insertion at AOX1 on the genome of Pichia. Gene insertion events at the AOX1 loci arise from a single crossover event between the loci and either of the two AOX1 regions on the pPICZ or pPICZA vectors: the AOX1 promoter or the AOX1 transcription termination region (TT). This event results in the insertion of one or more copies of the vector upstream or downstream of the AOX1 gene. This figure shows the result of an insertion of the plasmid 5’ to the intact AOX1 locus (Mut+) and the gain of PAOX1, a gene of interest, and the Zeocine™ resistance gene (Invitrogen, EasySelect™ Pichia Expression kit manual version A).

Figure 4  Construction of pPICZ-B plasmid vector. 5’ AOX1 indicates a 942 bp fragment containing the AOX1 promoter and this region targets plasmid integration to the AOX1 locus. myc epitope indicates C-terminal myc epitope tag (Glu-Gln-Lys-Leu-ILe-Ser-Glu-Glu-Asp-Leu-Asn) which permits detection of the fusion protein by the Anti-myc Antibody or Anti-myc-HRP Antibody. (His)6 indicates C-terminal polyhistidine tag which permits purification of the fusion protein on metal-chelating resin. AOX1 TT is AOX1 transcription termination which contains native transcription termination and polyadenylation signal from AOX1 gene that permits efficient 3’ mRNA processing, including polyadenylation for increased mRNA stability. PTEF1 indicates TEF1 promoter which is a transcription elongation factor 1 gene promoter from Saccharomyces cerevisiae that drives expression of the Sh ble gene in Pichia, conferring Zeocine™ resistance. PEM7 indicates EM7 promoter which is a constitutive promoter that drives expression of the Sh ble gene in E. coli, conferring Zeocine™ resistance. Zeocine indicates Sh ble (Streptocolloteichus hindustanus ble) gene for Zeocine™ resistance. cyc1 TT (transcription termination) indicates 3’ end of the S. cerevisiae CYC1 gene that allows efficient 3’ mRNA processing of the Sh ble gene for increased stability. ColEl indicates pUC19 origin which allows replication and maintenance of the plasmid in E. coli. XbaI in bold
letter in the multiple cloning site indicates the restriction enzyme site used for insertion of the fragment containing lysozyme gene.

**Figure 5** Construction of pPICZa-C plasmid vector. α-factor indicates native *S. cerevisiae* α-factor secretion signal which allows for efficient secretion of most proteins from *Pichia*. Other indicators are as described in Figure 4. XhoI in bold letter in the multiple cloning site indicates the restriction enzyme site used for insertion of the fragment containing lysozyme gene.

**Figure 6** Vibrational spectroscopic techniques and molecular energy levels. Infrared (IR) spectroscopy utilizes the absorption of incident electromagnetic radiation at a particular frequency ($v_i$), while Raman spectroscopy utilizes the inelastic scattering of the incident radiation. In Raman scattering, Stokes transitions occur as a result of excitation of molecule ($+v_i$) and anti-Stokes occurs as a result of de-excitation of molecule ($-v_i$). (Li-Chan, 1996)

**Figure 7** Construction of chicken egg white lysozyme cDNA in pUC19 vector. Chicken egg white lysozyme cDNA (387 base pairs) with stop codon (TGA) at the C-terminus was inserted into pUC19 multi-cloning site at *PstI* and *SalI*.

**Figure 8** A sequence map of pUC19 with restriction enzyme cleavage sites. Polylinker region indicates the multi-cloning site. Ap indicates β-lactamase gene conferring resistance toward ampicillin. ORI indicates an origin of replication gene from CoIE1 plasmid of *E.coli*. lacZ indicates a gene encoding β-galactosidase. (adopted from www.biology.ucsc.edu/classes/bio20U/content/molbio2/puc19.htm)

**Figure 9** Nucleotide sequence and the corresponding amino acids of chicken egg white lysozyme. Bold letters indicate the corresponding amino acid residues to the codon above. Ile at position 55 is shown underlined.

**Figure 10** Construction of insert fragments for pPICZ-B and pPICZa-C vectors. Insert DNA fragment was amplified between pPB and M4 primers for ligation into pPICZ-B vector DNA (A) and pPaC and M4 primers for ligation into pPICZa-C vector DNA (B). pPB and pPaC primers introduced restriction enzyme sites *EcoRI* and *XhoI* to the inserts, respectively. Fragments were digested with *EcoRI* or *XhoI* and *XbaI* and inserted into pPICZ-B/EX or pPICZa-C/XX vector DNA.

**Figure 11** Vector DNA with or without insert DNA after subcloning observed on 0.7% agarose gel. Insert DNA successfully inserted into vector DNA was observed in #1,2, and 4, while #3 did not have insert DNA.
Figure 12  Vector DNA with or without insert DNA after restriction enzyme digestion observed on 0.7% agarose gel. Insert DNA fragment cut out of vector and the vector fragment were observed in #1,2,4,5 and 6, while #3 shows a band longer than the vector fragment and no insert fragment.

Figure 13  Fragments amplified in Direct PCR screening method observed on 0.7% agarose gel. Wild and I55T clones were the positive controls which show the fragment of 1,555 base pairs and control clone was the negative controls with 1168 base pairs. Two of nine mutants were shown to have the same length fragment as wild and I55T clones. All the other mutants were also applied to Direct PCR screening method and showed the same positive fragment. Lane 1: wild clone (positive control); lane 2: I55T clone (positive control); lane 3: control clone (negative control); lane 4: mutant 1; lane 5: mutant 2; lane 6: λ/EH marker.

Figure 14  Construction of the amplified fragment from the gene integrated into Pichia genome and sites for primers used in sequencing. A gene (pPICZa-C carrying I55T lysozyme with A31F-N44I mutations) integrated into the P. pastoris genome was amplified using 5' and 3'-AOX primers. The lysozyme gene was sequenced using N44I-A or N44I-B primers.

Figure 15  Enzymatic activity/gram wet-cell weight of wild and I55T lysozyme in four day fermentation. Wild and I55T lysozymes were expressed extracellularly in P. pastoris (four days). Enzymatic activities were checked daily with Micrococcus lysodeiktus and specific activity was calculated by dividing the decrease of absorbance (dA) by wet cell weight (g). •: wild type lysozyme; •: I55T lysozyme. The linear regression r² (----) and the p values for wild type are 0.9205 and 0.041, and those for I55T lysozyme are 0.979 and 0.011, respectively.

Figure 16  Time course change of enzymatic activity/gram wet-cell weight of wild type (A) and I55T variant (B) lysozymes with various concentrations of Vectrase-P™ during a four day fermentation. • blank (no additives), • control (2 µl DMSO without Vectrase-P™), ▲ 0.05 mg/ml, × 0.1 mg/ml, * 0.2 mg/ml, ● 0.4 mg/ml.

Figure 17  Elution profile of wild type lysozyme expressed in P. pastoris from CM-Toyopearl ion-exchange chromatography. Wild type lysozyme extracellularly expressed in P. pastoris was harvested, diluted with 2-3 volumes of deionized distilled water and applied to CM-Toyopearl equilibrated with 50 mM potassium phosphate buffer (pH 6.2). Lysozyme was eluted using a linear gradient of 0-1 M NaCl at 1.0 ml/ min. The dotted line indicates the NaCl gradient.

Figure 18  SDS-PAGE (left) and Western blotting (right) of fractions of wild type lysozyme separated by CM-Toyopearl shown in Figure 18. Fractions 44 (peak A), 47 and 49 (peak B) from CM-Toyopearl were applied to SDS-PAGE (left); fraction 49 showed
one band at lysozyme (ca. 14.5 kDa) and fractions 44 and 47 showed lysozyme and a higher molecular weight band (ca. 16 kDa). Both bands were recognized by anti-chicken egg white lysozyme antibody during Western blotting. Lane 1: commercial lysozyme from chicken egg white; lane 2: fraction 49, lane 3: fraction 44; lane 4: fraction 47; lane 5: Low molecular weight marker; lane 6: fraction 43; lane 7: fraction 47.

Figure 19  Elution profile of peak fractions purified using hydroxyapatite chromatography. Fractions 46-54 from CM-Toyopearl of wild type lysozyme (Figure 17) were applied to hydroxyapatite chromatography equilibrated with 10 mM sodium phosphate buffer (pH 7.2). Lysozyme was eluted using a linear gradient of 10-400 mM sodium phosphate buffer (pH 6.8) at 1.0 ml/min. Fractions were plotted as the absorbance at 280 nm vs. fraction number. The dotted line indicates the sodium phosphate gradient.

Figure 20  SDS-PAGE image of fractions from hydroxyapatite chromatography fractions of wild type lysozyme. Fractions from wild type lysozyme purified with hydroxyapatite chromatography were applied to SDS-PAGE. Peak A fractions contained a lysozyme band (ca. 14.5 kDa) and peak B fractions contained higher molecular weight band (ca. 16 kDa). Peak A fractions 10-17 were collected, dialyzed and lyophilized for Raman analysis. Lane 1: native lysozyme from chicken egg white; lane 2: fraction 13; lane 3: fraction 16; lane 4: fraction 17; lane 5: low molecular weight marker; lane 6: fraction 18; lane 7: fraction 19; lane 8: fraction 20.

Figure 21  Elution profile of I55T lysozyme expressed in P. pastoris from CM-Toyopearl chromatography. I55T lysozyme intracellularly expressed in P. pastoris was harvested, lysed with glass beads, centrifuged, diluted with 2-3 volumes of the equilibration buffer and applied to CM-Toyopearl equilibrated with 50 mM potassium phosphate buffer (pH 6.2). Lysozyme was eluted using a linear gradient from 0 to 1 M NaCl at 1.2 ml/min. The dotted line indicates the NaCl gradient.

Figure 22  SDS-PAGE (left) and Western blotting (right) of fractions from CM Toyopearl chromatography of I55T lysozyme. CM-Toyopearl I55T lysozyme fractions 23, 26, 30, 34, and 40 were applied to SDS-PAGE and fractions 23, 26, 30 and 34 were applied to Western blotting. Fractions 23, 26 and 30 showed band at ca. 40 kDa reacting with the anti-lysozyme antibody. Lane 1: fraction 40; lane 2: fraction 36; lane 3: fraction 34; lane 4: fraction low molecular weight marker; lane 5: fraction 30; lane 6: fraction 26; lane 7: fraction 23; lane 8: commercial lysozyme from chicken egg white; lane 9: fraction 23; lane 10: fraction 26; lane 11: fraction 30; lane 12: fraction 34. (Note: SDS-PAGE gel and the membrane are not presented on the same scale.)

Figure 23  Second load fractions of CM-Toyopearl chromatography of I55T lysozyme. Fractions 32-40 (peak B) in Figure 22 were applied to second CM-Toyopearl chromatography (pH 6.2). Lysozyme was eluted using a linear gradient from 0 to 1
M NaCl at 0.3 ml/min. A single peak was observed and fractions were applied to SDS-PAGE. The dotted line indicates the NaCl gradient.

**Figure 24** SDS-PAGE of fractions from I55T lysozyme CM-Toyopearl second load fractions. Fractions 25, 26, 27, 28, 29, 30 and 31 from Figure 23 were applied to SDS-PAGE. Early fractions 25-27 were found to contain less lysozyme than late fractions 29-31. Lane 1: fraction 31; lane 2: fraction 30; lane 3: fraction 29; lane 4: fraction 28; lane 5: low molecular weight marker; lane 6: fraction 27; lane 7: fraction 26; lane 8: fraction 25; lane 9: commercial lysozyme from chicken egg white.

**Figure 25** Raman spectra of native, wild and I55T lysozyme. Commercial lysozyme from chicken egg white, wild type and I55T lysozymes expressed in *P. pastoris* were applied to Raman spectroscopy at the wavelength range of 400 cm\(^{-1}\) and 1700 cm\(^{-1}\).

**Figure 26** Western blotting and dot blotting images of wild, I55T and nine variant lysozymes with control. 1: A31F-N44I, 2: K33V-C115F, 3: C6L-Y23C, 4: A95S-G71F, 5: G126K-P79I, 6: I124E-G67I, 7: K116P-T40P, 8: P70F-R61H, and 9: T40E-A42K. Control was from *Pichia* control clone fermentation which does not synthesize lysozyme due to the lack of lysozyme gene. No lysozyme expression was observed in two and five from three sets of fermentations, while seven other mutants, I55T variant, wild type, and native lysozyme from chicken egg white show the reactivity with anti-lysozyme antibody.

**Figure 27** Average propensity values of chicken egg white lysozyme wild type and I55 mutants (Shih et al., 1995) in the region flanking site 55; the bar shows the melting temperature relative to I55T (dTm), other factors were indicated by different marks shown in the legend.

**Figure 28** Average propensity values of human lysozyme wild type and I56 mutants (Funahashi et al., 1999) in the region flanking site 56; the bar shows the melting temperature relative to I56T variant (dTm), other factors were indicated by different marks shown in the legend.
LIST OF APPENDICES

Appendix 1  Scheme of PCR primer checking method.  114

Appendix 2  Scheme of “Long PCR method” for site directed mutagenesis.  115

Appendix 3  E. coli competent cell preparation by rubidium chloride method.  116

Appendix 4  Transformation of gene into E. coli competent cells.  117

Appendix 5  DNA mini-prep method for plasmid DNA extraction from E. coli.  118

Appendix 6  PEG precipitation method for purification of plasmid DNA.  119

Appendix 7  Electroporation method for E. coli transformation.  120

Appendix 8  Genomic DNA extraction from yeast.  121
ACKNOWLEDGEMENTS

I would like to thank my research supervisor, Dr. Christine H. Scaman, whose support and encouragement helped me to achieve this great accomplishment in my life.

I also would like to thank Dr. Shuryo Nakai who gave valuable advices to guide me to this big achievement and Dr. Masahiro Ogawa who taught me all the basics and techniques. I am grateful for all the help from my committee members, Dr. Eunice Li-Chan and Dr. Hennie van Vuuren.

I especially thank my family, Stephanie Foudrinier and Gary Raisanen for their support, encouragement, and care.

I would like to thank all the people helped me and encouraged me inside and outside of Food Science.

Saeko Kanagawa
CHAPTER I INTRODUCTION

1.1 Outline and Objectives

The relationship between protein primary structure and functionality has been extensively investigated using various proteins as model systems. Today it is still impossible to predict the tertiary structure and function of a protein from the amino acid sequence; estimation and prediction of unknown protein functions from the known primary structure is one of the biggest hurdles in current proteomics. While the rapid improvement in DNA sequencing technology which resulted in automated sequencing has been observed in last three decades, the technique to analyze the relationship between the sequence and protein functionality has not been established.

Synthesized polypeptides are correctly folded in the endoplasmic reticulum by the action of molecular chaperons to become fully-functional proteins. Various factors are believed to be involved in protein folding; such as amino acid propensities (tendency to form helix, strand, or turn structure in a local conformation) and properties (hydrophobicity, bulkiness, charge), and non-covalent interactions such as van der Waals contacts, hydrogen bonding, ionic interactions and association of non-polar residues. It remains unknown which of these factors are most important in protein folding or how these factors interact with each other; therefore, it is still impossible to predict protein structure from amino acid sequence.

There are several amyloidogenic proteins found to cause serious diseases in human, such as Alzheimer’s disease and Creutzfeldt-Jakob disease. These amyloidogenic proteins form aggregates and deposit in various parts of the body. The mechanism of aggregation has not been elucidated yet. It is widely believed the aggregation of \( \beta \)-sheet structure from a globular molecule is the main mechanism of amyloid formation. An example of the complexities of protein folding has been revealed by the discovery of amyloidogenic lysozyme causing non-neuropathic systemic amyloidosis (Pepys et al., 1993). The single amino acid substitution in
human lysozyme of I56T or D67H results in the deposit of protein aggregates in the viscera, especially the liver, and the death of the patients by the fifth decade. This was a significant discovery in terms of both the medical implications and for proteomics. Naturally occurring single amino acid mutations can dramatically alter normal lysozyme into an amyloidogenic protein; therefore, investigating these substitutions may lead to an understanding of the significance of a certain amino acid residue to the whole protein structure. It was found that chicken egg white lysozyme mutants I55T and D66H variants (analogue of I56T and D67H human lysozyme mutants, respectively) are also amyloidogenic (Shih et al., 1995; Song et al., 2001). It was unpredictable that a single substitution in 129 amino acids (chicken egg white) or 130 amino acids (human) would drastically alter the properties of the protein. Therefore, it became very important to obtain information about the effects of these substitutions on protein conformation. There have been several studies focused on these lysozyme mutants; however, the mechanism of amyloid formation, and amino acid substitution effects on protein conformation, are still not fully understood.

Amyloidogenic chicken egg white lysozyme I55T variant was used in two studies. In the first study, chicken egg white lysozyme (wild type) and the I55T variant were expressed in Pichia pastoris, purified, and subjected to Raman spectroscopy to investigate structural differences between wild type and the I55T variant. This analysis should provide some information about the amyloid formation mechanism by the structural comparison between a normal protein (wild type) and an abnormal protein (I55T variant) using Raman spectroscopic analysis.

In the second part of this study, the Random Centroid optimization for Genetics (RCG) program was utilized to introduce nine sets of two-site random mutations into the lysozyme I55T variant with the aim of reducing the aggregation of the protein. RCG is a regulated random process which requires minimum information about the template to be optimized. Site directed
mutagenesis is very effective to investigate the protein structure-function relationship; however, the selection of mutation site and amino acid residue has to be based on the background information on the protein structure for efficient mutagenesis, otherwise enormous mutagenesis iterations, which are time-consuming, labor intensive, and expensive, are required. Therefore, application of RCG for the introduction of mutations into proteins is advantageous and could be very effective. These nine protein variants were expressed in *P. pastoris* and characterized in comparison with wild type and I55T lysozymes. It was expected that simultaneous introduction of two mutations may result in a more rapid alteration of the protein properties. Homology Similarity Analysis (HSA) program was used to analyze the effects of mutations based on the amino acid residue properties.

*P. pastoris* was used as the expression system for amyloidogenic proteins. In many studies, *Saccharomyces cerevisiae* is used as an expression host. However, the methylotrophic yeast *P. pastoris* has been reported to have higher foreign protein productivity (Cereghino *et al.*, 2000). This organism is reported to yield 5-100 times higher productivity of wild type chicken egg white lysozyme over *S. cerevisiae* (Mine *et al.*, 1999); however, up to now *P. pastoris* has not been extensively used in amyloidogenic protein studies. Therefore, the use of *P. pastoris* as an expression system for amyloidogenic lysozyme expression was examined.
1.2 Literature search

1.2.1 Lysozyme

Lysozyme (EC 3.2.1.17) is the ubiquitous enzyme which has been found in various biological fluids and tissues, including avian egg, plant and animal secretions, tears, saliva, and respiratory and cervical secretions (Jollés, 1996). It is reported that lysozyme has various biological activities such as bacteriolytic activity, antiviral activity, anti-inflammatory properties, anti-lipopolysaccharide endotoxic activity (Ito et al., 1997a), activity as an immunostimulant (Siwicki et al., 1998) and anti-metastatic activity (Pacor et al., 1996).

Chicken egg white lysozyme belongs to the c type (chicken- or conventional-type) lysozymes (Jollés, 1996; Prager & Jollés, 1996a). There are some other lysozyme families (g-type, v-type, λ-type, h-type, and b-type families) with the unique functions and structures. Human lysozyme differs from chicken egg white lysozyme by one additional amino acid Ala at 47 and the homology between these c-type lysozymes is about 60.8% (79 amino acids out of 130 residues). Chicken egg white lysozyme is a relatively low molecular weight protein (~14,300 Da) with a high isoelectric point (pI ≈ 10.7). Lysozyme has a higher extinction coefficient $E_{1\%}^{1\%}$ =26.4 at 280 nm than most proteins due to the high aromatic amino acid content. Lysozyme consists of 129 amino acid residues with four disulfide bonds between 6-127, 30-115, 64-80, and 76-94 residues. The molecule consists of two domains (domain A and B) linked by a long $\alpha$-helix and the active site lies in the cleft between these domains. Domain A is a hydrophobic box consisting of residues 1-39 and 85-129 surrounding tryptophan residues at 28, 108 and 111. X-ray structure analysis shows 4 $\alpha$-helices, 5 residues and one $3_{10}$ helix turn; helix A (4-15), B (24-36), C (88-99), D (108-115) and $3_{10}$ helix (120-125). Domain B is a flexible hydrophilic region consisting of a triple-stranded antiparallel $\beta$-sheet (41-60), small $\beta$-sheet (1-2 and 39-40), single turn $3_{10}$ helix (79-84) and the large coiled loop (61-78) (Wilson et al., 1992) (Figure 1-A).
Figure 1  Schematic views of the native structure of chicken egg white lysozyme. (A) a lysozyme molecule: helix A (4-15), helix B (24-36), helix C (88-99), helix D (108-115) and 310 helix (120-125) form the α-domain; triple-stranded antiparallel β-sheet (41-60), small β-sheet (1-2 and 39-40), single turn 310 helix (79-84) and the large loop (61-78) for the β-domain (Protein Explorer 1HEL); (B) a lysozyme molecule co-crystallized with tri-N-acetyl-chitotriose at pH 4.7 shown with catalytic residues, Glu-35 and Asp-52, and Ile-55 (Protein Explorer 1LZB).
Lysozymes have enzymatic activity which lyse the cell walls of Gram-positive bacteria by hydrolyzing the β 1, 4-glycosidic linkage between N-acetylmuramic acid (NAM) and N-acetyl-glucosamine (NAG) polymers (NAM-NAG)_n in the peptidoglycan, yielding polymers of N-acetyl-D-glucosamine (chitin) and N-acetylmuramic acid or homopolymers of N-acetyl-D-glucosamine. The mechanism of hydrolysis of β 1, 4-glycosidic linkage is described in Figure 2 (Imoto, 1996). The active site of chicken egg white lysozyme consists of six subsites A, B, C, D, E, and F and catalytic residues Glu-35 and Asp-52 (Fig 1-B). Subsite A consists of Asn-101 and Asn-103 and subsite B is Asn-103, Trp-62 and Trp-63. Subsite C which provides most of the binding energy is formed by the main chain of Asn-59, the main chain carbonyl of Ala-107, the deep hydrophobic hole formed by Ile-58 and Ile-98, and the indole ring of Trp-108. Subsite D includes the main and side chain of Val-109 and the hydrogen-bonded ‘platform’ of Asn-46, Ser-50, Asp-52 and Asn-59. Subsites E and F presumably consist of Asn-44, Gln-57 and Arg-114. While the six subsites bind six sugar residues in a substrate, Glu-35 in protonated form and Asp-52 in dissociated form cleave the bond between the oxygen and C1 on the D-sugar (Imoto, 1996; Ibrahim, 1997).

