A Study on the Effects of the N-Terminal Amino Acid Sequence on the Activation of Human T-Cell Leukemia Virus Type 1 Protease

Hidayah Muhammad Kendall

Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/chemistry_etds

Part of the Biochemistry Commons, Oncology Commons, Virology Commons, and the Virus Diseases Commons

Recommended Citation
Kendall, Hidayah M.. "A Study on the Effects of the N-Terminal Amino Acid Sequence on the Activation of Human T-Cell Leukemia Virus Type 1 Protease" (1999). Master of Science (MS), Thesis, Chemistry & Biochemistry, Old Dominion University, DOI: 10.25777/dnbv-cs29
https://digitalcommons.odu.edu/chemistry_etds/121

This Thesis is brought to you for free and open access by the Chemistry & Biochemistry at ODU Digital Commons. It has been accepted for inclusion in Chemistry & Biochemistry Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.
A STUDY ON THE EFFECTS OF THE N-TERMINAL AMINO ACID SEQUENCE ON THE ACTIVATION OF HUMAN T-CELL LEUKEMIA VIRUS TYPE 1 PROTEASE

by

Hidayah Muhammad Kendall
B.S. August 1997, Old Dominion University
A.S. May 1994, Tidewater Community College

A Thesis Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirement for the Degree of

MASTER OF SCIENCE

CHEMISTRY

OLD DOMINION UNIVERSITY
December 1999

Approved by:

Laura K. Moen (Director)

Mark S. Elliott (Member)

James H. Yuan (Member)
ABSTRACT

A STUDY ON THE EFFECTS OF THE N-TERMINAL AMINO ACID SEQUENCE ON THE ACTIVATION OF HUMAN T-CELL LEUKEMIA VIRUS TYPE 1 PROTEASE

Hidayah Muhammad Kendall
Old Dominion University, 1999
Director: Dr. Laura K. Moen

Human T-cell leukemia virus type 1 (HTLV-1) is dependent upon the enzymatic activity of its protease for maturation. Maturation of the protease is facilitated by cleavage of specific amino acid residues, followed by dimerization. The effects of the amino acid sequence located N-terminally to the cleavage site on the ability of the protease to become active were the focus of the current study. These amino acid sequences were contributed by the plasmid vector into which the protease gene was inserted.

Surface probability analyses (SPAs) of the vectors, as well as for native sequences which produce the mature protease and other viral proteins indicated that a specific surface topology may be required for cleavage site recognition by the protease. Using this information, two plasmid vectors, pET15b and pET19b, were selected for use. SPA of the pET15b vector indicated that amino acid sequence located N-terminally to the cleavage site was hydrophobic. In the pET19b vector, the sequence was hydrophilic. Three clones of the protease gene were synthesized and inserted into each vector, leading to six protease constructs.

The protease gene was expressed, purified, and activated in each of the constructs. The enzyme was activated against two different buffer systems. One buffer contained sodium citrate, an anti-chaotroph. The other contained 2-(4-morpholino)ethanesulfonic acid (MES). The use of two different buffer systems provided the opportunity to assess the effects of buffers on the activation process. Fluorometry was used to evaluate enzymatic activity.
Results of this study indicated that both the N-terminal amino acid sequence and the composition of the buffer against which the enzyme was activated may affect the level of enzyme activation in specific constructs. Protease constructs composed of a protease gene inserted into the pET15b vector exhibited two to four times more enzymatic activity than the same gene inserted into the pET19b vector. The use of sodium citrate to promote activation resulted in higher activity levels than when MES buffer was used for activation. The increased level of activation when citrate buffer was used ranged from approximately 1¼ times to almost three times the level of activity seen when MES buffer was used.
This thesis is dedicated to Elaine Lucas, without whom this project and degree would have remained a dream, and Kenny “Munchkin-man” Lucas, who is just a joy
ACKNOWLEDGMENTS

All praise and glory is due to Allah, sub hana wa ta’ala. It is due to Him, and none other. I testify and state that there is no deity worthy of any type of worship but Allah, and that Muhammad, sallalaahu ‘alayhe wa sallam, is the servant and final messenger of Allah.