In addition to the enzymatic activity against Gram-positive bacteria which is associated with the conformation of lysozymes described above, lysozymes possess antimicrobial activity which is possibly associated with the positive charge of lysozymes and not with its lytic activity (Cisani et al., 1984). It also has been reported that heat denatured chicken egg white lysozyme, with no enzymatic activity, showed enhanced antimicrobial activity against Gram-negative bacteria with no detrimental effect on its inherent action to Gram-positive bacteria (Ibrahim et al., 1996). The greatly enhanced surface hydrophobicity and exposure of tryptophan residues were observed in fluorescence spectra. The heat denatured lysozyme mainly consists of β-sheets and exists as a dimer with some high molecular weight oligomer and the action against Gram-negative bacteria has been possibly attributed to either the permeabilization of membranes.
Figure 2 Reaction mechanism of lysozyme (Imoto, 1996). Glu-35, which is located in a hydrophobic environment, participates in catalysis in protonated form and Asp-52, which is located in a hydrophilic environment, in dissociated form. Glu-35 donates its proton to the oxygen linking DE sugars. Then the bond between this oxygen and C1 on the D sugar is cleaved. A carbonium ion is formed on the D sugar and this ion is stabilized by the formation of an oxocarbonium ion. The negative charge on Asp-52 stabilizes the formation of a positive charge on the D sugar.
resulting from the amphiphilicity of the dimer and/or the stimulation of cell autolysis. A recent study using D52S variant indicated that the antimicrobial activity of lysozyme is due to structural factors which are independent from the enzymatic action (Ibrahim et al., 2001).

Both genetic and chemical modifications have been applied to lysozymes in order to understand the function of the enzyme including the binding and catalytic actions and to improve the properties of lysozymes such as thermal stability, enzymatic activity and antimicrobial activity. Site directed mutagenesis to substitute a specific amino acid at a specific position is a valuable technique used to better understand the structure-function relationship of proteins. In lysozyme research, a variety of substitutions have been applied to the amino acids that are involved in substrate binding and catalytic activity in order to elucidate the mechanism of the enzymes. Trp-62 in chicken egg white lysozyme is the residue most exposed to solvent among the three tryptophans at positions 62, 63 and 108, and is very susceptible to chemical reagents. Trp-62 is involved in the activity of lysozyme (Kumagai, 1987); hence, this residue has been extensively mutagenized. Trp-62 in chicken egg white lysozyme was mutated to Tyr (original amino acid at corresponding site 63 in human lysozyme), Phe, or His and these mutants were found to have enhanced hydrolytic activity (Kumagai et al., 1987; 1993). Trp-62 was further mutated to Leu, Ile, Val, Ala, and Gly and all substitutions but Gly (which caused a 50 % reduction in activity) increased the bacteriolytic activity against Micrococcus lysodeikticus (Maenaka et al., 1994). On the other hand, reduced activity against a simple substrate glycol chitin was observed in all of these Trp-62 mutants and X-ray crystallography analysis of some of these mutants (substitution by Tyr, Phe and His) in the substrate-analogue complex form showed a remarkable change in the substrate binding modes (Maenaka et al., 1994; 1998a; 1998b). This suggests the occurrence of local structural changes around the residue at 62. A triple mutant Trp62His/Asp101Gly/Asn37Gly showed enhanced lytic activity of three-fold more than that of wild type (Kumagai et al., 1992). X-ray crystallography showed that two mutants Trp62Gly and
Asp101Gly complexed with a substrate analogue \((N\text{-acetyl-D-glucosamine})_3\) have the identical structure with that of wild type; however, the substrate analogue, which is an inhibitor of wild type, had no inhibitory effect in the mutant complexes but was hydrolyzed as a substrate. Thus, the interactions of Asp-101 and Trp-62 in subsite B are not vital for the catalytic mechanism; yet co-operatively increase the affinity of the substrate in the productive binding mode (Maenaka et al., 1998a).

There are also numerous studies using site directed mutagenesis of lysozymes in order to understand and elucidate correlative relationships between structural change and protein stability. Conversion of buried residues 40, 55 and 91 by 24 sets of mutations in chicken egg white lysozyme provided three mutants with enhanced thermal stability and specific activity, and other mutants with decreased or unchanged stability and specific activity including I55T variant which exhibited the lowest stability and the aggregation property (Shih et al., 1995). This study of lysozyme core variants showed that the thermostability for most mutants is correlated with hydrophobicity and bulkiness of side-chain of the residues at a certain level; however, neither parameter fully accounts for the stability of all variants suggesting the possibility of other important stability determinants. Several amino acid substitutions have been done in various conformational regions in human lysozyme; two buried positions Ile-56 and Ile-59 (Funahashi et al., 1999), surface hydrophobic residues Val-2, -74 and -110 (Funahashi et al., 2000), and amino acid residues in α-helix and turn regions (Takano et al., 2001). Hydrophobic force, hydration structure through hydrogen bonding, and hydrogen bond interactions have been suggested to play important roles in protein stability; however, the present understanding and information of protein folding still cannot eliminate the involvement of other factors.

Genetic engineering technology has also been used on lysozymes to alter the properties by introducing specific amino acid sequences. N-glycosylation signal insertion by a single amino acid substitution has been used to enhance the stability and activities of lysozymes (Kato
et al., 1998; 2000; Ueda et al., 1996). For example, enhanced antibacterial activity has been achieved by the insertion of hydrophobic peptides at C-terminus resulting in enhanced bactericidal action of chicken egg white lysozyme (Arima et al., 1997; Ibrahim et al., 1994a), and the insertion of polyproline chain at the C-terminus of human lysozyme resulted in higher bactericidal activity (Ito et al., 1997b).

While genetic modification is an attractive tool in protein research, there are still many chemical modifications of proteins that increase the stability and enhance the functional properties of lysozyme. Examples include perillaldehyde-conjugated chicken egg white lysozyme with enhanced antimicrobial action against Gram-negative and Gram-positive bacteria (Ibrahim et al., 1994b), the Maillard-type lysozyme-polysaccharide conjugate with enhanced emulsifying properties and heat stability (Shu et al., 1996), lipophilized chicken egg white lysozyme with increased antimicrobial activity against Gram-negative bacteria (Ibrahim et al., 1991; Liu et al., 2000), and lysozyme-glucose stearic acid monoester conjugate with enhanced resistance to proteolysis, emulsifying activity and emulsion stability (Takahashi et al., 2000). These studies suggest the possibility of lysozyme with enhanced antimicrobial activity and/or some other functions in food and drug application.

1.2.2 Amyloidogenic chicken egg white lysozyme

The ability of a polypeptide to form a tertiary structure as a functional protein in vivo depends on its primary structure and the functions of molecular chaperons and folding-catalyzing proteins. After polypeptides are synthesized in the cell, they go through folding and modification steps before becoming functional proteins. Misfolding of polypeptides is detected by the cell quality control system involving the action of molecular chaperones, and misfolded polypeptides are degraded; however, some misfolded peptides are not recognized by the mechanism and cause several problems. It has been reported that the secretion of unstable
hydrophobic pentapeptide (Phe-Phe-Val-Ala-Pro) fused lysozyme at the C-terminus was enhanced by the combination of the introduction of N-glycosylation signal and the disruption of the calnexin gene in *S. cerevisiae* (Arima *et al.*, 1998). Calnexin is a major molecular chaperon associating with synthesized glycoproteins.

There have been many reports about the relationship between defective protein folding and various human diseases. There are three main categories of protein folding diseases; diseases caused by (1) inability of proteins to fold, (2) toxic folds of proteins, and (3) mislocalization of proteins due to misfolding (Thomas *et al.*, 1995). The most notorious diseases, Alzheimer’s disease and Creutzfeldt-Jacob disease (the human counterpart of mad cow disease), belong to the second category where toxic folds occur in proteins. These diseases are caused by the deposition of aggregated proteins, β-amyloid protein and prion protein, respectively, in the brain. Diseases caused by an accumulation of aggregated proteins are referred as amyloidosis, and proteins that result in the aggregate formation are called amyloidogenic proteins. These protein deposits are called amyloid (starch-like) fibrils, which are detected as a green birefringence under polarized light with Congo red staining (Klunk *et al.*, 1999).

In 1993, it was reported that hereditary non-neuropathic systemic amyloidosis is caused by naturally occurring variants of human lysozyme aggregated in liver (Pepys *et al.*, 1993). Two different single mutations of amino acid in these amyloid fibril forming human lysozymes were found to be the substitutions of threonine for isoleucine at 56 (I56T) and histidine for aspartic acid at 67 (D67H). Since then, many studies have been done on these amyloidogenic human lysozymes and it was found that these proteins have less thermal stability and more surface hydrophobicity than native human lysozyme (Booth *et al.*, 1997). The mechanism of amyloid fibril formation has been assumed to be due to misfolding of the peptide involving an increase in β-sheet structure of protein-folding intermediate, resulting in aggregation (Booth *et al.*, 1997; Bychkova & Ptitsyn, 1995; Guijarro *et al.*, 1998). However, the mechanism of amyloid fibril
formation is not fully understood in terms of what kind of forces are involved in plaque formation. The observation of chicken egg white lysozyme in an ethanol environment suggested one possible mechanism of amyloid fibril formation; as the concentration of ethanol was increased, first the helical content gradually increased; secondly, the tertiary structure was destroyed as the helical content further increased; thirdly, the helical structures were partially destroyed as β-structure increased; and lastly, β-sheets associated with one another and formed the aggregation with the characteristics of amyloid plaques (Goda et al., 2000a). It has been postulated that hydrophobic force play an important role in amyloid fibril formation since the exposure of hydrophobic amino acid residues and the polypeptide backbone buried inside in fully folded structure was observed in a small protein module PI3-SH3 (SH3 domain of the p85alpha subunit of bovine phosphatidylinositol 3-kinase) at low pH where PI3-SH3 generated fibrils (Guijarro et al., 1998). This study also indicated that non-disease related proteins can also form amyloid fibrils in certain environments.

It was reported that chicken egg white lysozyme mutant I55T, in addition to being amyloidogenic, also has reduced thermostability ($T_m$=60.9 °C vs. 74.0 °C for wild type), and increased hydrophobicity (Shih et al., 1995). Furthermore, the chicken egg white lysozyme D66H mutant, had greatly reduced secretion levels in S. cerevisiae, suggesting the occurrence of aggregation (Song et al., 2001). A mass spectrometry study of amyloidogenic human lysozymes I56T and D67H showed that the amyloidogenic behavior is closely related to either a reduction in the folding rate for I56T or an increase in the unfolding rate for D67H (Canet et al., 1999). A recent study showed that in vivo glycosylation improved the stability of amyloidogenic chicken egg white lysozymes I55T and D66H expressed in S. cerevisiae. The polymannosyl chains may suppress the formation of amyloid fibril in these amyloidogenic mutants (Song et al., 2001). As for the I55T/I56T mutations, a study using a synthesized protein showed the single substitution of Ile in the hydrophobic core of the cassette with Ala or Thr had the most dramatic effect on
protein stability (Kwok et al., 1998). In lysozymes, the effect of the substitution of Thr for Ile at hydrophobic core appears distinctively as a conversion of a normal protein into an amyloidogenic protein.

1.2.3  *Pichia pastoris*

Yeast expression systems of foreign genes have proven to be advantageous over prokaryotic and higher eukaryotic systems. The ability of yeast to perform many eukaryote-specific post-translational protein modifications such as proteolytic processing, folding, disulfide bond formation, and glycosylation can result in biologically active, native-form expressed proteins. *E. coli*, on the other hand, often produces high-level but insoluble and inactive eukaryotic proteins, and higher eukaryotic tissue culture expression systems are labor-intensive and costly compared to yeast expression system which requires less equipment and facilities.

Since the discovery of certain yeasts with the ability to utilize methanol as a sole source of carbon and energy in 1969, the methylotrophic yeast *P. pastoris* has been used to produce hundreds of foreign proteins. The methanol utilization pathway involves several unique enzymes such as alcohol oxidase (*AOX*) in the peroxisome to oxidize methanol to formaldehyde and hydrogen peroxide, catalase in the peroxisome to degrade hydrogen peroxide to oxygen and water, two cytoplasmic dehydrogenases to oxidize a portion of formaldehyde to formate and carbon dioxide, and dihydroxyacetone synthase (*DHAS*) in peroxisome to degrade the remaining formaldehyde. *AOX* and *DHAS* are present at high levels in cells grown on methanol, but are absent in cells grown on most other carbon sources such as glucose and glycerol (Cereghino & Cregg, 1999). The main advantages of *P. pastoris* over *S. cerevisiae* as a host of foreign proteins are the *AOX1* promoter and the possibility of high-density fermentation. *AOX1* promoter is tightly regulated and highly repressed under non-methanolic growth conditions; therefore, it is easy to control the growth of *P. pastoris* and induction of genes in this yeast. *P. pastoris* is not a
strong fermenter like *S. cerevisiae*. Ethanol generated in yeast fermentation could be toxic in high-density culture of *S. cerevisiae*, but *P. pastoris* expression strains are relatively easy to culture at cell densities of more than 130 g/litre (dry weight) which is desirable, especially for extracellular expression of proteins where the concentration of the protein secreted is roughly proportional to the concentration of cells (Cregg, 1999).

Genetic manipulation of *P. pastoris* requires similar techniques to that of *S. cerevisiae*. Expression vectors are integrated into the *P. pastoris* genome to maximize the stability of expression strains (Figure 3). The vector is cleaved at a unique site in either the marker gene such as *HIS4* or the *AOX1 5'* and transformed using either the spheroplast generation method or whole cell methods. The homologous recombination events stimulated by the free DNA termini (cleavage) leads to single crossover integration events into these loci at high frequencies (50–80% of *His*<sup>+</sup> transformant) (Cereghino & Cregg, 2000). Gene replacement appear to require longer terminal flanking sequences for efficient integration; replacement frequencies of greater than 20% typically require flanking segments of at least 500 base pairs at each terminus (Cregg *et al.*, 1999). Alternatively, the vector can be digested in such a way that the expression cassette and marker gene are released, flanked by 5' and 3' *AOX1* sequence and the resulting strains are deleted at *AOX1*, therefore the clones depend on the *AOX2* gene. These clones without *AOX1* gene metabolize methanol at a reduced rate, but sometimes express higher levels of heterologous proteins (Cregg & Higgins, 1995).

Several *P. pastoris* expression vectors for intracellular and extracellular expression systems are available and all of these vectors are shuttle plasmids designed for transformation into both *E. coli* and *P. pastoris*. pPICZ-B and pPICZa-C vectors are shown in Figure 4 and Figure 5, respectively. For maintenance in *E. coli*, the plasmids contain an origin of replication (ColE1) and a *Sh ble* gene (for Zeocine resistance). For expression of foreign genes, vectors contain an expression cassette composed of a fragment containing sequences 5' of the *AOX1*
Figure 3 Gene insertion at AOX1 on Pichia genome. Gene insertion events at the AOX1 loci arise from a single crossover event between the loci and either of the two AOX1 regions on the pPICZ or pPICZa vectors: the AOX1 promoter or the AOX1 transcription termination region (TT). This event results in the insertion of one or more copies of the vector upstream or downstream of the AOX1 gene. This figure shows the result of an insertion of the plasmid 5' to the intact AOX1 locus (Mut+) and the gain of PAOX1, a gene of interest, and the Zeocine™ resistance gene (Invitrogen, EasySelect™ Pichia Expression kit manual version A).
**Multiple cloning site**

| SfuI/EcoRI/PmlI/SfiI/BsmBI/Asp7181/KpnI/Xhol/SalI/NotI/XbaI | myc epitope | (His)$_6$ | STOP |

---

**Figure 4** Construction of pPICZ-B plasmid vector. 5' AOX1 indicates a 942 bp fragment containing the AOX1 promoter and this region targets plasmid integration to the AOX1 locus. myc epitope indicates C-terminal myc epitope tag (Glu-Gln-Lys-Leu-ILE-Ser-Glu-Glu-Asp-Leu-Asn) which permits detection of the fusion protein by the Anti-myc Antibody or Anti-myc-HRP Antibody. (His)$_6$ indicates C-terminal polyhistidine tag which permits purification of the fusion protein on metal-chelating resin. AOX1 TT is AOX1 transcription termination which contains native transcription termination and polyadenylation signal from AOX1 gene that permits efficient 3' mRNA processing, including polyadenylation for increased mRNA stability. PTEFI indicates TEF1 promoter which is a transcription elongation factor 1 gene promoter from *S. cerevisiae* that drives expression of the Sh ble gene in *Pichia*, conferring Zeocine™ resistance. PEM7 indicates EM7 promoter which is a constitutive promoter that drives expression of the Sh ble gene in *E. coli*, conferring Zeocine™ resistance. Zeocine indicates Sh ble (*Streptoalloteichus hindustanus* ble) gene for Zeocine™ resistance. cyc1 TT (transcription termination) indicates 3' end of the *S. cerevisiae CYC1* gene that allows efficient 3' mRNA processing of the Sh ble gene for increased stability. ColE1 indicates pUC19 origin which allows replication and maintenance of the plasmid in *E. coli*. XbaI in bold letter in the multiple cloning site indicates the restriction enzyme site used for insertion of the fragment containing lysozyme gene.
Figure 5  Construction of pPICZα-C plasmid vector. α-factor indicates native S. cerevisiae α-factor secretion signal which allows for efficient secretion of most proteins from Pichia. Other indicators are as described in Figure 4. XhoI in bold letter in the multiple cloning site indicates the restriction enzyme site used for insertion of the fragment containing lysozyme gene.
gene (5' AOXI) and just 3' of the gene for transcription termination (AOXI TT). For the secretion of foreign proteins (Figure 5), pPICZα-C and other vectors are available in which sequences encoding the secretion signals of *S. cerevisiae* α mating factor prepro (α-factor) sequence. The α-factor is a 13-residue peptide pheromone secreted from *S. cerevisiae* cells of mating type a and is essential for mating with cells of mating type a (Kjeldsen, 2000). The α-factor is the product of the *MFα1* gene which encodes a 165 residue polypeptide (prepro α-factor) which consists of a 19 amino acid pre-signal sequence, followed by a 66 amino acid pro-sequence with three N-linked glycosylation sites, and a dibasic Kex2 endoprotease processing site at the C-terminus, together with four repeats of the mature α-factor during secretion. In the prepro α-factor, each 13 residue α-factor is preceded by a spacer segment of 4-6 residues such as (Glu-Ala)₂₃ or (Asp-Ala)₂₃ which has a Kex2 dibasic processing site Lys-Arg at the N-terminus. Maturation of the α-factor involves three different proteolytic enzymes for N-terminal processing: the signal peptidase which cleaves between the pre-factor and pro-factor; Kex2 endoprotease, encoded by *KEX2* gene, which cleaves the glycosylated pro-α-factor between pro-factor and each α-factor; and the dipeptidyl aminopeptidase A, encoded by *STE13* gene, which removes Glu-Ala or Asp-Ala dipeptides from the N-terminus. It has been established that the *S. cerevisiae* α-factor prepro peptide is able to confer secretory competence on proteins expressed in *S. cerevisiae* and it has been used widely to facilitate secretion of heterologous proteins (Kjeldsen, 2000). In the extracellular expression system of *P. pastoris*, mature foreign proteins are secreted to the culture supernatant after the prepro α-factor processing. The fusion proteins are processed by the signal peptidase in the endoplasmic reticulum and by the Kex2 endoprotease in the late secretory pathway.

Numerous heterologous proteins expressed in *P. pastoris* were reported and summarized in Cereghino *et al.* (2000). In addition to the advantages of molecular and genetic manipulations with *Saccharomyces*, *P. pastoris* has the advantage of 10 to 100-fold higher heterologous protein
expression levels. The expression level of 20 mg/litre using α-factor signal sequence for chicken egg white lysozyme was reported (Mine et al., 1999), while S. cerevisiae AH22 and GRF180 strains were reported to synthesize and secrete 200 μg/litre and average 4 mg/litre of the protein, respectively (Song et al., 2001, Shih et al., 1993).

1.2.4 Raman spectroscopy

Vibrational spectra may be obtained either by infrared absorption (IR) or Raman scattering spectroscopy and these two techniques provide complementary information. Most photons are elastically scattered (Rayleigh scattering), while the inelastic scattering of photons by molecules (Raman effect) is a very small fraction (Figure 6) (Li-Chan, 1996). The energy of the scattered radiation is less than the incident radiation for the Stokes and is more for anti-Stokes. Since a molecule has to be vibrationally excited prior to irradiation, the anti-Stokes signal is much less intense than the Stokes signal. The inelastic scattering, which results in a Raman shift, is related to the energy of a vibrational transition within a molecule. Characteristic groups of atoms give rise to vibrational bands near the same frequency in spite of their location in the molecule and Raman spectrum is plotted as the intensity of scattered light versus Raman shift wavelength (Pelton et al., 2000). A particular molecule and its environment will determine what Raman signals will be observed. Raman scattering is proportional to the changes in the polarizability of functional groups as the atoms vibrate; therefore, non-polar groups give intense Raman bands. On the other hand, IR absorption is proportional to the changes in the dipole moment of the molecular; thus, nonpolar groups are not IR active. Raman spectroscopy has advantages over IR spectroscopy for *in vivo* or *in situ* study of biological systems, such as foods, since a water molecule is IR active due to its polarity.

A few Raman systems have been developed to overcome several problems in conventional Raman systems. Development of dispersive Raman spectroscopy systems allowed
Figure 6  **Vibrational spectroscopic techniques and molecular energy levels.** Infrared (IR) spectroscopy utilizes the absorption of incident electromagnetic radiation at a particular frequency ($\nu_i$), while Raman spectroscopy utilizes the inelastic scattering of the incident radiation. In Raman scattering, Stokes transitions occur as a result of excitation of molecule ($+\nu_i$) and anti-Stokes occurs as a result of de-excitation of molecule ($-\nu_i$). (Li-Chan, 1996)
the analysis of only microscopic amount of samples, in either liquid or dry form. Thus, the characterization of some difficult samples which cannot be applied to any other spectroscopic methods, such as insoluble samples or samples which are available in very small amounts, can be analyzed by Raman spectroscopy. A common problem of the interference from fluorescence from some samples has been reduced by the development of Fourier-transform (FT) Raman spectrometers which utilizes near-infrared excitation. The utilization of high power laser is beneficial in a fluorescence problem, yet the damage to samples during the analysis has to be considered.

There have been many applications of Raman spectroscopy in protein studies. Raman spectra can provide information about proteins' side chains and also backbone conformation; thus application of Raman spectroscopy is useful to estimate the secondary structure of proteins. Some Raman modes are available for an interpretation of protein structure from Raman spectra (Li-Chan, 1996). Some examples of Raman modes are the indole ring originating from tryptophan which exhibits sharp intense bands at 760, 880 and 1360 cm\(^{-1}\) when buried, and the ring from phenylalanine which exhibits a band at 1006 cm\(^{-1}\), which is independent from its environmental changes. Raman spectroscopy was successfully applied to structural investigation of thermally and/or chemically treated lysozyme and the result indicated increased exposure of aromatic residues and structure changes in treated lysozyme solution (Li-Chan & Nakai, 1991). The Raman optical activity (ROA) study indicated distinctive secondary structural alternations in the amyloidogenic prefibrillar partially denatured intermediate human lysozyme (57 °C, pH 2) (Blanch et al., 2000). ROA gives a complete vibrational optical activity spectrum of chiral molecules via circularly polarized Raman spectroscopy by observing a small difference in the intensity of Raman scattering from chiral molecules measured in right and left circularly polarized incident light. In addition to providing information on secondary backbone and side-group conformation, ROA spectra of proteins contain characteristics of loops and turns, so that
the tertiary fold and its loss and/or change under various conditions can be studied (Barron et al., 2000). The results showed the loss of hydrated α-helix, the formation of poly (L-proline) II-helical (PPII) conformation, and the conformational change of tryptophan residues which are located in the α-domain; however, no increase in β-structure was observed. This study postulated that the conversion of α-helix into β-sheet may be a key step in the amyloid formation, since a substantial loss of tertiary structure was observed in the α-domain but not in the β-domain. The interaction between PPII and β-sheet structure may also be involved in amyloid formation.

An investigation of heat or urea denatured lysozyme and whey proteins in Raman C-H stretching region (2800-3100 cm\(^{-1}\)) suggested the possible application of Raman spectroscopy to monitor the conformational changes of proteins (Howell et al., 1999). Raman spectroscopy was utilized to observe chicken egg white lysozyme binding to a reversed-phase matrix with the significant perturbation of the secondary structure; this demonstrated the possible application of Raman for \textit{in situ} characterization of protein secondary structures (Sane et al., 1999).

1.2.5 Random Centroid Optimization for Genetics (RCG) program

Global optimization is concerned with the determination of global optima (maxima or minima) of a function. Global optimization has increasingly become important both as the area of research and in applications in chemistry and biology. Instead of several time-consuming, and therefore expensive methods to obtain global optimum, random strategies with their simplicity and flexibility are logically more attractive. The algorithm for random global optimization called Random-Centroid Optimization (RCO) was proposed by Nakai et al. (1998a). RCO uses an algorithm that consists of regulated random search, centroid search, and mapping in each cycle and the cycle is repeated until the best response is obtained. The application of regulated random designs effectively avoids local optima, while the centroid design searches around the best results obtained in the process of random search. Mapping allows the visualization of progress
of the optimization sequence but heavily relies on subjective human judgment for the interpretation of the maps. It has been reported that RCO is suitable for application in chemical and engineering processes, and biological systems (Nakai et al., 1998b).

Directed molecular evolution in laboratories is an imitation of natural evolution. In protein engineering, the combination of random mutagenesis and optimization process is a powerful way to mimic natural evolution of proteins, and thus to obtain proteins with improved properties. Random mutagenesis of a protein is useful when little structural information is available for a specific protein. As a logical consequence, RCO program was modified so that the program could be applied to optimization of protein function in combination with site directed mutagenesis (Nakai et al., 1998b). The modified RCO program, Random Centroid Optimization for Genetics (RCG), suggests the site and amino acid for substitution to optimize the properties of proteins according to one of several amino acid scales. Hydrophobicity, helix and strand propensity, bulkiness and charge of amino acids are currently available in RCG program for the substitution process. It has been reported that RCG was successfully applied to improve thermostability of Bacillus stearothermophilus neutral protease with the use of a hydrophobic scale for the initial selection of amino acid substitutions (Nakai et al., 1998b). Since hydrophobic effects are generally believed to be the major stabilizing factor in the formation of tertiary structure of proteins, it is reasonable to utilize a hydrophobicity index of amino acids for random search. There are several amino acid hydrophobicity scales which have been proposed and there are significant differences between those scales. In the current RCG program, the hydrophobic index by Wilce et al. (1995) was used because it was obtained from an extensive HPLC study of peptides rather than a theoretically derived scale. However, it is possible to change scales in the following cycles of optimization, if it is found during mapping that another scale appears to play an important role in the specific function of interest.
The application of RCG to protein studies is beneficial because the studies do not require the structural information of proteins derived from X-ray crystallography or NMR; therefore, RCG is a powerful tool for modifying proteins without known three-dimensional features.

The RCG program can be downloaded from the site ftp://ftp.agsci.ubc.ca/foodsci/ and run on a PC Windows platform.

1.2.6 Homology Similarity Analysis (HSA)

The acquisition of DNA sequence, so called genome structural analysis, is an automated routine process owing to the development of the analytical apparatus. The identification of functional genes, so called genome functional analysis, has been accelerated since the development of DNA chip technology. It has also become important to obtain the technique to analyze the interaction and the binding characteristics of proteins with unknown functions in the field of proteomics. It is expensive, time-consuming, and labor intensive to experimentally elucidate the functions of an unknown gene; thus, some level of prediction prior to experimental analysis will be useful to narrow the possible functions of a protein to a reasonable size for practical analysis. A very useful technique has been to compare the sequence of a new protein to one with a similar sequence, with a known function and/or structure. Based on this idea, the program called Homology Similarity Analysis (HSA) is under development by Dr. S. Nakai (Food, Nutrition, and Health, Faculty of Agricultural Sciences, The University of British Columbia, 6650 N.W. Marine Drive, Vancouver, B.C., Canada V6T 1Z4) using Microsoft Visual Basic 6.0®. This program calculates the similarity coefficient of the amino acid pattern and average propensity within a selected sequence segment in terms of helix, strand, and turn propensities, hydrophobicity, hydrogen bond forming ability, the charge, and the bulkiness. The similarity is calculated as follows: propensity indices of amino acids in a sample segment are plotted against those of corresponding amino acids in a reference segment and the resulting
correlation coefficient (r) is used as the similarity coefficient. By comparing the similarity coefficients and/or average value of specific sequence segment, homology of specific sites among proteins which share different total sequence could be determined. There are several potential applications of this program: (1) identification of binding sites among proteins with the same functions, (2) prediction of conformational change trend by amino acid substitutions, and (3) estimation of unknown protein functions from template protein sequences with known functions. In this research, the HSA program was applied to analyze the nine double lysozyme mutants by calculating the average propensity of the sequence segment around the mutation sites in terms of the factors listed above. The data was compared to that of wild type so that the conformational effects of specific amino acid substitutions could be determined.
CHAPTER II  GENETIC MODIFICATION OF CHICKEN EGG WHITE LYSOZYME

2.1 Introduction

In order to obtain *P. pastoris* clones carrying the chicken egg white lysozyme gene with (I55T variant and other mutants) or without (wild type) mutations, genetic modification of chicken egg white lysozyme was carried out. First, the RCG program was applied and nine sets of random two-site mutations to be inserted into I55T lysozyme cDNA were suggested. These 18 random mutation sites are distributed over the entire sequence by the application of regulated random designs in RCG program. Second, *P. pastoris* vectors containing lysozyme genes with or without mutations were constructed. Lastly, the vectors were transformed into *P. pastoris* and these clones were subjected to protein expression described in the next chapter.

2.2 Materials and Methods

All the molecular cloning protocols are modified from general methods described in Sambrook *et al.* (1989) or from the manuals for the apparatus used, unless otherwise noted.

2.2.1 Mutations by Random Centroid Optimization for Genetics (RCG)

Nine sets of two-site mutations suggested by RCG program for 129 amino-acid chicken egg white lysozyme are listed following: A31F-N44I, K33V-C115F, C6L-Y23C, A95S-G71F, P79I-G126K, I124E-G67I, K116P-T40P, P70F-R61H, and T40E-A42K. The initial selection of an amino acid substitution was carried out using the hydrophobic scale of Wilce *et al.* (1995) shown as follows: -2.24 (His), -1.62 (Lys), -0.85 (Arg), -0.62 (Ser), -0.2 (Asp), -0.1 (Glu), 0.06 (Ala), 0.15 (Gly), 0.21 (Met), 0.25 (Asn), 0.31 (Gln), 0.49 (Cys), 0.65 (Thr), 0.71 (Pro), 1.59 (Val), 1.89 (Tyr), 2.29 (Trp), 3.0 (Ile), 3.5 (Leu) and 4.8 (Phe). Although the hydrophobic scale was used, the choice of the scale for the fist search cycle may not be critical as it is possible to
change the index used in the subsequent cycle after mapping.

2.2.2 Materials, Bacterial Strains and Plasmids

Chicken egg white lysozyme cDNA gene in pUC19 vector was donated by Dr. Kato, Yamaguchi University, Japan (Figure 7). The construction of pUC19 is shown in Figure 8 and the sequence of chicken egg white lysozyme cDNA and the corresponding amino acids is shown in Figure 9. *P. pastoris* X33 strain (wild type), pPICZ-B and pPICZα-C vectors, EasySelect™ *Pichia* Expression Kit (version A) and Zeocine™ were purchased from Invitrogen (San Diego, CA). *E. coli* host strain XL1-Blue (*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F’proAB lacIqZDM15 Tn10 (Tet’)]*) was purchased from Stratagene (La Jolla, CA). T4 ligase and alkaline phosphatase were purchased from Boehringer Mannheim (Indianapolis, IN) and other restriction enzymes were from Roche Diagnostics (Laval, QC). Three kinds of DNA polymerase enzymes were used: Accurase™ from DNAmp Ltd. (Missisauga, ON) and Expand™ High Fidelity PCR System (Roche) for site directed mutagenesis and amplification of fragments and Taq DNA polymerase (Roche) for Direct PCR screening method. dNTPs mixture was prepared from 100 mM ATP, CTP, GTP and TTP stock solutions from Invitrogen. Media components such as Luria Broth (LB) media, yeast extract, peptone, and agar were from Difco Laboratories (Detroit, MI). Agarose for electrophoresis was obtained from Promega Corp. (Madison, WI). All the other chemicals were from Sigma-Aldrich (ON, Canada) unless otherwise noted. Mupid®-2 mini agarose gel electrophoresis apparatus (Advance Co. Ltd, Japan) was used for routine DNA agarose gel electrophoresis and submarine gel electrophoresis apparatus NB-1011 (Nippon Eido, Tokyo, Japan) was used for DNA gel purification. Standard TAE (Tris-acetate-EDTA) buffer was used for the running and the gels were run at 100 V (mini gel) or 120 V (gel purification). DNA thermal cycler 480 (Perkin Elmer, Norwalk, CT) was used for all the PCR reactions.
Figure 7  Construction of chicken egg white lysozyme cDNA in pUC19 vector. Chicken egg white lysozyme cDNA (387 base pairs) with stop codon (TGA) at the C-terminus was inserted into pUC19 multi-cloning site at PstI and SalI.
Figure 8  A sequence map of pUC19 with restriction enzyme cleavage sites. Polylinker region indicates the multi-cloning site. Ap indicates β-lactamase gene conferring resistance toward ampicillin. ORI indicates an origin of replication gene from ColE1 plasmid of E.coli. lac Z indicates a gene encoding β-galactosidase.

(adopted from www.biology.ucsc.edu/classes/bio20L/content/molbio2/puc19.htm)
<table>
<thead>
<tr>
<th>Amino Acids</th>
<th>Nucleotide Code</th>
</tr>
</thead>
</table>
| AAA GTC TTT | KVFGRCELA
| GGA CGA TGT | NAMKRLDH
| GAG CTG GCA | GLDNY
| GCT ATG AAG | CTCTG
| CGA CAC GGA | TAT

Figure 9  **Nucleotide sequence and the corresponding amino acids of chicken egg white lysozyme.**

Bold letters indicate the corresponding amino acid residues to the codon above.

Ile at position 55 is shown underlined.
2.2.3 Primers for site directed mutagenesis

One set of two complementary primers containing I55T mutation and 17 sets of two complementary primers with a series of mutations suggested by RCG were designed to introduce single site mutations; in one case a single primer containing both mutations, T40E-A42K, was used for mutagenesis. All primers were designed to minimize the possibility of self-annealing. Primers were synthesized by Nucleic Acids Protein Services (NAPS) Unit (the University of British Columbia, Vancouver, BC).

The sequences of primers and Tm calculated by the equation (1) are shown in Table 1.

\[
Tm = 81.5 + 0.41 \times (%GC) - 675/N - %\text{mismatch} \tag{1}
\]

\(N\): primer length in base pairs, \(\%GC = (\text{GC in base pairs}/N) \times 100\), \(\%\text{mismatch} = (\text{mismatch}/N) \times 10\)

The concentrations of primers were estimated from the absorbance at 260 nm using Shimadzu UV-visible recording spectrophotometer UV-160 (Kyoto, Japan) according to the equation (2).

\[
\text{Oligonucleotide conc. (pmol/ul)} = \text{absorbance at 260nm} \times \left(\frac{100}{1.5N_A+0.71N_C+1.2N_G+0.84N_T}\right) \tag{2}
\]

\(N_A =\) the number of adenine; \(N_C =\) the number of cytosine; \(N_G =\) the number of guanine; \(N_T =\) the number of thymine

The primers were purified by ethanol precipitation method and the purity was checked according to the absorbance ratio \(OD_{260}/OD_{280}\). The \(OD_{260}/OD_{280}\) values ranged from 1.10 to 1.57 which is lower than the ideal value of 1.8; however, these primers were confirmed to be fully functional by “PCR primer checking method.” (Appendix 1). The PCR reaction was carried out with 20 μl reaction mixture containing 50 ng of template DNA, 125 ng of each primers to be tested, 250 μM dNTP mixture and 2.6 U of Expand High Fidelity PCR system enzyme with the temperature profiles [94 °C for 2 min for template denaturation, 10 cycles of 94 °C for 15 sec, 55 °C for 30
Table 1  Sequence and $Tm$ of primers used in site directed mutagenesis. The bold letters indicate the mutated bases.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>A (Forward, 5'→3')</th>
<th>B (Reverse, 5'→3')</th>
<th>$Tm$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I55T</td>
<td>CCGACTACGGTAACCCCTACAGATCAACAGC</td>
<td>GCTGTTGATCTGTAGGGTACCGTATCGG</td>
<td>73.9</td>
</tr>
<tr>
<td>A31F</td>
<td>AACTGGGTGTTCTGCAAATATCGAG</td>
<td>CTCGAGATTTCGCGAAACACACCCAGTT</td>
<td>67.3</td>
</tr>
<tr>
<td>N44I</td>
<td>ACCCAGGCTACAATCCGTAACACCGAT</td>
<td>ATCGGTGTACAGGATGGCTACGCTGGT</td>
<td>74.0</td>
</tr>
<tr>
<td>K33V</td>
<td>GTGCTGTGCCGAGCTTCTCGAGAGTAAC</td>
<td>GTTACTCTCAGAGACTGCGGCACACAC</td>
<td>68.2</td>
</tr>
<tr>
<td>C115F</td>
<td>TGGCGCAACCCGCTTCACAGGGCACCGAC</td>
<td>GTCGGTGCCCTTGAAAGCGTTGCGCCA</td>
<td>80.1</td>
</tr>
<tr>
<td>C6L</td>
<td>GTCTTTTGACGACTGGAGTGGCAGCA</td>
<td>TGCTGCCAGCTCCAGTCGTCCAAAGAC</td>
<td>69.7</td>
</tr>
<tr>
<td>Y23C</td>
<td>AACTATCGGAGATGCGAGCTGGGAAC</td>
<td>GTTCCCAGGCTGATCCCAGATGTT</td>
<td>75.5</td>
</tr>
<tr>
<td>A95S</td>
<td>AGCGTGAAACCTGCGGAAAGGAGTC</td>
<td>GACGATCTCTTTGCTGCAGGTCCT</td>
<td>66.6</td>
</tr>
<tr>
<td>G71F</td>
<td>GGCAGGACCCCATTTCTCAGGAACCTG</td>
<td>CAGGTTTCTCAGGAATGGGCTCTGGC</td>
<td>74.9</td>
</tr>
<tr>
<td>G126K</td>
<td>GCCTGGAATCAGAAAGTGCGGGTGCTGA</td>
<td>TCACAGCGGGACTTCTGTGATCCGCG</td>
<td>69.6</td>
</tr>
<tr>
<td>P79I</td>
<td>CTCTGCAACACTCATCTGCTGATCAGCCTG</td>
<td>CAGGGCTGAGCAGATGTGTGGCACAG</td>
<td>68.1</td>
</tr>
<tr>
<td>I124E</td>
<td>GTCCAAGCCTGTTGAGAGGGCTGCGG</td>
<td>CCGGCAAGCCTCTCTCCACGCGCTGGAC</td>
<td>75.7</td>
</tr>
<tr>
<td>G67I</td>
<td>TGGTGCAACAGATATCAGGACCCACAGCC</td>
<td>GCCTGGGTTCTGATATCGTGGTACCA</td>
<td>73.3</td>
</tr>
<tr>
<td>K116P</td>
<td>CGCAACCGCTGCCAGGACCGACGTC</td>
<td>GACGTCCGTTGGCTGGCAGCGTTGCG</td>
<td>75.7</td>
</tr>
<tr>
<td>T40P</td>
<td>AGTAACTTCAAACCCACAGGCTACACAC</td>
<td>GTTTGATGCTGTTGGAAGGTACT</td>
<td>63.6</td>
</tr>
<tr>
<td>P70F</td>
<td>GATGCGAGGACCCTTGCGGCACAGAC</td>
<td>GTTCCCGAGGCGGAGGTCTGCCATC</td>
<td>71.2</td>
</tr>
<tr>
<td>R61H</td>
<td>CAGATCAACAGGACACTTGGTGTCGAC</td>
<td>GTCGACCCAGCAGTGGCTTGGATCCTG</td>
<td>75.6</td>
</tr>
<tr>
<td>T40E/A42K</td>
<td>GAGAGTAACTTCAACGAGCAGAAGGACAAACCGTACAC</td>
<td>GGTGTACGTTGCTTTGCTGCTAGTGAAGTTACTCTC</td>
<td>67.7</td>
</tr>
</tbody>
</table>
sec, and 72 °C for 45 sec, then 15 cycles of 94 °C for 15 sec, 55 °C for 30 sec, and 72 °C for 45 sec plus 5 sec for each additional cycles, and 72 °C for 7 min for final elongation] for short fragments <750 base pairs and [94 °C for 2 min for template denaturation, 10 cycles of 94 °C for 15 sec, 55 °C for 30 sec, and 68 °C for 3 min, then 15 cycles of 94 °C for 15 sec, 55 °C for 30 sec, and 68°C for 3 min plus 5 sec for each additional cycle, and 72 °C for 7 min for final elongation] for longer fragments 3,000-4,000 base pairs.

2.2.4 Site directed mutagenesis

Site directed mutagenesis was carried out according to the "Long PCR method" described in Shimada et al. (1996) (Appendix 2). The main advantages of this method over the conventional site directed mutagenesis methods are: (1) there is no need to use specific vectors or hosts, (2) there is no need to prepare single-stranded DNA, and (3) only two primers are necessary.

The template plasmid DNA was directly applied to the PCR reaction with two complementary primers containing one or more mutations to be introduced and a whole plasmid DNA sequence with mutation(s) was amplified. The 50 μl of PCR reaction mixture contained 50 ng of template DNA, 125 ng of each two complementary primers, 200 μM dNTPs mixture, 5 μl of (10×) DNA polymerase buffer, and 1.25 U of Accurase™ or 2.6 U of Expand™ High Fidelity PCR system enzyme mix. The temperature profiles of PCR reaction varied with the enzyme used, as suggested by the manufacturer. For Accurase™, the temperature profile [first 95 °C for 30 sec for template denaturation and 18 cycles of 95 °C for 30 sec, 55 °C for 1 min, and 68 °C for 4 min] was used. For Expand™ High Fidelity PCR enzyme, the temperature profile [94 °C for 2 min for template denaturation, 10 cycles of 94 °C for 15 sec, 55 °C for 30 sec and 68°C for 3 min, 15 cycles of 94 °C for 15 sec, 55°C 30 sec and 68 °C for 3 min plus 5 sec for each additional cycle, and 72 °C for 7 min for final elongation] was used.
The amplification was checked by electrophoresis with 0.7% agarose gel containing 100 μg/ml ethidium bromide (EtBr) under UV light and the remaining PCR reaction mixture was purified by ethanol precipitation by addition of 3-fold volume of 100% ethanol and 0.3 M (final concentration) sodium acetate, incubation at −80 °C for 15 min, centrifugation at 16,000 ×g for 15 min at 4 °C, repeated washing of the precipitate with 70% ethanol and centrifugation, and drying of the precipitate. The DNA precipitate was re-suspended in sterilized water and applied to DpnI digestion for one hour at 37 °C to remove the template plasmid DNA. A portion of the digested mixture was applied on 0.7% agarose gel with EtBr for a rough estimation of DNA concentration.

*E. coli* competent cells were prepared according to a modified rubidium chloride method (Appendix 3). Approximately 1-2 μg of DNA with the volume of less than one-tenth of the volume of the cell suspension was transformed into *E. coli* competent cells according to a general "heat shock" method with some modifications (Appendix 4). The *E. coli* transformant mixture was plated onto LB media containing 100 μg/ml of ampicillin and the plate was incubated for 24 h at 37 °C to allow transformed colonies to grow. Several colonies were selected and grown in 5 ml of LB media containing 100 μg/ml ampicillin for 24 h at 37 °C with 250 rpm shaking. The cells were collected by centrifugation at 16,000 ×g for 5 min and used for DNA extraction according to a modified DNA mini-prep method (Appendix 5).

Extracted DNA plasmids from the *E. coli* transformants were further purified by a standard polyethylene glycol 6000 (PEG-6000) precipitation method (Appendix 6) for DNA sequencing reaction. The purified DNA was applied to DNA sequencing reaction according to the Sanger method (NAPS Unit, UBC). The DNA samples were prepared using BigDye® Terminator Cycle Sequencing chemistry (Taq premix) (Applied Biosystems, Foster City, CA) with pUC19's RV primer (5'-CAGGAAACAGCTATGAC-3'). Twenty microlitres of sequence mixture containing 2 μg of template DNA, 3.2 pmol of primer and 4 μl of Taq premix was
subjected to the temperature profile [25 cycles of 96 °C for 30 sec, 50 °C for 15 sec and 60 °C for 4 min]. The reaction mixture was purified with Centri-Sep™ column (Princeton Separations Inc., Adelphia, NJ) and dried in Speed vac® SC110 (Savant, Holbrook, NY). The DNA Sequencers employed at NAPS Unit are an Applied Biosystems (ABI) Model 373 Stretch and an ABI Prism 377.

2.2.5 Subcloning

Chicken egg white lysozyme wild type and I55T variant genes were subcloned into *P. pastoris* shuttle vectors pPICZ-B and pPICZa-C for intracellular and extracellular expression systems, respectively. The other nine mutants were subcloned into pPICZa-C vector only.