There is a hadith of the Prophet Muhammad, sallalaahu ‘alayhe wa sallam, which says, “If a person is not grateful to people, then he is not grateful to Allah.” The completion of this degree is the culmination of many years of hard work, sleepless nights, and incredibly high frustration levels. Although it would be nice to be able to take full credit for its accomplishment, the truth is that it would not have been possible without the wonderful network of supporting family, friends, co-workers, and colleagues who have provided me with emotional support, hot meals, and laughter over the years. I hardly feel that this acknowledgments section can truly convey my appreciation to each of you.

First, to my family. Thanks to my mother, Dorothy Kendall, who taught me to overcome obstacles while maintaining a smile. Your support during this course of study has been invaluable, and I appreciate the trips to the airport, last-minute errands, and meals we’ve shared. Thanks to my dad, David Couchman, and stepmother, Melanie Noble-Couchman, who helped me with my undergraduate degree. My brother and sister, Dan and Aviva; I am thankful to have you both in my life. And, my husband, Saeed, your contributions have been many and varied, but I especially appreciate your ability to make me laugh.

My grandparents have also contributed to this degree, each in different ways. My Poppop, Charles Kendall, who encouraged my love of books and reading when I was small; my Grandma, Margaret Kendall, who is ever-present with tireless love, support, and kindness; my Grandmother, Naomi Olingę, who has been one of my most ardent supporters; and my Grandfather, Charles Couchman, II, the first person in my family with a degree in chemistry.
Lastly, thanks to my Uncle Chuck, who writes the most elegantly beautiful letters of encouragement.

During both my undergraduate and graduate careers, I have had the opportunity to work with an amazing group of people, three of whom comprise my advisory committee. Dr. Moen, you have been an incredible advisor, not only during my current project, but also while I was an undergraduate. Your guidance and counsel have been greatly appreciated; thank you for taking the time to teach me. Dr. Yuan, thank you for giving me my first “real-life” laboratory experiences and for believing in my ability to achieve. That has meant more to me than you can imagine. Dr. Elliott, you always push me to learn more and to consider things in ways I haven’t before. I’m not sure you realize that, and it’s been a truly wonderful gift; thank you. Many thanks to each of you for your input, guidance, and influence.

Many thanks to my co-workers, Pinky Agbuya, Vivek Anatharaman, Barb Conyers, and Steve Scherer, each of whom has made valuable contributions to the success of my project.

Special notes of appreciation need to be given to Sultan Al-Saadi, Sa’ad Al-Sayari, Henri Parson, and Steve Scherer. Thank you, Sultan, for loaning me your car during the last couple of weeks when I was trying desperately to get this written, and mine had been flooded as a consequence of a very friendly visit from a hurricane known as Irene. Thank you, Sa’ad, for making sure that Sultan got everywhere he needed to be while I had his car. May Allah reward you both. Thank you, Henri, for helping me focus, get my work done, and for just being an extraordinary woman, friend, co-worker, and colleague. Your contributions to this project and my studies have been more extensive than I am sure you can even imagine. Thank you to both Henri and Steve, for taking the time one afternoon to help me pull about fifty journals so that I could put together my reference section.

And, with that said, enjoy.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
</tbody>
</table>