2.2.5.1 Insert DNA preparation. Insert DNA fragments containing chicken egg white lysozyme genes with or without mutations were first amplified using PCR reaction and digested with proper restriction enzymes for the ligation reaction. Lysozyme genes in pUC19 vector were amplified with 5′ primer containing *EcoRI* restriction site (namely pPB primer for pPICZ-B vector) or *XhoI* restriction site (namely pPaC primer for pPICZa-C) and pUC19 3′primer M4 primer. The sequences of these primers are: 5′-AGGAATTCCACCATGAAAGTCTTTTGGA CGATGT-3′ (pPB primer), 5′-GGGTATCACTCGAGAAGAGA|AAAGC7TTGGAC -3′ (pPaC primer), and 5′-GTTTTCCCAGTCACGAC-3′ (M4 primer). The bold letters indicate the restriction sites (*EcoI* and *XhoI*, respectively) and the italic letters indicate the lysozyme sequence. The underline in pPB primer indicates the start codon, and the vertical line in pPaC primer indicates the Kex2 endoprotease cleavage site. A spacer sequence coding Glu-Ala-Glu-Ala was not inserted, since Ste13 cleavage by dipeptidyl aminopeptidase A to remove the spacer sequence Glu-Ala dipeptides from N-terminus has been reported to be incomplete (Kjeldsen, 2000). The construction of insert fragments for pPICZ-B and pPICZa-C vectors are shown in
Figure 10. The PCR amplified DNA fragments were extracted with chloroform/isoamyl alcohol (24:1, vol/vol), precipitated by ethanol and digested with the first restriction enzyme (EcoRI or Xhol) at 37 °C for 24 h. The reaction mixture was purified by ethanol and applied to second restriction enzyme digestion (XbaI). The digested mixtures were directly applied to 50 μl-well agarose gel (1 %) and a single band containing the insert DNA fragment was cut out. DNA was recovered from the gel by squeezing the gel between pieces of parafilm, extracting with chloroform/isoamyl alcohol, and precipitation by ethanol. The concentration of purified insert DNA was visually estimated on agarose gel by comparison with a standard DNA with a known concentration. These inserts have 414 base pairs (insert DNA for pPICZ-B) and 415 base pairs (insert DNA for pPICZa-C).

2.2.5.2 Vector DNA preparation. Vector DNA was prepared for subcloning by digestion with appropriate restriction enzymes and purification steps. pPICZ-B and pPICZa-C were first digested with EcoRI or Xhol at 37 °C for 24 h, respectively, precipitated by ethanol, then digested with XbaI as well. Two-site cut vectors (pPICZ-B/EX and pPICZa-C/XX) were precipitated, separated as a single band in agarose gel (1 %), extracted from the gel, and purified with chloroform/isoamyl alcohol and ethanol. The concentration of purified insert DNA was visually estimated on agarose gel by the comparison with a standard DNA with a known concentration. The resulting vectors have 3267 base pairs (pPICZ-B/EX) and 3506 base pairs (pPICZa-C/XX).

2.2.5.3 Ligation. Insert DNA and vector DNA were run on agarose gel and the concentrations were estimated by visual comparison with the standard DNA with a known concentration. Vector DNA concentration was set at approximately 100 ng. The ratio of vector and insert DNA was set at either 1:1 or 1:3 (mol : mol). The ligation mixture containing vector and insert DNA,
Figure 10  Construction of insert fragments for pPICZ-B and pPICZα-C vectors. Insert DNA fragment was amplified between pPB and M4 primers for ligation into pPICZ-B vector DNA (A) and pPaC and M4 primers for ligation into pPICZα-C vector DNA (B). pPB and pPaC primers introduced restriction enzyme sites EcoRI and XhoI to the inserts, respectively. Fragments were digested with EcoRI or XhoI and XbaI and inserted into pPICZ-B/EX or pPICZα-C/XX vector DNA.
T4 ligase and the buffer with the total volume of 10 µl was incubated at 16 °C for overnight and directly transformed into \textit{E. coli} by electroporation using a Bio-Rad gene pulser® at charging voltage 2500 V, capacitance 25 µF and resistance 200 Ω (Appendix 7). The transformants were plated onto LB media containing 100 µg/ml Zeocine® and the plates incubated at 37 °C for 24 h.

2.2.5.4 Selection of transformants Zeocine®-resistant transformants were selected from LB plate media containing 100 µg/ml and grown in 5 ml of LB media for 24 h. The plasmid DNA was extracted and purified by DNA mini-prep method and checked on 0.7 % agarose gel with ethyldium bromide. The bands on the gel in samples from successful ligation reaction were located at slightly larger molecular weight compared to the control DNA without inserts (pPICZ-B/EX or pPICZa-c/XX) (Figure 11). Those samples with bands at larger molecular weight were digested with proper restriction enzymes (\textit{EcoRI} or \textit{Xhol} and \textit{Xbal}) to confirm the insertion of insert DNA fragment in vector DNA on the gel (Figure 12). These pPICZ-B and pPICZa-C with the lysozyme gene, with or without mutations, were transformed into \textit{E. coli} using the calcium chloride method. The sequences with lysozyme genes with or without mutations in pPICZ vectors were confirmed to be correct.

2.2.6 Transformation of gene into \textit{P. pastoris}

pPICZ vectors carrying the lysozyme gene, with or without mutations, were linearized by \textit{Dral} and transformed into \textit{P. pastoris} by electroporation method with Bio-Rad gene pulser II, as described in \textit{EasySelect™ Pichia Expression kit manual version A}. \textit{P. pastoris} X33 strain was prepared for electroporation and using 1500V for charging voltage, 25 µF for capacitance and the resistance 200Ω were used as described in the manual. Linearized pPICZa-C vector without lysozyme gene was also transformed as a control clone. These \textit{P. pastoris} transformants were selected on YPDS media (YPD and 1M sorbitol) with 100 µg/ml Zeocine®.
Figure 11  Vector DNA with or without insert DNA after subcloning observed on 0.7% agarose gel. Insert DNA successfully inserted into vector DNA was observed in #1,2, and 4, while #3 did not have insert DNA.
Figure 12  Vector DNA with or without insert DNA after restriction enzyme digestion observed on 0.7% agarose gel. Insert DNA fragment cut out of vector and the vector fragment were observed in #1, 2, 4, 5 and 6, while #3 shows a band longer than the vector fragment and no insert fragment.
2.2.6.1 *Direct PCR screening method.* The integration of lysozyme gene into *Pichia* genome was confirmed by a modified Direct PCR screening method, originally described in Linder *et al.*, (1996). The single colony of transformants was suspended in 10 μl of sterilized water, 5 μl of 5 U/μl Lyticase was added to the suspension, and the mixture was incubated at 30 °C for 10-15 min then −80 °C for more than 10 min. The mixture was directly applied to a PCR reaction. Fifty microlitres of reaction mixture containing 5 μl of (10×) Taq DNA polymerase buffer, 500 μM dNTPs mix, 10 pmol of both 5' and 3' AOX primers, 5 μl cell lysate, and 0.8 U of Taq DNA polymerase was applied to PCR reaction with "hot start." The temperature profile used was [94 °C for 5 min for template denaturation, 25 cycles of 95 °C, 54 °C and 72°C for 1 min, and 72 °C for 7 min for final elongation]. The resultant PCR amplified fragments were 1555 base pairs for wild type and all the other mutants and 1168 base pairs for the control clone without lysozyme gene (Figure 13).

2.2.6.2 *Sequence of integrated gene in Pichia genome.* The sequence of an integrated lysozyme gene in pPICZa-C vector in *Pichia* genomic of the randomly selected clone (A31F-N44I) was confirmed to be correct by the following method. For the DNA sequencing, the fragment between 5'-AOX and 3'-AOX priming sites was amplified by PCR from the extract of *Pichia* genomic DNA. Yeast genomic DNA was extracted according to protocols in Invitrogen *Pichia* manual and Hoffman and Winston (1987) with some modifications (Appendix 8). Fifty microlitres of PCR sequencing reaction mixture containing 5 μl of (10×) Expand® High Fidelity buffer, 0.2 μg yeast genomic DNA, 15 pmol 5'AOX and 3'AOX primers, 500 μM dNTPs mix, and 3.5 U of Expand® High Fidelity PCR system enzyme was applied to PCR reaction with hot start. The temperature profile used was the same as in the Direct PCR screening method. The fragment was extracted from the agarose gel as described in the section 2.2.5.1, purified with PEG precipitation method, and then applied to the sequence reactions using the primers (N44I-A...
Figure 13  Fragments amplified in Direct PCR screening method observed on 0.7% agarose gel. Wild and I55T clones were the positive controls which show the fragment of 1,555 base pairs and control clone was the negative controls with 1168 base pairs. Two of nine mutants were shown to have the same length fragment as wild and I55T clones. All the other mutants were also applied to Direct PCR screening method and showed the same positive fragment. Lane 1: wild clone (positive control); lane 2: I55T clone (positive control); lane 3: control clone (negative control); lane 4: mutant 1; lane 5: mutant 2; lane 6: λ/EH marker.
and -B) within the lysozyme gene (Figure 14). Twenty microlitres of PCR sequencing reaction mixture containing 50 ng of the purified fragment, 3.2 pmol of N44I-A or -B primers, and 4 µl Taq Dye premix was applied to PCR sequencing reaction temperature profile. The construction of the gene integrated into *Pichia* genome was confirmed to be correct.

### 2.3 Results and Discussion

Wild type, I55T, nine mutants and control *Pichia* clones were successfully produced and used for protein production as described in the following chapter.

For site directed mutagenesis, there were problems with some primers, as the wrong mutations or no mutation were observed in some sequences. No mutations observed in DNA sequencing might be due to incomplete *DpnI* digestion; however, the wrong mutations around the primer priming sites observed in a few sequence results cannot be explained. The PCR primer checking method was very useful to make sure that primers were not degraded or contaminated. Some modifications in site directed mutagenesis long PCR reaction were made to increase the efficiency. It was found that the concentrations of template DNA and primers were extremely important in the reaction. The optimal concentrations of template DNA and primers slightly varied with the primers used; however, the condition described in the section 2.2.4 was optimal in most of the reactions.

For ligation, there were numerous trials and errors in order to obtain the colonies of transformants. Finally, the ligation problem in this research was found to be due to the low transformation efficiency into *E. coli*; therefore, the problem was solved by application of the electroporation method. Before the cause of the problem was determined, a variety of modifications in the ligation reaction condition were attempted. Different vector/insert DNA ratios at 1:1, 1:3, 1:5 and 1:10 (mol/mol), different reaction temperature profiles (16 °C or 4 °C overnight, or thermo-cycle method (Lund *et al.*, 1996)), and different total DNA concentration
Figure 14  Construction of the amplified fragment from the gene integrated into *Pichia* genome and sites for primers used in sequencing. A gene (pPICZα-C carrying I55T lysozyme with A31F-N44I mutations) integrated into the *P. pastoris* genome was amplified using 5' and 3'-AOX primers. The lysozyme gene was sequenced using N44I-A or N44I-B primers.
ranging from 50 ng to 500 ng were tried. ATP was added to the reaction mixture in the case that ATP in T4 ligase buffer was degraded by thaw-and-freeze cycles; however, addition of ATP was found not to be necessary. Alkaline phosphatase treatment originally included in the ligation protocol was omitted, since there may be a possible inhibitory effect caused by the alkaline phosphatase buffer in the ligation reaction and there was very low possibility of self-ligation in this case. After these trials with no results, the ligation mixture was directly applied to PCR reaction to confirm the insert in the vector DNA using α-factor and 3’-AOX primers (Sun & Lolis, 1995). The α-factor primer locates on insert DNA and the 3’-AOX primer locates on vector DNA; therefore, the detection of free insert DNA in PCR was prevented. The amplified fragment with the size of insert DNA shown on the agarose gel confirmed that the ligation was successful; thus, the ligation problem was addressed in the transformation efficiency and DNA concentration. Since the lower concentration of total DNA gives a more successful ligation reaction in general, vector DNA concentration was fixed at 100 ng. The competency of E. coli competent cell prepared by calcium chloride method was found to be $3 \times 10^3$ cfu (colony forming unit/μg pUC19 plasmid DNA), which might be too low for transformation with the low DNA concentration ligation solution. Therefore, in order to achieve higher transformation efficiency with the low DNA concentration, the electroporation method was applied and Zeocine®-resistant transformants were obtained.

The transformation into P. pastoris with EasyComp™ chemical method (Invitrogen manual) was first used for several trials; however, no transformants colonies were observed in these attempts; therefore, electroporation was applied and all the transformants were successfully obtained. The reason for the difficulty in transformation with the EasyComp™ kit is unknown.

The Direct PCR screening method was used to confirm the integration of the whole pPICZ plasmid into Pichia genome. One argument against application of Direct PCR screening method is the possibility of detecting non-linearized or re-ligated vector DNA, remaining in the
*Pichia* cell. The survival of linearized vector DNA in the cell is impossible unless it re-ligates and the incidence of re-ligation of transformed DNA in the *Pichia* cell is very rare (Invitrogen, personal communication). As well, it is possible to detect cells that have been transformed by plasmids that were not linearized. In case of incomplete linearization of the vector, the intact plasmid DNA might be detected in the screening method; however, the survival of the plasmid DNA is highly unlikely since pPICZ plasmids have no yeast origin of replication (ColEl gene is *E. coli* origin). Therefore, observation of bands at the correct molecular weight in each *Pichia* clone in the Direct PCR screening method indicates the integration of the lysozyme gene into the *Pichia* genome was successful.
CHAPTER III  EXPRESSION OF WILD TYPE AND I55T LYSOZYMES IN PICHIA PASTORIS AND RAMAN SPECTROSCOPIC ANALYSIS

3.1 Introduction

Wild type and genetically engineered I55T lysozyme genes transformed into P. pastoris were overexpressed and the proteins were subjected to Raman spectroscopy analysis in order to compare the structural properties between wild type and I55T lysozymes. First, the expression of lysozymes from Pichia clones was examined in a small-scale expression system, and the enzymatic activity was determined to verify the production of biologically active lysozymes. Second, wild type and I55T lysozymes were expressed in large-scale expression system in order to obtain samples for Raman analysis. The effect of Vectrase-P™ was investigated in the protein expression levels of wild type and I55T lysozymes by measuring enzymatic activity.

3.2 Materials and Methods

3.2.1 Expression media

For yeast growth and expression, YPD or BMGY (1 % yeast extract, 2 % peptone, 100 mM potassium phosphate buffer (pH 6.0), 1.34 % yeast nitrogen base, 0.4 mg/ g biotin, and 1 % glycerol) and BMMY (1 % yeast extract, 2 % peptone, 100 mM potassium phosphate buffer (pH 6.0), 1.34 % yeast nitrogen base, 0.4 mg/ g biotin, and 0.5 % methanol) media listed in the Invitrogen Pichia manual (1996) were used, respectively. Yeast nitrogen base was from Difco, methanol and glycerol were from Fisher Scientific Canada (Ontario), and all the other highly purified chemicals were from Sigma-Aldrich Canada unless otherwise noted.
3.2.2 Protein separation and purification

CM-Toyopearl 650M from Tosoh (Tokyo, Japan) and Econo Pac® CHT-II cartridge 5 ml hydroxyapatite chromatography from Bio-Rad (Hercules, CA) were used for protein purification. The open-column ion-exchange chromatography loaded with CM-Toyopearl resin (2.8 cm x 23 cm) was carried out using Bio-Rad Econo system (EP-1 Econo Pump, EM-1 Econo UV monitor (280 nm), Model 2110 fraction collector, and Model 1327 Chart recorder). Spectra/Por® dialysis tube with MWCO 3500 (Spectrum laboratories Inc. Ft. Lauderdale, FL) was used for dialysis of expressed lysozymes.

Concentration of sample containing lysozyme was carried out with Centricon™-centrifugal filter units (MWCO 3,000) for 2 ml scale or Ultrafree-CM (MWCO 5,000) for 400 µl scale of initial volume (Millipore, Bedford, MA). All protein separation and purification, and dialysis processes were carried out at 4 °C unless otherwise noted.

3.2.3 Protein detection

SDS-PAGE was carried out with Hoefer® Mighty Small II SE-250 SDS-PAG electrophoresis system (Pharmacia Biotech, San Francisco, CA) with the standard Laemmli method (Ausubel, 1995). The gels for SDS-PAGE consist of 15 % resolving gel and 5 % stacking gel with the size of 8 x 10 x 0.75 cm. The samples were mixed with the same amount of (2x) sample loading buffer, heated at 95 °C for 5 min, and loaded onto the gel. After running at 20 mA until the tracking dye reached to the bottom, the gel was transferred into the coomassie brilliant blue (CBB) solution (0.25 % CBB-R250, 50 % methanol, 10 % acetic acid) for staining, and then destained in the destaining solution (20 % methanol, 10 % acetic acid). In some cases, silver staining was carried out after CBB staining according to a standard protocol using silver nitrite solution (Ausubel, 1995).

Mini Trans-Blot® electrophoretic transfer cell, Sequi-blot PVDF membrane, and blot
absorbent filter papers from Bio-Rad were used for blotting. The blotting buffer (25 mM Tris-base, 190 mM glycine, 20% (v/v) methanol), TBS buffer (20 mM Tris, 150 mM NaCl at 7.6) with or without 0.05% Tween-20, and the blocking buffer (1% fat-free skim milk in TBS-Tween) were used. Rabbit anti-chicken egg white lysozyme polyclonal antibodies from Rockland (Gilbertsville, PA) and Chemicon International (Temecula, CA) and rabbit anti-human lysozyme polyclonal antibody from Dako Diagnostics (Mississauga, ON) were utilized as a primary antibody, although the antibody from Rockland was used routinely in this research. Alkaline phosphatase conjugated goat anti-rabbit IgG as a secondary antibody and Sigma Fast™ BCIP/NBT buffered substrate tablet as a precipitating substrate for the detection of alkaline phosphatase activity on membrane were purchased from Sigma (St. Louis, MO).

For Western blotting, proteins were transferred from SDS-PAGE gel to the PVDF membrane for one hour at 100 V. After the blotting, the membrane was rinsed with TBS buffer and soaked in the blocking solution with gentle shaking for at least one hour and the gel was stained with CBB solution to check the blotting efficiency. The membrane was rinsed with TBS-Tween and soaked in the primary antibody in TBS-Tween for at least one hour with gentle shaking. After rinsing in TBS-Tween 5 times for 5 min each with shaking, the membrane was soaked in the secondary antibody in TBS-Tween for one hour with gentle shaking. After rinsing in TBS-Tween 5 times for 5 min each with shaking, the membrane was rinsed in TBS twice for 4 min each. Then, the membrane was soaked in BCIP/NBT solution with shaking until the bands were detected and the reaction was stopped by washing in deionized distilled water.

3.2.4 Lysozyme enzymatic activity assay and estimation of concentration

Lysozyme bacteriolytic activity assay was carried out according to the method using *M. lysodeikticus* cells described in Shih et al. (1993) and Ibrahim et al. (1996) with some modifications. *M. lysodeikticus* cells were suspended in potassium phosphate buffer (pH 6.2) to
achieve an absorbance of 0.75-0.8 at 600 nm. The cell suspension was equilibrated at 28°C in a 1 cm-pathlength cuvette (1 ml). Fifty microlitres of sample solution containing lysozyme was added to the cell solution, mixed by five pipetting action and five inversions and then immediately the decrease of absorbance (-ΔAU) at 600 nm was recorded for 2 min using UNICAM UV/Vis spectrometer UV2 (Thermo Spectronic, NY). The specific activity was calculated as ΔAU divided by the protein concentration. Chicken egg white lysozyme purified from chicken egg white (commercial lysozyme) was donated by Canadian Inovatech Inc. (Abbotsford, BC).

Lysozyme concentration was calculated using the extinction coefficient of $E_{1mg/ml}^{1cm} = 2.64$ at 280 nm (Shih et al., 1995). Minitab version 13 (Minitab Inc. PA) ANOVA one-way T-test was used for statistical analysis.

3.2.5 Small-scale expression of wild type and I55T lysozymes

Prior to the large-scale expression, the expression of lysozymes was confirmed in a small-scale expression system by SDS-PAGE. Single colonies of each *Pichia* transformant transformed with the pPICZα-C vector containing wild type or I55T lysozyme gene were grown in 5 ml of BMGY to achieve $OD_{600}$ of 1.5-2.0. The cells were collected by centrifugation at 2,000 xg for 5 min, resuspended in BMMY to achieve $OD_{600}$ of 1.3-1.5. Sampling of 500 μl of homogenous culture suspension was carried out every 24 h followed by the addition of 0.5 % methanol. The time course change in enzymatic activity of wild type and I55T lysozymes during four-day expression was examined in crude media.

3.2.6 Small-scale expression of wild type and I55T lysozyme with Vectrase-P™

Vectrase-P™ is a synthetic reducing agent, dithiol (+/-)-trans-1,2-bis (2-mercaptoacetamindo) cyclohexane, with the low molecular weight of 262 Da, reduction potential
of -0.24 V, and pK$_a$ values of 8.3 and 9.9. The synthesis of this compound, and its ability to increase the heterologous production of proteins with multiple disulfide bonds in *S. cerevisiae* were reported by Woycechowsky *et al.* (1999). Protein disulfide isomerase (PDI; EC 5.3.4.1) has an essential role to catalyze the unscrambling of non-native disulfide bonds in proteins in the endoplasmic reticulum. Two active-site thiols of PDI are most important in the catalytic activity and a small-molecule dithiol with a low thiol pK$_a$ value and high reduction potential value was synthesized to mimic the activity of PDI which has a molecular weight of 57 kDa. Although not confirmed, it was postulated that a molecule such as Vectrase-P™ with small mass and little charge could gain access to the ER of a yeast cell (Woycechowsky *et al.*, 1999). Addition of this compound at 0.1 mg/ ml to *S. cerevisiae* growth media increased the production of *S. pombe* acid phosphatase, with eight disulfide bonds, by over threefold, which was equivalent to the increase achieved with 15-fold overexpression of protein disulfide isomerase (Woycechowsky *et al.*, 1999).

Vectrase-P™ was kindly provided by BioVectra del (Charlottetown, PE). Vectrase-P™ was added to expression media of wild type and I55T lysozymes extracellular expression system at the concentrations of 0.05, 0.1, 0.2 and 0.4 mg/ ml from the stock solution in DMSO. The specific activity was examined daily to determine the effect of addition of Vectrase-P™. Single colonies of each *Pichia* transformants carrying pPICZa-C gene containing wild type and I55T lysozyme were grown in 5 ml of BMGY to achieve OD$_{600}$ of 1.5-2.0. The cells were collected by centrifugation at 2,000 ×g for 5 min, resuspended in BMMY to achieve OD$_{600}$ of 1.4±0.15. Vectrase-P™ was suspended in DMSO and added to cell suspension at 0.05, 0.1, 0.2, and 0.4 mg/ ml. Control sample contained 2 μl DMSO and blank sample had no additions. Every 24 h, 500 μl of homogenous culture suspension was taken for an activity assay and 0.5 % of methanol was added to each sample for 4 days.
3.2.7 Large-scale expression of wild type lysozyme and purification

A *P. pastoris* transformant carrying the wild type lysozyme gene with α-factor secretion signal sequence (pPICZα-C wild type lysozyme transformant) was inoculated into BMGY media from a single colony. The transformant was grown at 30 °C overnight to achieve OD$_{600}$=1.5-2.0. Cells were collected by centrifugation at 2,000 ×g for 5 min and resuspended into BMMY media to achieve OD$_{600}$=1.3-1.5. The culture was incubated at 30 °C with shaking at 250 rpm for four days and added 0.5 % methanol every 24 h.