## Chapter

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>THE RETROVIRAL LIFE CYCLE</td>
<td>3</td>
</tr>
<tr>
<td>THE GENOME OF HTLV-1</td>
<td>5</td>
</tr>
<tr>
<td>HTLV-1 PROTEASE</td>
<td>6</td>
</tr>
<tr>
<td>PREVIOUS RESEARCH</td>
<td>9</td>
</tr>
<tr>
<td>THE DEVELOPMENT OF AN EXPERIMENTAL METHODOLOGY</td>
<td>13</td>
</tr>
<tr>
<td>PROJECT OVERVIEW</td>
<td>16</td>
</tr>
<tr>
<td>II. MATERIALS AND METHODS</td>
<td>17</td>
</tr>
<tr>
<td>PREPARATION OF COMPETENT CELLS</td>
<td>17</td>
</tr>
<tr>
<td>PREPARATION OF PLASMIDS</td>
<td>18</td>
</tr>
<tr>
<td>EXPRESSION OF PROTEASE GENE</td>
<td>19</td>
</tr>
<tr>
<td>PURIFICATION OF PROTEASE</td>
<td>27</td>
</tr>
<tr>
<td>ACTIVATION OF PROTEASE</td>
<td>30</td>
</tr>
<tr>
<td>ACTIVITY ASSAYS USING FLUOROMETRY</td>
<td>32</td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>38</td>
</tr>
<tr>
<td>EXPRESSION OF PROTEASE GENE AT 25°C AND 37°C</td>
<td>38</td>
</tr>
<tr>
<td>RESULTS OF PURIFICATION</td>
<td>43</td>
</tr>
<tr>
<td>RESULTS OF PROTEASE ACTIVATION</td>
<td>45</td>
</tr>
<tr>
<td>RESULTS OF ACTIVITY ASSAYS</td>
<td>45</td>
</tr>
<tr>
<td>IV. DISCUSSION OF RESULTS</td>
<td>59</td>
</tr>
<tr>
<td>V. CONCLUSIONS</td>
<td>67</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>70</td>
</tr>
</tbody>
</table>

## APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. POSITIVE CONTROL FOR THE ACTIVITY ASSAY USING CITRATE BUFFER</td>
<td>74</td>
</tr>
<tr>
<td>B. POSITIVE CONTROL FOR THE ACTIVITY ASSAY USING MES BUFFER</td>
<td>75</td>
</tr>
<tr>
<td>C. NEGATIVE CONTROL FOR THE ACTIVITY ASSAY USING CITRATE BUFFER</td>
<td>76</td>
</tr>
</tbody>
</table>
APPENDICES, CONTINUED

D. NEGATIVE CONTROL FOR THE ACTIVITY ASSAY USING MES BUFFER ................................................................. 77
E. SURFACE PROBABILITY ANALYSES FOR THE IKEDA (IK) PROTEASE CONSTRUCTS ..................................................... 78
F. SURFACE PROBABILITY ANALYSES FOR THE NDE PROTEASE CONSTRUCTS ................................................................. 79
G. SURFACE PROBABILITY ANALYSES FOR THE OPT PROTEASE CONSTRUCTS ................................................................. 80