The harvesting and separation of secreted protein was carried out using the protocol described in Shih *et al.* (1993) with some modifications. On the fourth day, the culture was harvested by centrifugation at 4 °C at 5,000 ×g for 10 min. The precipitate was washed with an equal volume of 0.5 M NaCl in 50 mM potassium phosphate buffer (pH 6.2) and the weight was recorded, and the second supernatant portion was combined with the first portion. The combined supernatants were centrifuged at 16,000 ×g for 60 min to remove residual cells and debris and diluted with 2 volumes of deionized distilled water. The diluted supernatant was loaded onto an open column containing CM-Toyopearl resin equilibrated with 50 mM potassium phosphate buffer (pH 6.2) at 4 °C at the rate of 1.0 ml/ min. The column was washed with 2-3 bed volumes of the equilibration buffer at 1.0 ml/ min and proteins were eluted with the about three-column volumes of 0 to 1 M NaCl linear gradient at the rate of 0.3 ml/ min. Several peak fractions were applied to SDS-PAGE and Western blotting and fractions containing lysozyme were selected.

CM-Toyopearl peak fractions containing lysozyme were collected and dialyzed in 10 mM sodium phosphate buffer (pH 7.2) at 4 °C for further purification with hydroxyapatite chromatography (HAC). The dialyzed sample was applied to HAC equilibrated with 10 mM sodium phosphate buffer (pH 7.2) and proteins were eluted by 10 mM to 400 mM gradient of sodium phosphate buffer (pH 6.8) at room temperature.
3.2.8 Large-scale expression of I55T lysozyme and purification

Expression of I55T lysozyme was carried out with *P. pastoris* transformant carrying I55T lysozyme gene with α-factor secretion signal sequence (pPICZa-C I55T lysozyme transformant) as described above. However, it was not possible to separate the secreted I55T by any separation method (ion-exchange, hydroxyapatite, hydrophobic interaction, or gel filtration chromatographies, ammonium sulfate precipitation method, or filter concentration method) due to aggregation and precipitation; thus, I55T lysozyme was expressed intracellularly and harvested using the method described in *Pichia* manual with some modifications. One part of harvested cells were suspended in 2 parts of yeast breaking buffer (50 mM sodium phosphate (pH 7.4), 1 mM PMSF, 1 mM EDTA, and 1 % glycerol) and 3 parts of acid-washed glass beads were added. The cells were lysed either in BioSpec (Bartlesville, OK) Bead-Beater® (total volume of 350 ml) or in tubes (for smaller scales) with a vortex mixer. The cell lysis was carried out with 10-15 cycles of 30 sec of beating and at least 1 min of incubation on ice. The cell lysates were centrifuged at 4 °C at 12,000 ×g for 10 min. The supernatant was pooled and the precipitate was washed with 50 mM potassium phosphate buffer (pH 6.2) and centrifuged as well. The supernatants were combined and centrifuged at 20,000 ×g for 60 min to remove residual cell lysates and debris. The clear supernatant was diluted with the same volume of 50 mM potassium phosphate buffer (pH 6.2) and applied to CM-Toyopearl as described above.

3.2.9 Raman spectroscopy

Three Raman systems were used to obtain spectra from commercial, wild type and I55T lysozymes; a JASCO Model NR-1100 laser Raman spectrophotometer with excitation from the 488 nm line of a Spectra-Physics Model 168B argon ion laser, a dispersive Renishaw micro Raman system was used with various laser wavelengths (514, 575, 785 nm) between 200 nm and
2,000 nm, and lastly a Nicolet FT-Raman system Nexus 670 FT-IR with Nexus FT-Raman interface (Thermo Nicolet, Madison, WI) with NdYVO₄ laser exciting at 1.064 μm was utilized.

3.3 Results and Discussion

After confirmation of lysozyme expression in wild type and I55T variant fermentation, both lysozymes were expressed in small-scale extracellular expression system with the presence of various concentrations of Vectrase-P™ and in large-scale extracellular expression system to obtain samples for Raman analysis. The effect of Vectrase-P™ in the expression levels of wild type and I55T variant was investigated in terms of lysozyme enzymatic activity. Wild type lysozyme was expressed in the extracellular expression system of \textit{P. pastoris} and the secreted protein was separated from the expression media and further purified. I55T variant lysozyme was expressed in extracellular expression system of \textit{P. pastoris} and applied to the separation step; however, the secreted protein was impossible to isolate using several chromatographic methods. Therefore, the expression of I55T variant was carried out in intracellular expression system of \textit{P. pastoris}. With extra steps in the separation, I55T variant was successfully separated, yet no purification could be carried out. Commercial, wild type and I55T lysozymes were applied to three Raman spectroscopy systems (JASCO, Renishaw, and Nicolet) and spectra obtained were analyzed.

3.3.1 Enzymatic activity of wild type and I55T lysozyme

Initially the assay method using glycol chitin as a substrate (Imoto & Yagishita, 1971) was used to measure the lysozyme enzymatic activity in crude media sample. The method has been introduced as a simple reproducible reducing group analysis method. However, it was found that the colour reagent containing potassium fericyanide used in this method reacts with media components, yeast extract and peptone and therefore gave false high activity readings.
Thus, *M. lysodeikticus* activity assay method, which has been widely used, was employed in this research.

Time course change in enzymatic activity during induction with 0.5 % methanol is shown in Figure 15. The activity of lysozymes in crude media was expressed based on the cell weight (g). The activities of wild type and I55T lysozymes in expression media were increased linearly with time ($r^2 = 0.9205$, $p= 0.041$ for wild type and $r^2 = 0.979$, $p= 0.011$ for I55T variant). The enzymatic activity of I55T was 2.5 to 5 times lower than that of wild type (20-40% of wild type). It was reported that I55T showed 91% of specific activity of wild type lysozyme expressed in *S. cerevisiae* (Shih et al., 1995).

3.3.2 Expression of wild type and I55T lysozymes with Vectrase-P™

Enzymatic activities per gram wet cell weight of each sample with different Vectrase-P™ concentrations were determined for 4 days (Figure 16). No variation in wet cell weight during 4-day culture was observed; indicating that addition of any concentrations of Vectrase-P™ caused no negative effect on the cell growth.

Addition of Vectrase-P™ to wild type and I55T lysozymes extracellular expression system did not improve the productivity in terms of enzymatic activity per gram wet cell weight. Addition of 0.05 and 0.1 mg/ ml did not influence the expression levels in both wild type and I55T lysozymes, while addition of 0.2 mg/ ml or more appeared to have a negative effect on expressions of wild type and I55T variant. Although it is impossible to investigate the statistical significance because culture samples were not duplicated due to the limitation of Vectrase-P™, a trend of decreasing expression level with increasing amounts of Vectrase-P™ was observed.

Decreased expression levels in terms of enzymatic activity with the addition of higher concentration of Vectrase-P™ may be due to some negative effect of the compound on the yeast expression system. The compound was found in the lysate of the cells in insect cells expression
Figure 15  Enzymatic activity/ gram wet-cell weight of wild and I55T lysozyme in four day fermentation. Wild and I55T lysozymes were expressed extracellularly in *P. pastoris* (four days). Enzymatic activities were checked daily with *Micrococcus lysodeiktus* and specific activity was calculated by dividing the decrease of absorbance (dA) by wet cell weight (g). ♦: wild type lysozyme; •: I55T lysozyme. The linear regression $r^2$ (----) and the p values for wild type are 0.9205 and 0.041, and those for I55T lysozyme are 0.979 and 0.011, respectively.
Figure 16  Time course change of enzymatic activity/ gram wet-cell weight of wild type (A) and I55T variant (B) lysozymes with various concentrations of Vectrase-P™ during a four day fermentation. ♦ blank (no additives), • control (2 μl DMSO without Vectrase-P™), ▲ 0.05 mg/ml, × 0.1 mg/ml, ● 0.2 mg/ml, ○ 0.4 mg/ml.
system, while CHO cells showed the compound in the supernatant, and in both cases the protein production was enhanced as the concentration of Vectrase-P™ was increased from 0 to 100 µM (personal communication with BioVectra dcl). The side effects of this compound are unknown, since there is not much information available for this compound yet.

No improvement in the expression of lysozymes observed in this study indicates that the compound does not enhance expression of all proteins, and that the lower expression level of I55T lysozyme in terms of enzymatic activity may not be due to the scrambled disulfide bond formation upon the synthesis of proteins in the cell. Alternatively, the condition for the protein expression with Vectrase-P™ may not have been optimal for the *P. pastoris* expression system.

3.3.3 Expression and purification of wild type lysozyme

Wild type lysozyme was successfully expressed in *Pichia* extracellular expression system, separated and purified for Raman analysis.

CM-Toyopearl chromatography of wild type lysozyme expressed extracellularly in *P. pastoris* is shown in Figure 17. A major peak B (fractions 46-54) and a small peak A (fractions 40-45) were observed. Fractions 44, 47 and 49 were applied to SDS-PAGE and Western Blotting (Figure 18). The peak A fractions were shown to contain an approximately 16 kDa protein which interacted with anti-lysozyme antibody, and the main peak fractions contained mainly approximately 14.5 kDa lysozyme. Peak B fractions 46-54, exhibiting lysozyme activity, were collected and dialyzed in 2 litre of 10 mM sodium phosphate buffer (pH 7.2) at 4 °C, repeated at least twice, for further purification with hydroxyapatite chromatography (HAC).

The dialyzed sample was applied to HAC and the absorbance at 280 nm was recorded and plotted as shown in Figure 19. One major peak (fractions 10-17, peak A) and one small peak (fractions 18-22, peak B) were observed. Three fractions from each peak were applied to SDS-PAGE (Figure 20). A major band at approximately 14.5 kDa was observed in major peak
Figure 17  Elution profile of wild type lysozyme expressed in *P. pastoris* from CM-Toyopearl ion-exchange chromatography. Wild type lysozyme extracellularly expressed in *P. pastoris* was harvested, diluted with 2-3 volumes of deionized distilled water and applied to CM-Toyopearl equilibrated with 50 mM potassium phosphate buffer (pH 6.2). Lysozyme was eluted using a linear gradient of 0-1 M NaCl at 1.0 ml/min. The dotted line indicates the NaCl gradient.
Figure 18  SDS-PAGE (left) and Western blotting (right) of fractions of wild type lysozyme separated by CM-Toyopearl shown in Figure 18. Fractions 44 (peak A), 47 and 49 (peak B) from CM-Toyopearl were applied to SDS-PAGE (left); fraction 49 showed one band at lysozyme (ca. 14.5 kDa) and fractions 44 and 47 showed lysozyme and a higher molecular weight band (ca. 16 kDa). Both bands were recognized by the anti-chicken egg white lysozyme antibody during Western blotting. Lane 1: commercial lysozyme from chicken egg white; lane 2: fraction 49; lane 3: fraction 44; lane 4: fraction 47; lane 5: Low molecular weight marker; lane 6: fraction 43; lane 7: fraction 47.
Figure 19  Elution profile of peak fractions purified using hydroxyapatite chromatography. Fractions 46-54 from CM-Toyopearl of wild type lysozyme (Figure 17) were applied to hydroxyapatite chromatography equilibrated with 10 mM sodium phosphate buffer (pH 7.2). Lysozyme was eluted using a linear gradient of 10-400 mM sodium phosphate buffer (pH 6.8) at 1.0 ml/min. Fractions were plotted as the absorbance at 280 nm vs. fraction number. The dotted line indicates the sodium phosphate gradient.
Figure 20  **SDS-PAGE image of fractions from hydroxyapatite chromatography fractions of wild type lysozyme.** Fractions from wild type lysozyme purified with hydroxyapatite chromatography were applied to SDS-PAGE. Peak A fractions contained a lysozyme band (ca. 14.5 kDa) and peak B fractions contained higher molecular weight band (ca.16 kDa). Peak A fractions 10-17 were collected, dialyzed and lyophilized for Raman analysis. Lane 1: native lysozyme from chicken egg white; lane 2: fraction 13; lane 3: fraction 16; lane 4: fraction 17; lane 5: low molecular weight marker; lane 6: fraction 18; lane 7: fraction 19; lane 8: fraction 20.
fractions, while the 16 kDa band was observed in small peak fractions. The higher molecular weight band was assumed to be the one observed in ion-exchange chromatography fractions. Fractions 10-17 were collected, dialyzed in 1 litre of distilled water at 4 °C, repeated twice, and lyophilized to obtain purified powder for Raman analysis.

During the purification, some fractions containing a protein with a molecular weight of 16 kDa were observed. The protein with higher molecular weight than lysozyme showed reactivity with anti-lysozyme antibody; thus, it is possible to assume that the protein is lysozyme expressed with extra residues. A possibility of incomplete Kex2 processing has been reported (Kjeldsen, 2000); however, the extra residues which may have been attached to N-terminus of lysozyme in 16 kDa band protein cannot be explained as an incomplete Kex2 processing since the molecular weight of extra residues of pro-α-factor would have been much larger than 2 kDa. The total yield of the protein was calculated to be 7.7 and 9.3 mg/litre in CM-Toyopearl fractions from 2 separate fermentations. Average yield was higher than the reported yield from *S. cerevisiae* expression system (Song, 2001 and Shih *et al.*, 1993). Purified wild type lysozyme possessed specific enzymatic activity, which was comparable to that of commercial lysozyme from chicken egg white (data not shown).

3.3.4 Expression and purification of I55T lysozyme

The expression and purification of I55T lysozyme was much more difficult due to the amyloidogenic nature of the protein.

For extracellular expression system, the expression media containing I55T was applied to ion-exchange chromatography; however, the elution of the protein was not observed. Several chromatographic methods were applied to separate the secreted I55T lysozyme in the expression media. Application of CM-Toyopearl, CM-Sepharose (Pharmacia Fine Chemicals, Sweden) and Q-Sepharose (Pharmacia Biotech) resin at pH 4.8, 6.2 and 7.5, and HAC and HiTrap® test kit
hydrophobic interaction chromatography (butyl Sepharose 4 FastFlow) (Bio-Rad) were not useful for I55T variant separation. Concentration of media containing I55T lysozyme was carried out either using an Amicon ultra-filtrate concentration system 8050 or using salting-out with ammonium sulfate precipitation. Precipitates obtained were partially solubilized by SDS-PAGE sample loading buffer and were confirmed to contain I55T lysozyme by SDS-PAGE. However, the precipitate could be only partially resolubilized by 6M guanidine chloride; most stayed as an insoluble mass. Resolubilized I55T precipitate in 1 M guanidine chloride was applied to FPLC® system (Pharmacia Biotech) with Superdex 200 column (1.5×28 cm) (Amersham) equilibrated with 50 mM sodium acetate buffer (pH 4.8) containing 0.1 M NaCl and 1 M guanidine chloride. One major peak with a leading shoulder peak was observed. Fractions from the shoulder peak and the major peak were applied to SDS-PAGE (data not shown). The major peak contained a band with the molecular weight of lysozyme, while the shoulder peak showed the band with lysozyme molecular weight together with faint one between 20 and 30 kDa. The higher molecular weight band may be a dimer form of lysozyme; however, most of lysozyme was eluted at the correct molecular weight. The purification of I55T with gel chromatography was successful, yet impossible for large-scale application because the fractions containing I55T could not be concentrated enough to obtain a sufficient amount of purified sample. It was calculated that over 500 injections would be required to obtain the quantity (eg. 1 mg) of I55T necessary for Raman analysis. Consequently, the intracellular expression system was used for I55T lysozyme expression.

Separation of I55T lysozyme expressed intracellularly was successful. CM-Toyopearl chromatography of I55T lysozyme expressed intracellularly expressed in P. pastoris showed a main peak B (31-40) with an early shoulder peak A (21-30) (Figure 21) and fractions were applied to SDS-PAGE and Western blotting (Figure 22). Shoulder fractions showed a heavy band at around 20 kDa, while the major peak fractions showed a heavy band at the molecular
Figure 21  Elution profile of I55T lysozyme expressed in *P. pastoris* from CM-Toyopearl chromatography. I55T lysozyme intracellularly expressed in *P. pastoris* was harvested, lysed with glass beads, centrifuged, diluted with 2-3 volumes of the equilibration buffer and applied to CM-Toyopearl equilibrated with 50 mM potassium phosphate buffer (pH 6.2). Lysozyme was eluted using a linear gradient from 0 to 1 M NaCl at 1.2 ml/min. The dotted line indicates the NaCl gradient.
Figure 22  SDS-PAGE (left) and Western blotting (right) of fractions from CM Toyopearl chromatography of I55T lysozyme. CM-Toyopearl I55T lysozyme fractions 23, 26, 30, 34, and 36, and 40 were applied to SDS-PAGE and fractions 23, 26, 30 and 34 were applied to Western blotting. Fractions 23, 26 and 30 showed band at ca. 40 kDa reacting with the anti-lysozyme antibody. Lane 1: fraction 40; lane 2: fraction 36; lane 3: fraction 34; lane 4: fraction low molecular weight marker; lane 5: fraction 30; lane 6: fraction 26; lane 7: fraction 23; lane 8: commercial lysozyme; lane 9: fraction 23; lane 10: fraction 26; lane 11: fraction 30; lane 12: fraction 34.
(Note: SDS-PAGE gel and the membrane are not presented on the same scale.)
weight of lysozyme. The reactivity of the 40 kDa band with anti-lysozyme antibody was observed. Main peak fractions (31-40) were collected and dialyzed in 10 mM sodium phosphate buffer (pH 7.2) at 4 °C for HAC purification. During dialysis, the formation of precipitate was observed. The precipitate could not be re-solubilized in HAC buffer; therefore, HAC purification was not carried out. Instead of HAC purification, CM-Toyopearl peak fractions were diluted with 2-3 volume of 50 mM potassium phosphate buffer (pH 6.2) and applied to CM-Toyopearl (second load). One major peak was observed (Figure 23) and fractions were applied to SDS-PAGE (Figure 24). The peak fractions were confirmed to contain lysozyme and other proteins, especially ones with the molecular weight of approximately 20 kDa and 40 kDa; however, further purification was impossible due to the aggregation nature of I55T. The formation of precipitation during storage at 4 °C was observed and SDS-PAGE showed that the precipitate contained lysozyme (data not shown); therefore, the insoluble precipitation was collected and lyophilized for Raman analysis.

The two CM-Toyopearl applications gave a single peak; however, the fractions still showed several impurity bands in SDS-PAGE (Figure 24). Formation of a precipitate in fractions during storage was due to aggregation of I55T and this phenomenon might have helped the exclusion of impurities from I55T lysozyme. For Raman spectroscopy, homogeneity of samples is desirable but not necessary; therefore, the precipitate was directly applied to Raman analysis. The eluting concentration of NaCl for wild type lysozyme in CM-Toyopearl chromatography was approximately 0.6 M, while the concentrations for I55T were 0.45 M in the first application and 0.35 M in the second application. The reason of varied eluting concentrations is unknown, so application of commercial lysozyme to the CM-Toyopearl chromatography in order to obtain the eluting condition will be required.

HAC and HIC were used for the separation of I55T in cell lysate; however, the separation of lysozyme was not successful because the cell lysate contained many more proteins. This
Figure 23  Second load fractions of CM-Toyopearl chromatography of I55T lysozyme.
Fractions 32-40 (peak B) in Figure 22 were applied to second CM-Toyopearl chromatography (pH 6.2). Lysozyme was eluted using a linear gradient from 0 to 1 M NaCl at 0.3 ml/min. A single peak was observed and fractions were applied to SDS-PAGE. The dotted line indicates the NaCl gradient.
Figure 24  SDS-PAGE of fractions from I55T lysozyme CM-Toyopearl second load fractions. Fractions 25, 26, 27, 28, 29, 30 and 31 from Figure 23 were applied to SDS-PAGE. Early fractions 25-27 were found to contain less lysozyme than late fractions 29-31. Lane 1: fraction 31; lane 2: fraction 30; lane 3: fraction 29; lane 4: fraction 28; lane 5: low molecular weight marker; lane 6: fraction 27; lane 7: fraction 26; lane 8: fraction 25; lane 9: commercial lysozyme from chicken egg white.
suggested that additional chromatographic methods would be required for further purification of lysozyme from crude samples. The total yield of I55T was calculated to be 9.4 mg/litre from CM-Toyopearl fractions. However, these fractions contained more impurities than those from extracellular expression system; therefore, this calculated value is possibly higher than the true concentration of I55T lysozyme. Although I55T variant prepared for Raman analysis was not completely soluble, the enzymatic activity and the protein concentration was measured and the specific enzymatic activity was about 2.7 times lower than that of wild type lysozyme (data not shown). During the isolation of wild type and I55T lysozymes, it was obvious that I55T had very different properties compared to the wild type.

Different properties were observed between the I55T lysozymes expressed in Pichia extracellular expression system and intracellular system in the separation process. It has been reported that the interactions between denatured proteins and cellular components, such as heparin and nucleic acids, promote the protein aggregation (Takase, 1998). Thus, some cellular components of P. pastoris may be responsible for the enhanced aggregation property observed in I55T lysozyme expressed in Pichia extracellular expression system.

3.3.5 Raman spectroscopic analysis

In the JASCO system, commercial lysozyme, purified from chicken egg white, gave the expected spectrum; however, wild type and I55T did not give any signals and strong fluorescent background signals were observed. In the Renishaw system, wild type lysozyme gave clear signals, while I55T gave signals but with high fluorescent background. Lastly, in the Nicolet FT-Raman system clear signals from commercial, wild type and I55T were obtained with less background noise. Spectra were obtained from three accumulations of a 20 sec scan from 400 cm\(^{-1}\) to 2000 cm\(^{-1}\). The Raman spectra of commercial, wild and I55T lysozymes showed distinctive features in the range between 400 and 1700 cm\(^{-1}\) (Figure 25). The tentative
Figure 25  Raman spectra of commercial, wild and I55T lysozyme. Commercial lysozyme from chicken egg white, wild type and I55T lysozymes expressed in P. pastoris were applied to Raman spectroscopy at the wavelength range of 400 cm\(^{-1}\) and 1700 cm\(^{-1}\).
Assignment of peaks in commercial chicken egg white lysozyme Raman spectrum was made with reference to Li-Chan and Nakai (1991) (Table 2). The major signals in this region are: $\nu\text{ S-S (g-g-g)}$ at 508 cm$^{-1}$, $\nu\text{ S-S (g-g-t)}$ at 528 cm$^{-1}$, Trp at 761, 879, 1014, 1343 and 1553 cm$^{-1}$, and Phe at 1005 cm$^{-1}$.

Since the data shown here are based on a single spectroscopic analysis for each sample, it therefore will require confirmation. However, there are some distinct differences observed among these spectra. The signals observed in commercial and wild type lysozymes were similar, while 155T showed distinctive differences in these signals. There are some indications of structural changes in wild type lysozyme compared to commercial lysozyme. A sharp band at 1005 cm$^{-1}$ in all lysozymes is assigned to phenylalanine, which is conformation insensitive, and therefore useful as an internal standard (Li-Chan, 1996). The ratios of intensities at 761 and 1005 cm$^{-1}$ for commercial lysozyme and wild type are approximately 1.28 and 1.86, respectively. Although acquisition of more spectra is required to observe a significant difference in the ratio values, the intensified tryptophan signal at 761 cm$^{-1}$ in wild type is suggested. In addition, stronger tryptophan signals are observed at 1014 cm$^{-1}$, 1343 cm$^{-1}$ and 1365 cm$^{-1}$. This may indicate increased hydrophobicity around the buried tryptophan residue(s) in wild type compared to commercial lysozyme. A strong signal appears at 480 cm$^{-1}$ and this band is assigned to disulfide bonds which are under conformational strain (Li-Chan, 1996). The ratios of intensities at 507 cm$^{-1}$ and 525 cm$^{-1}$ in commercial lysozyme and wild type lysozyme are approximately 3:1 and 2:1, respectively. The band at 507 cm$^{-1}$ and 525 cm$^{-1}$ are assigned to S-S stretch in gauche-gauche-gauche conformation and in gauche-gauche-trans conformation, respectively; therefore, the relative intensities of these bands in commercial lysozyme indicated that three of the four disulfide bonds were in the lowest potential energy conformation (g-g-g) and the fourth in the g-g-t conformation. The change in the ratio of intensities at 507 cm$^{-1}$ and 525 cm$^{-1}$ and the band at 480 cm$^{-1}$ in wild type lysozyme may indicate that one of the disulfide bonds is under
Table 2  Tentative assignment of some major peak bands in Raman spectrum of chicken egg white lysozyme (Li-Chan & Nakai, 1991).

<table>
<thead>
<tr>
<th>wavenumber (cm$^{-1}$)</th>
<th>tentative assignment$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>508</td>
<td>$v$ S-S, $g$-$g$-$g$</td>
</tr>
<tr>
<td>528</td>
<td>$v$ S-S, $g$-$g$-$t$</td>
</tr>
<tr>
<td>538, 543, 754</td>
<td>Trp</td>
</tr>
<tr>
<td>626</td>
<td>Phe</td>
</tr>
<tr>
<td>644</td>
<td>Tyr</td>
</tr>
<tr>
<td>696</td>
<td>$v$ C-S, Met</td>
</tr>
<tr>
<td>761</td>
<td>Trp</td>
</tr>
<tr>
<td>839, 859</td>
<td>Tyr</td>
</tr>
<tr>
<td>879</td>
<td>Trp</td>
</tr>
<tr>
<td>900, 936</td>
<td>$v$ C-C ($\alpha$-helix)</td>
</tr>
<tr>
<td>980</td>
<td>$v$ C-C</td>
</tr>
<tr>
<td>1005</td>
<td>Phe</td>
</tr>
<tr>
<td>1014</td>
<td>Trp</td>
</tr>
<tr>
<td>1030</td>
<td>Phe</td>
</tr>
<tr>
<td>1080, 1107, 1130(sh)$^a$</td>
<td>$v$ C-N</td>
</tr>
<tr>
<td>1178</td>
<td>Tyr</td>
</tr>
<tr>
<td>1200, 1210</td>
<td>Tyr, Phe</td>
</tr>
<tr>
<td>1238, 1258, 1278(sh)</td>
<td>amide III</td>
</tr>
<tr>
<td>1343, 1365</td>
<td>Trp and $\delta$ C-H</td>
</tr>
<tr>
<td>1447, 1460(sh)</td>
<td>$\delta$ CH$_2$</td>
</tr>
<tr>
<td>1553</td>
<td>Trp</td>
</tr>
<tr>
<td>1582</td>
<td>Trp</td>
</tr>
<tr>
<td>1620(sh)</td>
<td>Tyr, Trp, Phe</td>
</tr>
<tr>
<td>1660</td>
<td>amide I</td>
</tr>
</tbody>
</table>

$^a$ (sh) refers to shoulder.
$^b$ $v$ and $\delta$ are stretching and bending vibrations, respectively. $g$-$g$-$g$ and $g$-$g$-$t$ refer to the gauche or trans conformations of the C-S-S-C bonds of the disulfide groups.
conformational strain. Another structural changes are indicated by the changes in amide I and amide III regions, which are useful for the investigation of secondary structure. The sharp signal at 1650 cm\(^{-1}\) in amide I region, which is assigned to α-helix structure, and the decreased signal around 1240 cm\(^{-1}\) in amide III region, which is assigned to anti-parallel β-sheet structure at 1235±5 cm\(^{-1}\) (sharp) and disordered structure at 1245±4 cm\(^{-1}\) (broad) (Li-Chan, 1996), may indicate that wild type lysozyme has higher helix content and lower strand contend compared to commercial lysozyme.

The most significant change in these spectra is observed in I55T spectra. The decrease in the intensity of the peaks at 761, 879, 1014, 1343 and 1553 cm\(^{-1}\) observed in I55T indicates increased exposure of tryptophan residues to the aqueous environment (Li-Chan, 1996). The importance of available hydrophobic groups on the molecular surface for interactions, including those leading to β-sheet formation has been suggested in Raman analysis of egg white proteins (Li-Chan and Qin, 1998). Thus, this result may explain the amyloidogenic property of I55T lysozyme. A strong signal at ~1400 cm\(^{-1}\), which is assigned to C=O stretching of COO\(^-\), is observed in I55T variant. The appearance of this signal may have indicated local change in one of the catalytic residues Glu-35, which is normally localized in hydrophobic area.

It has been reported that heat denatured lysozyme with enhanced antimicrobial activity showed the exposure of tryptophan residues along with dimer formation (Ibrahim et al., 1996). The same phenomenon of tryptophan residue(s) exposure was observed both in heat denatured and I55T lysozymes, which may contribute to the aggregation/unstable property of I55T. In this project, antimicrobial assays were not carried out but it would be worthwhile to investigate I55T antimicrobial functions in comparison with wild type lysozyme.

X-ray crystallography analysis indicated that there was little conformational change in I56T human lysozyme variant compared to wild type, except for the new hydrogen bond formation between threonine side-chain and an intramolecular water molecule (Funahashi et al.,
This seems to contradict the results obtained in this research where the Raman spectrum of I55T chicken egg white lysozyme variant showed distinct changes in protein conformation. It is possible that the condition of crystallization may have contributed to the retention of native structure in I55T variant. There may be differences in conformations of I55T chicken egg white lysozyme variant and I56T human lysozyme variant. These structural changes observed in I55T variant may have been related to the sample preparation process. The I55T lysozyme sample prepared for Raman analysis was aggregated and insoluble, while the heat denatured lysozyme which exhibited the exposure of tryptophan residue(s) was irreversibly dimerized but stayed soluble (Ibrahim et al., 1996). It is possible that the structural change observed in I55T variant was due to the artifact upon the aggregation formation. However, it is still obvious that the I55T lysozyme expressed in P. pastoris has different properties than that expressed in S. cerevisiae.

The application of different expression systems may affect the characteristics of the expressed proteins, since S. cerevisiae (AH22R) was used as a host to express I56T human lysozyme in the X-ray crystallography study. I55T and I56T variants secreted from S. cerevisiae have been successfully purified using the same conditions as used for wild type purification (Shih et al., 1995; Funahashi et al., 1996; 1999); on the other hand, in this work, I55T variant secreted from P. pastoris could not be separated using the method described for variants secreted from S. cerevisiae (see section 3.2.8). It is possible that the aggregation property of I55T observed in separation process in this research may have been particular to the amyloidogenic protein expressed in P. pastoris. It has been reported that the interactions between denatured proteins and some cell components, such as heparin and nucleic acids, promote insoluble aggregate formation of the proteins (Takase, 1998). In addition, the structural changes observed in wild type lysozyme compared to commercial lysozyme may have been related to Pichia extracellular expression system, since the wild type lysozyme with identical amino acid sequence to commercial lysozyme was expected to have identical spectra to that of commercial lysozyme.
Therefore, Raman analysis of wild type lysozyme expressed in *Pichia* intracellular expression system, separated and purified may be useful to investigate the interactions between the protein expressed and other cell components of *P. pastoris*.

In this study, the confirmation of lysozyme expression was carried out by enzymatic activity assay, SDS-PAGE, and Western blotting, but N-terminal protein sequencing was not carried out. The lysozyme genes were designed to have no extra residues at both C- and N-terminus upon the protein expression; however, it is possible that extra residue at N-terminus may have been added if the Kex2 cleavage of α-factor signal sequence was not complete in extracellular expression system. It has been reported that the addition of four extra residues (Glu-Ala-Glu-Ala), which is the alpha-factor spacer sequence, to the N-terminus of human lysozyme expressed in *P. pastoris* cause destabilization (Goda, 2000b). Although no spacer sequence was added in the N-terminus of lysozymes in this study, it remains one of the possibilities that extra residues attached may have contributed to the destabilization of I55T variant and/or the structural difference of I55T variant observed in Raman analysis. In wild type expression, a 16 kDa band was detected by the anti-lysozyme antibody, and it is possible that this band results from the addition of extra residues at the N-terminus. It will be useful to carry out protein sequence on both Raman sample fraction and 16 kDa fraction proteins.

More detailed structural information about wild type and I55T from Raman spectroscopy requires optimized running conditions and the spectra range of 400-1700 nm and 2500-3350 nm. Identification of tryptophan residue(s) exposed in I55T is important for the structural study and this task may be achieved by utilization of Nuclear Magnetic Resonance (NMR) spectrometry or X-ray crystallography, which will bring complemental structural information to Raman spectral analysis. However, the solubility of I55T lysozyme may be a problem in the application of other spectroscopic methods which require the sample to be in a solvent.
CHAPTER IV  EXPRESSION OF NINE LYSOZYME MUTANTS SUGGESTED BY RCG

4.1 Introduction

Using the I55T lysozyme gene as a template, nine sets of two-site mutations, as suggested by RCG, were introduced. *Pichia* clones carrying the I55T gene with two additional mutations, designed for extracellular expression, were successfully obtained. The expression levels and enzymatic activities of the secreted proteins were investigated, and then the protein sequence data were analyzed for the protein structure-function relationship elucidation. First, each mutation used in this work was compared with other naturally occurring lysozyme c family sequences reported in the literature. Next, the effects of each single mutation of the 9 mutants were examined using Homology Similarity Analysis (HSA) program in terms of α-helix, β-strand, and turn propensities, hydrophobicity, hydrogen bond forming ability, charge, and bulkiness of the amino acid sequence space around the mutation. As well, HSA was used for the analysis of 155 chicken egg white lysozyme mutants and 156 human lysozyme mutants reported in the literature, to evaluate the performance of the HSA program with a single mutation analysis.

4.2 Materials and Methods

4.2.1 Protein expression and detection

Expression of nine mutants, wild type and I55T lysozymes were carried out for 4 days with the control clone in 5 ml small-scale expression system described in section 3.2.5. During the induction phase of fermentation, 500 μl sample of homogenous cell culture was taken every 24 h just prior to the addition of 0.5 % methanol. The samples were centrifuged at 4 °C at 10,000 ×g for 5 min and supernatant was immediately assayed for the enzymatic activity. On the fourth day, the harvested supernatant was concentrated with a filter concentrator and applied to
Western blotting described in section 3.2.3 or directly applied to dot blotting. For dot blotting, the protein solution was directly applied onto the PVDF membrane equilibrated with the blotting buffer. After drying, the blocking solution was applied to the membrane, and the same protocol as described for Western blotting was used.

4.2.2 Homology Similarity Analysis (HSA) program

HSA program calculates an average propensity value of an indicated residue segment for each amino acid factor including helix, strand, and turn, hydrophobicity, hydrogen bonding forming ability, charge, and bulkiness. The similarity constants were calculated using wild type lysozyme sequence as a reference for each amino acid factor listed above. A similarity constant is a correlation coefficient (r) derived from the linear regression line of propensity indices for amino acids in a mutant segment plotted against those of corresponding amino acids in the wild type segment. Since this is the first application of this program, some lysozyme mutants reported previously were analyzed prior to the nine mutant analysis in order to verify the functionality of this program. First, 155 chicken egg white lysozyme mutants (Shih et al., 1995) and 156 human lysozyme mutants (Funahashi et al., 1999) were applied to this program in order to evaluate the application of this program for the analysis of effects of single mutations. The sequence segment around the mutation (site 55 for chicken egg white lysozyme and site 56 for human lysozyme) used in HSA was chosen to give the greatest difference in similarity constants. Then, nine double I55T lysozyme mutants were applied to this program and resulting average propensity values and similarity constants were manually compared. The sequence segment analyzed was limited to five amino acid residues including the mutation site and two residues before and two residues after the mutation sites (e.g. a mutation at site 31 has a sequence space of residues from 29 to 33).

The values of helix, strand, and turn propensities, hydrophobicity, hydrogen bond
forming ability, charge, and bulkiness for each amino acid are shown in Table 3 with the source references. The values of helix and strand scales are derived from the energy required to form the structure; therefore, a higher value indicates a low propensity of the residue. For ease of comparison, the helix and strand propensity values of a sequence segment calculated by HSA program were normalized by dividing by the range of each scale; 1.2 for helix and 2.3 for strand, and the calculated values were presented as 1-(calculated value) so that the higher values indicate higher propensity for the structure. The hydrophobicity scale is the same used in RCG program shown in 2.2.1. For the turn propensity values and the hydrogen bond forming ability values, higher value indicates the higher propensity for the factor. Bulkiness is compressibility of each amino acid and charge is represented by isoelectric points of each amino acid residue.

4.3 Results and Discussion

From the result of immuno-blotting, no expression was observed in two mutants K33V-C115F and P79I-G126K, while other mutants had the expression level equivalent to I55T lysozyme (Figure 26). In order to see the effect of methanol concentration in stimulating the production, methanol level was increased from 0.5 to 1.0 and 2.0 %; however, no change in the protein production was observed. Cell density in the expression media was increased from OD<sub>600</sub> of 1.3-1.5 to OD<sub>600</sub> of 2.0-2.5 to see the effect on the protein productivity; however, the adjustment did not have any effect on the lysozyme productivity. The α-factor signal region in one mutant without lysozyme expression (K33V-C115F) was sequenced and confirmed to be correct. Cell homogenates of these two mutants that did not express lysozyme were examined using Western blotting; however, it was not possible to determine if lysozyme had accumulated in the cell due to non-specific binding by the anti-lysozyme antibody caused by the PBS buffer. It was obvious that none of the mutation sets improved the structural stability of I55T lysozyme; thus, further RCG optimization could not be carried out.
Table 3  The values of helix, strand, and turn propensities, hydrophobicity, hydrogen bond forming ability, charge and bulkiness used in the HSA program for each amino acid.

<table>
<thead>
<tr>
<th></th>
<th>Helix</th>
<th>Strand</th>
<th>Hydrophobicity</th>
<th>Turn</th>
<th>H-bond</th>
<th>Charge</th>
<th>Bulkiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>0.617</td>
<td>0.978</td>
<td>0.06</td>
<td>0.740</td>
<td>0.00</td>
<td>6.0</td>
<td>11.50</td>
</tr>
<tr>
<td>Arg</td>
<td>0.753</td>
<td>0.780</td>
<td>-0.85</td>
<td>0.960</td>
<td>0.00</td>
<td>10.8</td>
<td>14.28</td>
</tr>
<tr>
<td>Asn</td>
<td>1.089</td>
<td>0.915</td>
<td>0.25</td>
<td>1.440</td>
<td>0.93</td>
<td>5.4</td>
<td>12.82</td>
</tr>
<tr>
<td>Asp</td>
<td>0.932</td>
<td>1.038</td>
<td>-0.20</td>
<td>1.425</td>
<td>1.03</td>
<td>2.8</td>
<td>11.68</td>
</tr>
<tr>
<td>Cys</td>
<td>1.107</td>
<td>0.573</td>
<td>0.49</td>
<td>1.080</td>
<td>0.80</td>
<td>5.1</td>
<td>13.46</td>
</tr>
<tr>
<td>Gln</td>
<td>0.770</td>
<td>0.863</td>
<td>0.31</td>
<td>1.040</td>
<td>0.64</td>
<td>6.0</td>
<td>14.45</td>
</tr>
<tr>
<td>Glu</td>
<td>0.675</td>
<td>0.962</td>
<td>-0.10</td>
<td>1.030</td>
<td>0.68</td>
<td>3.2</td>
<td>13.57</td>
</tr>
<tr>
<td>Gly</td>
<td>1.361</td>
<td>1.405</td>
<td>0.15</td>
<td>1.600</td>
<td>0.00</td>
<td>6.0</td>
<td>3.40</td>
</tr>
<tr>
<td>His</td>
<td>1.034</td>
<td>0.724</td>
<td>-2.24</td>
<td>0.820</td>
<td>0.10</td>
<td>7.6</td>
<td>13.67</td>
</tr>
<tr>
<td>Ile</td>
<td>0.876</td>
<td>0.502</td>
<td>3.00</td>
<td>0.580</td>
<td>0.08</td>
<td>6.0</td>
<td>21.40</td>
</tr>
<tr>
<td>Leu</td>
<td>0.740</td>
<td>0.766</td>
<td>3.50</td>
<td>0.640</td>
<td>0.08</td>
<td>6.0</td>
<td>21.00</td>
</tr>
<tr>
<td>Lys</td>
<td>0.784</td>
<td>0.841</td>
<td>-1.62</td>
<td>0.820</td>
<td>0.00</td>
<td>9.7</td>
<td>15.70</td>
</tr>
<tr>
<td>Met</td>
<td>0.730</td>
<td>0.729</td>
<td>0.21</td>
<td>0.780</td>
<td>0.18</td>
<td>6.0</td>
<td>16.25</td>
</tr>
<tr>
<td>Phe</td>
<td>0.968</td>
<td>0.585</td>
<td>4.80</td>
<td>0.690</td>
<td>0.45</td>
<td>5.5</td>
<td>19.80</td>
</tr>
<tr>
<td>Pro</td>
<td>1.780</td>
<td>2.613</td>
<td>0.71</td>
<td>1.360</td>
<td>0.00</td>
<td>6.0</td>
<td>17.43</td>
</tr>
<tr>
<td>Ser</td>
<td>0.980</td>
<td>0.790</td>
<td>-0.62</td>
<td>1.300</td>
<td>0.62</td>
<td>6.0</td>
<td>9.47</td>
</tr>
<tr>
<td>Thr</td>
<td>1.053</td>
<td>0.569</td>
<td>0.65</td>
<td>0.990</td>
<td>0.67</td>
<td>6.0</td>
<td>15.80</td>
</tr>
<tr>
<td>Trp</td>
<td>0.910</td>
<td>0.671</td>
<td>2.29</td>
<td>0.760</td>
<td>0.35</td>
<td>6.0</td>
<td>21.61</td>
</tr>
<tr>
<td>Tyr</td>
<td>1.009</td>
<td>0.560</td>
<td>1.89</td>
<td>0.850</td>
<td>0.52</td>
<td>6.0</td>
<td>18.03</td>
</tr>
<tr>
<td>Val</td>
<td>0.940</td>
<td>0.444</td>
<td>1.59</td>
<td>0.480</td>
<td>0.08</td>
<td>6.0</td>
<td>21.50</td>
</tr>
</tbody>
</table>

Reference  a  a  b  c  d  e  f

* a: Muñoz & Serrano (1994); b: Wilce et al. (1995); c: Monné et al. (1999); d: Vijayakumar et al. (1999); e: isoelectric point; f: Gromiha & Ponnuswamy (1993).
Figure 26 Western blotting and dot blotting images of wild, I55T and nine variant lysozymes with control. 1: A31F-N44I, 2: K33V-C115F, 3: C6L-Y23C, 4: A95S-G71F, 5: G126K-P79I, 6: I124E-G67I, 7: K116P-T40P, 8: P70F-R61H, and 9: T40E-A42K. Control was from Pichia control clone fermentation which does not synthesize lysozyme due to the lack of lysozyme gene. No lysozyme expression was observed in two and five from three sets of fermentations, while seven other mutants, I55T variant, wild type, and native lysozyme from chicken egg white show the reactivity with the anti-lysozyme antibody.
Only one mutant, T40E-A42K, displayed measurable enzymatic activity in the crude media; however, the activity was much lower than that of I55T (7.2-16.8 times lower, data not shown) on the basis of enzymatic activity per wet cell weight.

Although there have been many lysozyme mutants produced, there is only one study which reported the lack of secretion of lysozyme mutants. Random mutants of chicken egg white lysozyme which were not secreted from yeast were expressed in *E. coli* and applied to a folding study (Kunichika *et al.*, 1999). The results suggested that these mutants may be unstable and/or have a defect in the folding pathway, so that they were detected by the yeast quality control system and degraded before secreted. After proteins are synthesized in the cytoplasm, they travel to their ultimate locations in or outside of the cell. The majority of proteins are first transported into the lumen of the endoplasmic reticulum (ER) (translocation). Proteins are translocated into ER through a multi-subunit protein complex on the ER membrane and several members of this complex, known as molecular chaperones, play an important role in protein folding, maturation, and degradation. The ER is a highly discriminatory organelle that allows passage only to proteins with an essentially native conformation and molecular chaperones appear to facilitate and/or monitor protein folding and assembly (Gething & Sambrook, 1990). Misfolded proteins are detected and degraded by the actions of molecular chaperones; however, proteins and peptides that are aggregated into amyloid plaques are often resistant to degradation (Thomas *et al.*, 1995). This is an important quality control system, since proteins such as growth factors, toxins, and hormones are translocated into ER and their mislocalization or misfolding can lead to diseases such as cystic fibrosis, atherosclerosis, cancer and diabetes (Thomas *et al.*, 1995). Therefore, results in this research suggests that these mutations decreased the stability of I55T and/or added some defects in the folding pathway so that expressed proteins were degraded by the quality control system of the yeast cell. Therefore, results of this work may be suggesting that mutations in K33V-C115F and P79I-G26K which did not yield any lysozyme expression are
critical residues in folding. It is obvious the amyloidogenic property of I55T was not reduced by the introduction of any of the mutations.

It was unexpected to observe no enzymatic activity in all mutants except for one. It was confirmed that I55T variant possesses reduced enzymatic activity in this research, as it has been reported by others (Shih et al., 1995). This result may indicate that the additional mutations altered the protein conformation, which was conserved in I55T variant at some level, and that the conformation change interfered with the binding of substrate and/or the accessibility of the substrate to catalytic sites.

Although many amino acid substitutions in lysozyme have been done, there have been no mutations applied to the I55T form of lysozyme. Raman study in this project showed that I55T variant has a very different structural conformation compared to commercial lysozyme with more exposed tryptophan residue(s). The altered substrate binding mode of lysozyme Trp-62 mutants has been reported as mentioned previously (Maenaka et al., 1994; 1998a; 1998b). Thus, it may be possible that I55T variant possesses different substrate binding mode and the introduction of additional mutations caused additional problems with the 3-D conformation of the protein which affected substrate binding or catalysis. To verify this hypothesis, more structural information about the I55T variant interacting with a substrate may be required.

Sequences for many other c-type lysozymes are available. Many of these are naturally occurring variants at the mutation sites used in this work. If a single mutation introduced in this work is observed naturally in other c-type lysozymes, the effects of the mutation is possibly less than a mutation which is not present in c-type lysozymes. In order to investigate the significance of single amino acid substitutions on the whole protein structure, naturally occurring variants of the amino acid residues observed at each single mutation sites among all animal lysozymes c were compiled (Prager, 1996b). These are listed in Table 4, together with the original amino acid present in chicken egg white lysozyme, the mutation constructed in this work, and the
Table 4  Summary of mutation site, original amino acid residue, amino acid residues observed among all animal lysozymes, substitution amino acid residue, expression level, and enzymatic activity level of nine mutants.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Site</th>
<th>CEWL</th>
<th>Others</th>
<th>Mutation</th>
<th>Expression</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>44</td>
<td>N</td>
<td>G</td>
<td>I</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>115</td>
<td>C</td>
<td></td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>Y</td>
<td></td>
<td>EFNQR</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>95</td>
<td>A</td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>126</td>
<td>G</td>
<td></td>
<td>DENS</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>124</td>
<td>I</td>
<td></td>
<td>LTVE</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>116</td>
<td>K</td>
<td></td>
<td>EQR</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>70</td>
<td>P</td>
<td></td>
<td>HKR</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>42</td>
<td>A</td>
<td></td>
<td>KSV</td>
<td>Yes</td>
<td>Low</td>
</tr>
</tbody>
</table>
observed expression level and enzymatic activity level. Three mutations introduced at 42, 61, and 71 would probably have less influence than other mutations, since the substituted residues appear naturally in other lysozyme sequences.

Two mutations were made in conserved positions 95 and 115, and the mutation at 115 was one of the mutants which exhibited no lysozyme expression. Cys-115 forms a disulfide bond with Cys-30, and the substitution of residue 115 with Phe may have caused instability of the K33V-C115F mutant. Another mutant C6L-Y23C has the substitution of a Cys residue which forms a disulfide bond as well, and did show expression; however, Cys-6 is not a conserved residue in c-type lysozyme. Thus, it is postulated that a disulfide bond between Cys-30 and Cys-115 is structurally more important for protein stability than that between Cys-6 and Cys-127. In order to analyze the significance of the double mutations which were not expressed, characterization of lysozyme mutants containing single mutation of these four sites may be required. With the available data, it is not possible to understand the significance of the mutants that did not express lysozyme.

4.3.1 Homology Similarity Analysis

Further investigation of the changes made by amino acid substitutions in terms of structural propensity was carried out using the Homology Similarity Analysis (HSA) program. Prior to analysis of the 9 mutants, I55 chicken egg white lysozyme mutants and I56 human lysozyme mutants were applied to HSA program in order to investigate the applicability of this program in single mutant analysis.

First, several I55 chicken egg white lysozyme mutants and I56 human lysozyme mutants reported previously were analyzed using the HSA program. Six I55 chicken egg white lysozyme variants I55L, I55V, I55M, I55F, I55A and I55T from Shih et al. (1995) were compared with wild type in terms of the correlation between HSA data and thermal stability. For each amino
acid propensity factor, the segment range that exhibited the most difference among similarity constants of these mutants was selected. Differences in average propensity values and similarity constants are shown in Table 5 and Table 6, respectively, along with the sequence segment range.

For better visual comparison, average propensity values of these mutants for each factor are shown in Figure 27. A decrease in α-helix similarity constants was observed as the protein stability was reduced, while changes in the average helix propensity do not seem to follow any trend. However, I55T variant which has the lowest thermal stability has the lowest average α-helix propensity and similarity constant. The average propensity and similarity constant for β-strand similarity constant do not show distinctive trends when compared to thermal stability. The most distinctive change in these variants is the hydrophobicity; the hydrophobicity propensity and similarity constant were reduced as the stability was reduced, except for I55F variant. Charge propensity changed only in I55F mutant, which may be related to the low stability of this mutant. No distinctive trends are observed in hydrogen bond forming ability. Increase in turn propensity and decrease in similarity constants are observed as the thermal stability was lowered, except for I55V variant. Average bulkiness tends to decrease as the stability was decreased, yet no trend in the similarity constants is exhibited. Therefore, HSA data seem to indicate some correlation between a change in hydrophobicity, turn propensity and bulkiness and protein stability in this region.

I56 human lysozyme variants I56L, I56G, I56V, I56M, I56T, and I56F (Funahashi et al., 1999) were applied to the HSA program as well and the results for differences in average propensity values and similarity constants are shown in Table 7 and Table 8, respectively. For better visual comparison, average propensity values of these mutants for each factor are shown in Figure 28. Similar to the I55T chicken egg white lysozyme variant, human I56T variant exhibited low helix propensity and similarity constant. I56G mutant exhibits relatively low
Table 5  Differences in average propensity values of sequence segments flanking site 55 in I55 chicken egg white lysozyme mutants comparing to wild type (wild type-mutant).

(dTm is a difference in the melting temperature from that of I55T)

<table>
<thead>
<tr>
<th>mutant</th>
<th>I55L</th>
<th>I55V</th>
<th>I55M</th>
<th>I55F</th>
<th>I55A</th>
<th>I55T</th>
<th>segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helix</td>
<td>-0.03</td>
<td>0.02</td>
<td>-0.04</td>
<td>0.03</td>
<td>-0.07</td>
<td>0.05</td>
<td>54-56</td>
</tr>
<tr>
<td>Strand</td>
<td>0.04</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.01</td>
<td>0.07</td>
<td>0.01</td>
<td>54-56</td>
</tr>
<tr>
<td>Hydrophobicity</td>
<td>-0.17</td>
<td>0.47</td>
<td>0.93</td>
<td>-0.60</td>
<td>0.98</td>
<td>0.78</td>
<td>54-56</td>
</tr>
<tr>
<td>H-bond forming</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.02</td>
<td>-0.07</td>
<td>0.01</td>
<td>-0.01</td>
<td>53-57</td>
</tr>
<tr>
<td>Charge</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.10</td>
<td>0.00</td>
<td>0.00</td>
<td>52-56</td>
</tr>
<tr>
<td>Turn propensity</td>
<td>-0.01</td>
<td>0.02</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-0.08</td>
<td>52-56</td>
</tr>
<tr>
<td>Bulkiness</td>
<td>0.08</td>
<td>-0.02</td>
<td>1.03</td>
<td>0.32</td>
<td>1.98</td>
<td>1.12</td>
<td>52-56</td>
</tr>
<tr>
<td>dTm*</td>
<td>11.9</td>
<td>10.7</td>
<td>7.1</td>
<td>6.6</td>
<td>1.5</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

*Reference: Shih et al., 1995
Table 6  Similarity constants of sequence segments flanking site 55 in I55 chicken egg white lysozyme mutants with wild type as a reference. 
(dTm is a difference in the melting temperature from that of I55T)

<table>
<thead>
<tr>
<th>mutant</th>
<th>I55L</th>
<th>I55V</th>
<th>I55M</th>
<th>I55F</th>
<th>I55A</th>
<th>I55T</th>
<th>segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helix</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
<td>0.96</td>
<td>0.93</td>
<td>54-56</td>
</tr>
<tr>
<td>Strand</td>
<td>0.96</td>
<td>1.00</td>
<td>0.97</td>
<td>1.00</td>
<td>0.84</td>
<td>1.00</td>
<td>54-56</td>
</tr>
<tr>
<td>Hydrophobicity</td>
<td>0.99</td>
<td>0.85</td>
<td>0.16</td>
<td>0.96</td>
<td>0.08</td>
<td>0.40</td>
<td>54-56</td>
</tr>
<tr>
<td>H-bond forming</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.84</td>
<td>0.99</td>
<td>0.63</td>
<td>53-57</td>
</tr>
<tr>
<td>Charge</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>52-56</td>
</tr>
<tr>
<td>Turn propensity</td>
<td>1.00</td>
<td>1.00</td>
<td>0.98</td>
<td>1.00</td>
<td>0.99</td>
<td>0.91</td>
<td>52-56</td>
</tr>
<tr>
<td>Bulkiness</td>
<td>1.00</td>
<td>1.00</td>
<td>0.95</td>
<td>1.00</td>
<td>0.77</td>
<td>0.94</td>
<td>52-56</td>
</tr>
<tr>
<td>dTm*</td>
<td>11.9</td>
<td>10.7</td>
<td>7.1</td>
<td>6.6</td>
<td>1.5</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

* Reference: Shih et al., 1995
Figure 27  Average propensity values of chicken egg white lysozyme wild type and I55 mutants (Shih et al., 1995) in the region flanking site 55; the bar shows the melting temperature relative to I55T (dTm), other factors were indicated by different marks shown in the legend.
Table 7  Differences in average propensity values of sequence segments flanking site 56 in I56 human lysozyme mutants comparing to wild type (wild type-mutant).
(dTm is a difference in the melting temperature from that of I56T)

<table>
<thead>
<tr>
<th>mutant</th>
<th>I56L</th>
<th>I56G</th>
<th>I56V</th>
<th>I56M</th>
<th>I56A</th>
<th>I56T</th>
<th>I56F</th>
<th>segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helix</td>
<td>-0.03</td>
<td>0.13</td>
<td>0.02</td>
<td>-0.04</td>
<td>-0.07</td>
<td>0.05</td>
<td>0.03</td>
<td>54-56</td>
</tr>
<tr>
<td>Strand</td>
<td>0.04</td>
<td>0.13</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.01</td>
<td>0.07</td>
<td>0.01</td>
<td>54-56</td>
</tr>
<tr>
<td>Hydrophobicity</td>
<td>-0.17</td>
<td>0.95</td>
<td>0.47</td>
<td>0.93</td>
<td>0.98</td>
<td>0.78</td>
<td>-0.60</td>
<td>54-56</td>
</tr>
<tr>
<td>H-bond forming</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
<td>-0.02</td>
<td>0.02</td>
<td>-0.12</td>
<td>-0.07</td>
<td>56-61</td>
</tr>
<tr>
<td>Charge</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.10</td>
<td>52-56</td>
</tr>
<tr>
<td>Turn propensity</td>
<td>-0.01</td>
<td>-0.20</td>
<td>0.02</td>
<td>-0.04</td>
<td>-0.03</td>
<td>-0.08</td>
<td>-0.02</td>
<td>52-56</td>
</tr>
<tr>
<td>Bulkiness</td>
<td>0.08</td>
<td>3.60</td>
<td>-0.02</td>
<td>1.03</td>
<td>1.98</td>
<td>1.12</td>
<td>0.32</td>
<td>52-56</td>
</tr>
<tr>
<td>dTm*</td>
<td>12.20</td>
<td>9.80</td>
<td>8.90</td>
<td>6.70</td>
<td>0.00</td>
<td>0.00</td>
<td>-2.50</td>
<td></td>
</tr>
</tbody>
</table>

* Reference: Funahashi et al., 1999
Table 8  Similarity constants of sequence segments flanking site 56 in I56 human lysozyme mutants with wild type as a reference.

(dTm is a difference in the melting temperature from that of I56T)

<table>
<thead>
<tr>
<th>mutant</th>
<th>I56L</th>
<th>I56G</th>
<th>I56V</th>
<th>I56M</th>
<th>I56A</th>
<th>I56T</th>
<th>I56F</th>
<th>segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helix</td>
<td>0.98</td>
<td>0.25</td>
<td>0.99</td>
<td>0.98</td>
<td>0.96</td>
<td>0.93</td>
<td>0.98</td>
<td>54-56</td>
</tr>
<tr>
<td>Strand</td>
<td>0.96</td>
<td>0.45</td>
<td>1.00</td>
<td>0.97</td>
<td>0.84</td>
<td>1.00</td>
<td>1.00</td>
<td>54-56</td>
</tr>
<tr>
<td>Hydrophobicity</td>
<td>0.99</td>
<td>0.13</td>
<td>0.85</td>
<td>0.16</td>
<td>0.08</td>
<td>0.40</td>
<td>0.96</td>
<td>54-56</td>
</tr>
<tr>
<td>H-bond forming</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
<td>0.54</td>
<td>0.81</td>
<td>54-58</td>
</tr>
<tr>
<td>Charge</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.99</td>
<td>52-56</td>
</tr>
<tr>
<td>Turn propensity</td>
<td>1.00</td>
<td>0.31</td>
<td>1.00</td>
<td>0.98</td>
<td>0.99</td>
<td>0.91</td>
<td>1.00</td>
<td>52-56</td>
</tr>
<tr>
<td>Bulkiness</td>
<td>1.00</td>
<td>0.32</td>
<td>1.00</td>
<td>0.95</td>
<td>0.77</td>
<td>0.94</td>
<td>1.00</td>
<td>52-56</td>
</tr>
</tbody>
</table>

| dTm*         | 12.20 | 9.80 | 8.90 | 6.70 | 0.00 | 0.00 | -2.50 |

* Reference: Funahashi et al., 1999
Figure 28  Average propensity values of human lysozyme wild type and I56 mutants (Funahashi et al., 1999) in the region flanking site 56; the bar shows the melting temperature relative to I56T variant (dTm), other factors were indicated by different marks shown in the legend.
similarity constants in most of the factors, yet this mutant is second most stable mutant among 7 mutants. With some exceptions such as I56G and I56F, it seems that changes in hydrophobicity, bulkiness, and turn propensity influence in protein stability in this region.

The data obtained from HSA for I55 and I56 lysozyme mutants was compared to the analysis published with the original data. It has been postulated that the reduced stability of I55T variant was due to the introduction of hydrophilic residue (Thr) into a hydrophobic core (Booth et al., 1997). Hydrophobic interactions are considered to be the main contributors to protein stability (Golovanov et al., 2000). The bulkiness of a substituted amino acid residue is also reported to correlate with protein stability (Takano et al., 1997; Vlassi et al., 1998); reduced bulkiness of the residue creates a cavity, which leads to the destabilization of the protein. However, in many cases the cavity created by mutations could be compensated by the introduction of new water molecule or an internal water molecule shift so that the formation of hydrogen bonding between the side-chain and the water molecule contributed to the stabilization of the protein molecule (Takano et al., 1997; 1999; Vlassi et al., 1998). Only small structural rearrangements were observed locally around the mutations site in I56 mutants from X-ray crystallography analysis (Takano et al., 1995) and it has been postulated that the destabilization mechanism of the mutant protein differs depending on the location of the mutation sites. Therefore, decreased hydrophobicity and bulkiness are likely the main destabilization factors in I55 and I56 lysozyme variants, since sites 55 and 56 are located in the tightly packed hydrophobic core of lysozymes (Funahashi et al., 1999). The effect of turn propensity on protein stabilization has not been discussed extensively; however, increase in turn propensity observed in HSA for I55 and I56 lysozyme mutants may have contributed to protein instability. The introduction of an aromatic amino acid to site 55 and 56 also results in the destabilization of the protein; although both the hydrophobicity propensity and the similarity constant remain high. A decrease in charge property and similarity constant is observed in I55F and I56F. Therefore, the
introduction of aromatic amino acids into a packed hydrophobic core may cause some conformational constriction which leads to the protein destabilization, possibly by the alteration of charge property in this region where the charge distribution may be critical for conformational stabilization.

The chicken and human amyloidogenic I55T and I56T variants exhibited a decrease in α-helix propensity; however, the β-sheet propensity did not increase, although aggregation has been proposed to occur from β-sheet interaction (Blanch et al., 2000). Therefore, it is postulated that reduced hydrophobicity, bulkiness and a possibly increased turn propensity, may be related to protein destabilization, while the process of amyloid fibril formation occurs later. Reduced α-helix propensity in I55T and I56T variants may accelerate the β-sheet aggregation and other regions with different secondary structures may be converted into β-sheet structure due to the secondary structure adaptation to the environment.

Next, nine RCG double mutants were analyzed using the HSA program for each mutation site; average propensities calculated are shown as the difference from that of wild type in Table 9 and similarity constants are shown in Table 10. The data are shown in three categories from left; a mutant which exhibited the enzymatic activity (T40E-A42K), mutants which were expressed but exhibited no enzymatic activity (A31F-N44I, C6L-Y23C, G71F-A95S, G67I-I124E, T40P-K116P, and R61H-P70F), and mutants which did not express (K33V-C115F and P79I-G126K). Since each sequence segment was determined to be five residues with the mutation site in the middle, calculated values for T40E-A42K segments includes both mutations (segment residues 38-42 and 40-44).

From the differences in average propensity values, one significant difference in the P79I-G26K mutant, which did not express protein, is that a higher charge was introduced by the lysine residue comparing to wild type. Although the similarity constant for charge with this mutation is high, the charge distribution in this region may be critical for retention of conformation, resulting
Table 9  Differences in average propensity values of five residue sequence segments flanking each mutation in nine double mutants compared to the wild type (wild type-mutant).
(Bold letters indicate the values which exhibit the largest and second largest variations in each factor)

<table>
<thead>
<tr>
<th>mutant</th>
<th>T40E-A42K</th>
<th>A31F-N44F</th>
<th>C6L-Y23C</th>
<th>G71F-A95S</th>
<th>G67I-I124E</th>
<th>T40P-K116P</th>
<th>R61H-P70F</th>
<th>K33V-C115F</th>
<th>P79I-G126K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helix</td>
<td>-0.04</td>
<td>-0.04</td>
<td>0.06</td>
<td>-0.06</td>
<td>-0.07</td>
<td>0.12</td>
<td>0.05</td>
<td>0.03</td>
<td>-0.15</td>
</tr>
<tr>
<td>Strand</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.03</td>
<td>-0.07</td>
<td>-0.08</td>
<td>0.18</td>
<td>0.00</td>
<td>0.03</td>
<td>-0.18</td>
</tr>
<tr>
<td>Hydrophobicity</td>
<td>0.49</td>
<td>0.49</td>
<td>-0.95</td>
<td>-0.93</td>
<td>-0.57</td>
<td>-0.01</td>
<td>0.28</td>
<td>-0.64</td>
<td>-0.46</td>
</tr>
<tr>
<td>H-bond</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.09</td>
<td>0.14</td>
<td>-0.02</td>
<td>0.13</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>Charge</td>
<td>-0.18</td>
<td>-0.18</td>
<td>0.10</td>
<td>-0.18</td>
<td>0.00</td>
<td>0.00</td>
<td>0.74</td>
<td>0.74</td>
<td>0.00</td>
</tr>
<tr>
<td>Turn</td>
<td>-0.02</td>
<td>-0.02</td>
<td>0.01</td>
<td>0.09</td>
<td>0.20</td>
<td>-0.07</td>
<td>0.03</td>
<td>0.07</td>
<td>0.16</td>
</tr>
<tr>
<td>Bulkiness</td>
<td>-0.39</td>
<td>-0.39</td>
<td>-1.66</td>
<td>-3.28</td>
<td>-3.60</td>
<td>-0.33</td>
<td>0.12</td>
<td>-1.16</td>
<td>-0.79</td>
</tr>
</tbody>
</table>
Table 10  Similarity constants of five residue sequence segments flanking each mutation in nine double mutants with wild type as a reference.
(Bold letters indicate the values which exhibit the smallest and second smallest similarity constants in each factor)

<table>
<thead>
<tr>
<th>mutant</th>
<th>T40E-A42K</th>
<th>A31F-N44I</th>
<th>C6L-Y23C</th>
<th>G71F-A95S</th>
<th>G67I-H24E</th>
<th>T40P-K116P</th>
<th>R61H-P70F</th>
<th>K33V-C115F</th>
<th>P79I-G126K</th>
</tr>
</thead>
<tbody>
<tr>
<td>T40E</td>
<td>0.41</td>
<td>0.70</td>
<td>0.84</td>
<td>0.90</td>
<td>0.35</td>
<td>0.76</td>
<td>0.28</td>
<td>0.92</td>
<td>-0.31</td>
</tr>
<tr>
<td>A42K</td>
<td>0.49</td>
<td>0.90</td>
<td>0.99</td>
<td>0.66</td>
<td>0.95</td>
<td>0.06</td>
<td>0.42</td>
<td>0.97</td>
<td>0.81</td>
</tr>
<tr>
<td>A31F</td>
<td>0.71</td>
<td>0.53</td>
<td>0.97</td>
<td>1.00</td>
<td>0.91</td>
<td>-0.10</td>
<td>0.41</td>
<td>-0.02</td>
<td>0.29</td>
</tr>
<tr>
<td>N44I</td>
<td>0.53</td>
<td>0.53</td>
<td>0.91</td>
<td>0.84</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>C6L</td>
<td>0.77</td>
<td>0.94</td>
<td>0.42</td>
<td>0.96</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>Y23C</td>
<td>0.84</td>
<td>0.99</td>
<td>0.42</td>
<td>0.96</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>G71F</td>
<td>0.90</td>
<td>0.66</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>A95S</td>
<td>0.35</td>
<td>0.95</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>G67I</td>
<td>0.76</td>
<td>0.06</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>H24E</td>
<td>0.28</td>
<td>0.42</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>T40P</td>
<td>0.96</td>
<td>0.75</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>K116P</td>
<td>0.72</td>
<td>1.00</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>R61H</td>
<td>0.96</td>
<td>0.75</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>P70F</td>
<td>0.72</td>
<td>1.00</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>K33V</td>
<td>0.96</td>
<td>0.75</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>C115F</td>
<td>0.72</td>
<td>1.00</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>P79I</td>
<td>0.96</td>
<td>0.75</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>G126K</td>
<td>0.72</td>
<td>1.00</td>
<td>0.51</td>
<td>0.31</td>
<td>0.51</td>
<td>0.31</td>
<td>1.00</td>
<td>0.72</td>
<td>0.81</td>
</tr>
</tbody>
</table>
in the lack of expression in the mutant P79I-G126K. The T40E-A42K mutant, which exhibited enzyme activity, did not show any large variation in all factors compared to other mutations (bold letters indicate the largest and second largest values in each factor). Although A42K mutation alone introduces a higher charge, this mutation is observed in other c-type lysozymes; therefore, the charge of amino acid residues may not play an important role in this region. In addition, the combination of both mutations resulted in modest amino acid propensity changes in this region, which may have contributed to the conservation of enzymatic activity in T40E-A42K. On the other hand, the similarity constants of T40E and A42K segments are relatively low in all the factors because these segments include two mutations instead of one as in other segments. The fact that T40E-A42K mutant is the only mutant exhibiting activity may indicate that this region is not critical in the conformational stability of lysozyme. K33V-C115F mutant exhibits relatively high similarity constants in all factors; however, the deletion of the disulfide bond between Cys-115 and Cys-30 may be critical for expression of the protein. Although A42K, G71F, and R61H mutations exist in c-type lysozymes, similarity constants in these regions for some factors are relatively low. This may indicate that these regions may not be important in the conformational stability of the protein. However, it is very difficult to interpret the difference in effects of amino acid substitutions among the mutants with or without the enzymatic activity and with or without lysozyme expression solely from these sequence segment analysis.

As postulated above, the trigger for protein destabilization, which may precede amyloidosis, may be related to a number of factors, not just increase in β-sheet propensity. On the other hand, an occurrence of amyloidosis may be accelerated by increase in β-sheet propensity and/or decrease in α-helix propensity. Therefore, changes in α-helix and β-strand propensities in I55T variant may affect the amyloidogenic property. In two mutants K33V-C115F and P79I-G126K with lack of lysozyme expression, an increase in β-strand propensity
was observed. On the other hand, T40E-A42K variant with enzymatic activity indicated a decrease in β-strand propensity and an increase in α-helix propensity. Other mutants did not follow any trend. It is postulated that introduction of K33V-C115F and P79I-G126K mutations may have accelerated the amyloidosis process by increasing the β-strand propensity; thus the mutants could not be secreted from the yeast. Changes in α-helix and β-sheet may be a good indicator for predicting effects of mutations on amyloidogenic property of proteins.

HSA results of I55/ I56 lysozyme mutants and nine RCG mutants suggested that the effects of amino acid substitution in protein stability depend more on where the mutation site is located in a protein, rather than the propensities of the sequence space around the mutations. Therefore, it suggests that the secondary structure can be altered depending on the environment. The importance of secondary structure adaptation to its environment has been emphasized by the study in which synthesized peptides of α-helical regions in hamster prion proteins were found to form β-sheet and amyloids suggesting the requirement of the intact tertiary structure of the protein to form helical structure (Gasset et al., 1992). The secondary structure adaptation to its environment including interactions with neighboring residues is also still unpredictable (Peterson et al., 1999). To overcome this problem, a de novo designed protein model system named Structural Cassette Mutagenesis (SCM) model has been suggested (Kwok et al., 1998). This system successfully demonstrated that an amino acid sequence with strong β-sheet propensity formed an α-helical conformation when placed in an appropriate environment. Therefore, there are limitations to examine the protein stability from local amino acid propensities.

Protein stability has been investigated by both practical and theoretical approaches, and hydrogen bonding, hydrophobic force, secondary structure propensities, and electrostatic interactions are generally accepted as important protein stability factors. However, it is debatable what factor is the most important in protein stability, and it is not possible to improve a particular factor without disturbing or influencing any other factors even if the model protein
structure is known. In one of numerous protein stability studies, molecular hydrophobicity potential (MHP) approach has been suggested to investigate the protein stability in terms of hydrophobic interactions as the most important protein stability factor (Golovanov et al., 2000). This study proposed criteria for the choice of the mutation site (should take part in hydrophobic contacts, should not be strongly exposed to solvent, should not participate in hydrogen bonding with the side chains, should not be proline, and should not be important in functional activity) and choice of new residue (should have more hydrophobic CH$_2$ and /or CH$_3$ groups, should not be proline or charged, and should not cause steric bumps). These protein stability studies indicate the complex interactions among all these stability factors to determine overall protein stability.

4.3.2 Summary

Comparison of all the mutations in nine mutants with other lysozyme c family members (Table 4) indicated that the lack of expression in K33V-C115F variant may be due to the deletion of the disulfide bond between Cys-115 and Cys-30. The C6L-Y23C mutant, that also lost the disulfide bond between Cys-6 and Cys-127, exhibited lysozyme expression; therefore, it may indicate that the disulfide bond between Cys-115 and Cys-30 is more critical in protein stability than that between Cys-6 and Cys-127. A95 and C115 are conserved residues in other c-type lysozymes; therefore, the introduction of mutations in these amino acid positions may be more critical than other mutations. On the other hand, A42K, G71F, and R61H mutations are observed in c-type lysozymes; thus, the effect of these mutations could be less than other mutations.

The HSA program was utilized to further investigate the effects of amino acid substitution in a specific region flanking a mutation. From a preliminary application of the HSA program on I55 chicken egg white lysozyme mutants and I56 human lysozyme mutants, it was
found that the introduction of amino acids with smaller bulkiness, less hydrophobicity, and higher turn propensity may be the main contributor to the destabilization of the protein at positions 55 or 56. A decrease in α-helix propensity was observed in both I55T and I56T lysozymes, while no significant change in β-strand propensity was observed. Therefore, it is postulated that the process of amyloidosis in these proteins may be initiated by these destabilization factors. After the initial destabilization, α-strand and β-strand propensities may play an important role in amyloid formation. The decreased α-helix propensity in I55T and I56T lysozymes may affect the aggregation property of these proteins.

In the analysis of nine mutants by the HSA program, the introduction of higher charge in the P79I-G126K variant, which did not exhibit lysozyme expression, was observed; therefore, the lack of expression in this mutant may be due to the introduction of higher charge. Among all nine mutants, T40E-A42K, the only one that exhibited measurable enzymatic activity, could be considered a conservative mutant. This was the only mutant which did not have the highest or second highest change in amino acid average propensities in the segments, when compared to wild type. Although A42K mutation introduces higher charge in the region, this mutation is observed in c-type lysozymes and the propensity changes introduced by two mutations were cancelled out by each other. Thus, the preservation of partial enzymatic activity in T40E-A42K mutant may be explained by less variation in propensities. The effects of mutations seem to be largely depending on where the mutation is located in a whole protein molecule, rather than on the characteristics of the sequence space around the mutation.

The effects of mutations on amyloidogenic property have been observed in terms of changes in α-helix and β-strand propensities. T40E-A42K mutations resulted in α-helix propensity increase and β-strand propensity decrease, while K33V-C115F and P79I-G126K mutations resulted in β-strand propensity increase. K33V-C115F and P79I-G126K mutants may
have been accelerated the amyloidosis process; thus, these mutants could not be secreted from the yeast.

For efficient alternation of the protein properties, two sites were introduced at once into I55T lysozyme in this study; however, each mutation should be looked at individually to determine the effect on lysozyme stability and activity. This study showed the difficulty in reducing amyloidogenic nature of I55T lysozyme by introducing random mutations.
CHAPTER V CONCLUSION AND RECOMMENDED FUTURE WORK

This is the first research attempting to reduce the amyloidogenic nature of chicken egg white lysozyme variant I55T by introducing random mutations suggested by Random Centroid Optimization for Genetic (RCG) program. This project provided the first preliminary structural information of I55T variant in comparison with commercial and wild type lysozymes by Raman spectroscopy. The application of P. pastoris as an expression host for amyloidogenic lysozyme in this project gives future prospect in utilization of P. pastoris.

The sequence of the wild type lysozyme, I55T variant and other nine lysozyme variants containing I55T and additional two site mutations were genetically manipulated for protein expression in P. pastoris. Wild type and I55T lysozyme were expressed, separated and purified (not I55T lysozyme) for Raman spectroscopic analysis. The preliminary information obtained from the Raman analysis showed the exposure of hydrophobic tryptophan residue(s) to an aqueous environment in I55T variant. This result suggests the conformational alternation induced by the single amino acid residue substitution, I55T, and this may be related to the aggregation tendency in amyloidogenic I55T lysozyme. I55T lysozyme showed very different properties comparing to wild type lysozyme during the expression process; secreted I55T variant could not be separated using several ion-exchange and affinity chromatographies. The Raman spectrum of I55T variant showed a strong band at ~1400 cm\(^{-1}\), which is assigned to C=O stretch of COO\(^{-}\) and does not appear strongly in commercial and wild type lysozymes. This may be related to a change in one of the catalytic residues Glu-35 which is usually localized in hydrophobic region, and may have been affected the enzymatic activity of I55T variant.

The Raman spectrum of wild type was similar to that of commercial lysozyme; however, some changes in wild type lysozyme compared to commercial lysozyme were suggested; including enhanced hydrophobicity around tryptophan residue(s), possible conformational strain
around one of four disulfide bonds, and changes in α-helix and β-strand contents. Since wild
type lysozyme was expressed extracellularly and I55T lysozyme was intracellularly, Raman
analysis of wild type lysozyme expressed intracellularly may be useful for the investigation of
protein folding of secreted proteins compared to intracellularly expressed proteins.

Further Raman spectroscopic analysis is required for verification of the reproducibility of
the information obtained in this project. The identification of the tryptophan residue(s) which is
exposed will be significant in order to understand the nature of amyloidogenic lysozyme and
other spectroscopic methods may be useful to identify the tryptophan residue(s) exposed, such as
NMR spectrometry or X-ray crystallography; however, it is necessary to overcome the problem
with the solubility of I55T lysozyme for future application of any other spectroscopic methods
which require samples to be dissolved in a solvent.

It may be important to investigate the antimicrobial activity of I55T comparing to wild
type lysozyme, since denatured lysozyme and mutants with tryptophan residue exposed have
been reported to possess enhanced antimicrobial activity against both Gram-negative and Gram-
positive bacteria (Ibrahim et al., 1996).

The utilization of P. pastoris in amyloidogenic protein study may not be suitable, since
enhanced tendency of aggregation were observed in I55T variant expressed from P. pastoris. In
addition, conformational differences in wild type lysozyme compared to commercial lysozyme
were suggested by Raman analysis; therefore, P. pastoris extracellular expression system may
involve some cellular components which affect protein folding. Although P. pastoris is an
attractive expression system as it has higher heterologous protein expression levels, application
of P. pastoris in amyloidogenic protein expression may not be advantageous unless the condition
to reduce the aggregation tendency of expressed protein is achieved.

Two out of nine lysozyme variants exhibited no protein expression level and six of the
seven variants that were expressed did not possess enzymatic activity. These results indicated
the difficulty of reducing the amyloidogenic property of I55T lysozyme by introducing random mutations. The result of protein expression level strongly indicates that the mutations of K33V-C115F and P79I-G126K have significant negative effects on protein stability. At the start of the project, it was anticipated that more significant improvements in protein stability of the I55T mutant would be achieved by modifying two sites simultaneously. However, knowing the difficulty of altering the amyloidogenic property of I55T, it may be useful to introduce a single mutation at a time in order to find out the critical residues in folding of I55T. This approach would verify whether the effect of the double mutations was a result of independent effects, or whether the combination of two mutations was required for the decrease in stability. The reduced enzymatic activity in most of the mutants may suggest the mutations caused conformational changes that affected the active site.

Substrate binding of I55T lysozyme should be investigated to understand the effect on enzymatic activity in this mutant, since it has been reported that the substrate binding mode was altered in some lysozyme mutants (Kumagai et al., 1993; Maenaka et al., 1998b).

This first application of the HSA program in examining the effects of each single mutation within the sequence segment flanking the mutation site suggested some possible explanations of the effect of single amino acid substitution. The analysis of 155 chicken egg white lysozyme mutants and 156 human lysozyme mutants indicated that hydrophobicity, bulkiness, and turn propensity of amino acid residues are important factor in the packed core region in protein stability. It was also postulated that the charge distribution in this region may play an important role. Low α-helix propensity was observed in I55T and I56T amyloidogenic mutants, while β-strand propensity did not exhibit any significant alteration. Therefore, it was postulated that the destabilization factors such as lower hydrophobicity and bulkiness and higher turn propensity may initiate the amyloidosis process, followed by an association of β-sheet structure and a conversion of α-helix structure. Low α-helix propensity in amyloidogenic
lysozyme mutants may indicate that the transition between α-helix and β-strand propensities are important once the amyloidosis process started.

The analysis of nine mutants suggested that the introduction of high charge at Gly-126 may have lead to destabilization of P79I-G126K mutant, and the moderate change in propensities by the combination of two vicinal mutations in T40E-A42K may be related to the conservation of enzymatic activity in this mutant. HSA results also indicated that some residues may not play an important role in protein stability, such as Ala-42, Gly-71 and Arg-61, since naturally occurring amino acid residues in c-type lysozymes (Lys, Phe, and His, respectively) at these positions exhibited relatively low similarity constants. Therefore, HSA data obtained from 155 chicken egg white lysozyme mutants, 156 human lysozyme mutants and nine RCG double mutants seem to indicate that the effect of a single mutation is dependent on where the site is located in a whole protein molecule than the local characteristics of a sequence segment around the site. The effects of these mutations on amyloidogenic property of I55T in terms of α-helix and β-strand propensities suggested that the introduction of higher β-strand propensity may have accelerated the amyloidosis process. This first application of the HSA program may provide preliminary information for further development of the program.

There were two possibilities for the lack of lysozyme expression in K33V-C115F and P79I-G126K mutants; destabilization and enhanced amyloidogenic property of I55T lysozyme. Destabilization factors may vary depending on the site of mutation, and some possible factors are described above. Acceleration of amyloidosis process may be indicated in HSA results as increase in β-sheet propensity in these mutants. Thus, it may be possible to predict an effect of a mutation on amyloidogenic proteins in terms of a change in β-sheet propensity.

In order to alter the amyloidogenic nature of I55T, a more efficient approach is required. RCG may work better with the Homology Similarity program to introduce semi-random mutations efficiently by narrowing down the sequence space to be mutated. Instead of
introducing two sites at once, single mutation introduced to I55T lysozyme may effectively suggest the significant residues in the protein folding process.

In this research, RCG program was utilized as a new approach to alter the amyloidogenic property of I55T chicken egg white lysozyme. Although the reduction of amyloidogenic property of the lysozyme by introduction of two site random mutations could not be achieved, the results obtained in this project will be important information in future proteomics and amyloidogenic protein research. The preliminary information obtained from Raman spectroscopic study of I55T amyloidogenic lysozyme may be useful, accompanied with further structural investigations, in amyloidogenic protein research. In order to elucidate the mechanism of amyloid formation and prevent and/or eliminate the occurrence of diseases such as Alzheimer’s disease, a more detailed understanding of protein structure-function relationship will be necessary.
Reference


Appendix 1  Scheme of Primer checking method.

(A) PCR amplification of fragment with A-5' and B-3' primers
(B) PCR amplification of fragment with B-5' and A-3' primers

(A) 300 base pairs fragment
(B) 2,750 base pairs fragment

Length of primers = 25 base pairs
Appendix 2  Scheme of “Long PCR method” for site directed mutagenesis.

primer 1

primer 2

template DNA (double-stranded) amplified in E.coli (methylated)

X: mutation to be introduced

PCR amplification of whole plasmid sequence (long PCR)

Dpn I digestion (5'-G^mATC-3' restriction site)

Transformation into E. coli

Selection media (LB media containing ampicillin 100μg/ml)

Incubation at 37 °C overnight

Culture ampicillin-resistant transformants for DNA mini-prep extraction (Appendix I-4)

Selection by restriction enzyme* and/or DNA sequencing

* Only in the case alternation of bases creates or deletes a restriction enzyme site.
Appendix 3 Preparation of *E.coli* competent cell by modified rubidium chloride method.

XL1-blue single colony or glycerol stock (50 μl)

- Inoculate
- 2 ml LB media
  - Grown at 37 °C with 250 rpm shaking overnight
- Inoculate 0.5 ml into 50 ml SOB media
  - Incubate at 37 °C with 250 rpm shaking
  - Until OD$_{550}$=0.5–0.6
- Incubate on ice for 10 min
  - Centrifuge at 10k xg for 5 min at 4 °C
  - Supernatant  Precipitate
  - Discard  Resuspend in 30 ml RF1 solution
  - Incubate on ice for 2 h
  - Centrifuge at 10k xg for 5 min. at 4 °C
  - Supernatant  Precipitate
  - Discard  Resuspend in 8 ml RF2 solution
  - Incubate on ice for 15 min.
  - Portion 200 μl in each tube
  - Snap-freeze in liquid-nitrogen or ethanol with dry ice
  - Store at –80 °C

*SOB media consists of 2 % peptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl$_2$, and 10 mM MgSO$_4$.

*RF1 solution consists of 11.2 % (w/w) RbCl, 30 mM potassium acetate buffer (pH 7.5), 0.99 % (w/w) MnCl$_2$·4H$_2$O, 0.15 % (w/w) CaCl$_2$·2H$_2$O, and 15 % (w/w) glycerol. The solution is adjusted to pH 5.8 and filtered for sterilization.

*RF 2 solution consists of 10mM MOPS or MES buffer (pH 6.8), 0.12 % (w/w) RbCl, 1.1 % (w/w) CaCl$_2$·2H$_2$O, and 15 % (w/w) glycerol. The solution is adjusted to pH 6.8 and filtered for sterilization.
Appendix 4 Transformation of gene into *E. coli* competent cells.

- Thaw *E. coli* competent cell (200 µl) on ice
- Add Plasmid DNA (max. volume 20 µl)
- Incubate on ice for 10 min.
- "Heat shock" treatment
  - Incubate at 42 °C for 90 sec.
- Incubate on ice for minimum 5 min.
- Add 800 µl SOC media
- Static culture at 37 °C for 30 min.
- Plate on LB plate containing appropriate antibiotics at various concentrations
- Incubate at 37 °C overnight
- Check the colonies on plates
- Apply transformants to next step (DNA mini-prep)

SOC media consists of 2 % peptone, 0.5 % yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, and 20 mM glucose.
Appendix 5  DNA mini-prep method for plasmid DNA extraction from E. coli.

Overnight culture of E. coli carrying plasmid DNA in 5 ml LB media with appropriate antibiotics.

Collect cells by centrifugation at room temperature at 16k xg for 3 min.

Cell pellet → Supernatant (discard)

Add 200 µl Lysis buffer a and vortex well
Add 300 µl Alkaline-SDS b solution and mix gently by inverting
Add Sodium acetate buffer (pH 8.0) and mix gently by inverting
Incubate on ice for 10 min.
Centrifugation at 4 °C for 15 min at 16k xg

Supernatant → Precipitate (discard)

Add equal volume of Chloroform/IsoAmino Alcohol (24:1) (Chl/IAA)
Mix thoroughly by shaking
Centrifugation at room temperature for 10 min. at 16k xg
Upper layer → Lower layer (discard)
Add 2-3 volumes of 100 % cold ethanol
Incubate at -80 °C for 8 min.
Centrifugation at 4 °C for 15 min. at 16k xg
Precipitate → Supernatant (discard)
Add 500 µl of 70 % cold ethanol
Centrifugation at room temperature for 5 min. at 16k xg
Precipitate → Supernatant (discard)
Dry in vacuum desiccator
Resuspend the pellet in 200 µl TE buffer
Add 10 µl RNAseA (final concentration 0.1 mg/ml)
Incubate at 37 °C for 20 min.
Apply 10 µl aliquot to agarose gel electrophoresis
Add 200 µl Chl/IAA
Check DNA concentration on the gel
Repeat Chl/IAA extraction
Add 0.3 M sodium acetate buffer (pH 4.8)
Repeat ethanol precipitation
Resuspend the pellet in 10 µl sterilized water
Apply 1 µl aliquot to agarose gel electrophoresis for estimation of the concentration

a Lysis buffer contains 0.9 % glucose (w/w), 10 mM EDTA-2Na (pH 8), and 25 mM Tris-HCl (pH 8), and the solution is autoclaved and stored at 4 °C.
b Alkaline-SDS solution contains 0.8 % NaOH (w/w) from 10 N NaOH stock solution and 1 % SDS (w/w) from 10 % autoclaved SDS solution. The solution is mixed in a sterilized container and stored at room temperature.
Appendix 6  PEG precipitation method for purification of plasmid DNA.

DNA solution
↓
Add 16 μl of 4 M NaCl solution (autoclaved)
↓
Add 80 μl of 13 % PEG 8000 solution (autoclaved)
↓
Add sterilized water to make total volume of 160 μl
↓
Incubate on ice for 20-30 min.
↓
Centrifugation at 4 °C for 30 min at 16k xg
↓
Precipitate  
Supernatant (discarded)
↓
Add 500 μl of 70 % cold ethanol
↓
Centrifugation at room temperature for 5 min.
↓
Precipitate  
Supernatant (discarded)
↓
Dry in desiccator for 5 min.
↓
Add appropriate volume of sterilized water
↓
Resuspend in sterilized water
↓
Check with Electrophoresis for estimation of DNA concentration
Appendix 7  Electroporation method for *E. coli* transformation.

**Cell preparation**

*E. coli* single colony

400 ml LB media

$\text{OD}_{600}=0.5-0.8$

Incubate on ice for 30 min.

Centrifugation at 4 °C at 4000 $\times g$ for 15 min.

Precipitate Supernatant (discard)

Resuspend in 300 ml cold sterilized water

Centrifugation at 4 °C at 4000 $\times g$ for 15 min.

Precipitate Supernatant (discard)

Resuspend in 100 ml cold sterilized water

Centrifugation at 4 °C at 4000 $\times g$ for 15 min.

Precipitate Supernatant (discard)

Resuspend in 4 ml of 10 % glycerol (cold)

Centrifugation at 4 °C at 4000 $\times g$ for 15 min.

Resuspend in 2.4 ml of 10 % glycerol (cold)

Aliquot 40 μl per tube for one transformation

**Electroporation**

40 μl *E. coli* competent cell

5 pg/μl –0.5 ng/μl of DNA solution

Incubate on ice for 5 min.

Transfer to cold electroporation cuvette

Electroporation

Add 1 ml SOC media

Incubate at 37 °C for 30–60 min. (static)

Plate on LB media containing appropriate antibiotics

Incubate overnight to form colonies
Appendix 8 Genomic DNA extraction from yeast.

5 ml YPD media

↓

Single colony of yeast

Incubate at 30 °C at 250 rpm

↓

OD_{600}=2-3

Centrifugation at RT\(^a\) at 1200 \times g for 5 min.

↓

Cell pellet

Supernatant (discard)

Resuspend in 0.5 ml sterilized water

Centrifugation (pulse) for 5 sec.

↓

Cell pellet

Supernatant (discard)

Resuspend in 0.2 ml Breaking buffer

Add 0.3 g acid washed glass beads and 0.2 ml PCI\(^b\)

Vortex for 3 min.

Add 0.2 ml TE buffer

Vortex for 5 sec.

Centrifugation at RT at 16k \times g for 5 min.

↓

Supernatant

Precipitate (discard)

Add 1 ml 100 % ethanol and mix by inversion

Add 30 \mu l of 1 mg/ml RNaseA

Incubate at 37 °C for 5–30 min.

Add 400 \mu l PCI and mix by inversion

Centrifugation at RT at 16k \times g for 5 min.

↓

Supernatant

Precipitate (discard)

Add 10 \mu l of 4 M ammonium acetate and 1 ml 100 % ethanol and mix by inversion

Centrifugation at RT at 16k \times g for 3 min.

↓

Precipitate

Supernatant (discard)

Wash with 1 ml 70 % ethanol (twice)

Dry in vacuum desiccator

Resuspend in 100 \mu l sterilized water slowly

\(^a\)RT: room temperature, \(^b\)PCI=phenol: chloroform: isoamyl alcohol (25:24:1) (vol/vol/vol)