VITA .................................................................................................................................. 81
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Dilutions of Samples for 14-18 Hour Inductions</td>
<td>22</td>
</tr>
<tr>
<td>2. Experimental Parameters for Activity Assay on Activated Protease</td>
<td>33</td>
</tr>
<tr>
<td>3. Volumes of Bovine Serum Albumin (BSA), MilliQ Water, and Bradford Reagent Used to Generate a Standard Curve</td>
<td>35</td>
</tr>
<tr>
<td>4. Comparison of Activity Levels of Protease Produced by Constructs</td>
<td>56</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>General Mechanism of Retroviral Genome Replication</td>
<td>2</td>
</tr>
<tr>
<td>2.</td>
<td>Mechanism of Retroviral Proteases</td>
<td>8</td>
</tr>
<tr>
<td>3.</td>
<td>Surface Probability Analyses for HTLV-1 Native Cleavage Sequences</td>
<td>11</td>
</tr>
<tr>
<td>4.</td>
<td>Surface Probability Analyses of Protease Constructs Used in Previous Research Experiments</td>
<td>12</td>
</tr>
<tr>
<td>5.</td>
<td>Surface Probability Analyses of the N-Terminal Amino Acid Sequences of Plasmid Vectors pET15b and pET19b</td>
<td>15</td>
</tr>
<tr>
<td>6.</td>
<td>Gel Analysis of Whole Cell Samples from Inductions at 37°C</td>
<td>39</td>
</tr>
<tr>
<td>7.</td>
<td>Gel Analysis of Soluble and Insoluble Fractions from Samples Induced at 37°C</td>
<td>40</td>
</tr>
<tr>
<td>8.</td>
<td>Gel Analysis of Whole Cell Samples from Inductions at 25°C</td>
<td>41</td>
</tr>
<tr>
<td>9.</td>
<td>Gel Analysis of Soluble and Insoluble Fractions from Samples Induced at 25°C</td>
<td>42</td>
</tr>
<tr>
<td>10.</td>
<td>Evaluation of Purification Protocol Using Gel Electrophoresis</td>
<td>44</td>
</tr>
<tr>
<td>11.</td>
<td>Evaluation of Ikl5 Protease Autoprocessing Using Gel Electrophoresis</td>
<td>46</td>
</tr>
<tr>
<td>12.</td>
<td>Activity Assay of Protease Produced from the Ikl5 Construct and Activated Against Citrate Buffer</td>
<td>47</td>
</tr>
<tr>
<td>13.</td>
<td>Activity Assay of Protease Produced from the Ikl5 Construct and Activated Against MES Buffer</td>
<td>48</td>
</tr>
<tr>
<td>14.</td>
<td>Activity Assay of Protease Produced from the Ikl9 Construct and Activated Against Citrate Buffer</td>
<td>49</td>
</tr>
<tr>
<td>15.</td>
<td>Activity Assay of Protease Produced from the Ikl9 Construct and Activated Against MES Buffer</td>
<td>50</td>
</tr>
<tr>
<td>16.</td>
<td>Activity Assay of Protease Produced from the Nde15 Construct and Activated Against Citrate Buffer</td>
<td>51</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>17.</td>
<td>Activity Assay of Protease Produced from the Nde15 Construct and Activated Against MES Buffer</td>
<td>52</td>
</tr>
<tr>
<td>18.</td>
<td>Activity Assay of Protease Produced from the Nde19 Construct and Activated Against Citrate Buffer</td>
<td>53</td>
</tr>
<tr>
<td>19.</td>
<td>Activity Assay of Protease Produced from the Nde19 Construct and Activated Against MES Buffer</td>
<td>54</td>
</tr>
<tr>
<td>20.</td>
<td>Positive Control for the Activity Assay Using Citrate Buffer</td>
<td>74</td>
</tr>
<tr>
<td>21.</td>
<td>Positive Control for the Activity Assay Using MES Buffer</td>
<td>75</td>
</tr>
<tr>
<td>22.</td>
<td>Negative Control for the Activity Assay Using Citrate Buffer</td>
<td>76</td>
</tr>
<tr>
<td>23.</td>
<td>Negative Control for the Activity Assay Using MES Buffer</td>
<td>77</td>
</tr>
<tr>
<td>24.</td>
<td>Surface Probability Analyses for the Ikeda (Ik) Protease Constructs</td>
<td>78</td>
</tr>
<tr>
<td>25.</td>
<td>Surface Probability Analyses for the Nde Protease Constructs</td>
<td>79</td>
</tr>
<tr>
<td>26.</td>
<td>Surface Probability Analyses for the Opt Protease Constructs</td>
<td>80</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

In 1980, the first isolates of human T-cell leukemia virus (HTLV-1) were reported by Poiesz et al. It was isolated from a 28-year-old man who had been diagnosed with a rare T-cell lymphoma (33, 39) which eventually became fatal. Since this initial report, HTLV-1 has been isolated from patients diagnosed with Sézary T-cell leukemia (32) and tropical spastic paraparesis (39). A coincident infection with HTLV-1 has also been shown in patients diagnosed with multiple sclerosis (22). In addition, HTLV-1 has been linked to the development of adult T-cell leukemia (20), HTLV-1 myelopathy (20), uveitis (6), arthritis (6), and infective dermatitis (6). Individuals infected with HTLV-1 also have an increased risk of developing opportunistic infections (39). To date, no cure for HTLV-1 infection or HTLV-1-induced disease has been discovered.

Although HTLV-1 has not been studied extensively, several of its characteristics have been elucidated. HTLV-1 is a retrovirus, and its entire genome is contained within a dimer complex composed of identical, single-stranded ribonucleic acid (ssRNA) molecules (40). As a retrovirus, it is completely dependent on the genetic machinery of its host cell for the production of proteins. Protein synthesis of retroviruses is facilitated by the conversion of ssRNA to single-stranded deoxyribonucleic acid (ssDNA) through the formation of a double-stranded hybrid containing one strand of RNA and one strand of DNA. Upon degradation of the RNA strand by the enzyme, RNase H, the remaining ssDNA acts as a template to which a complementary strand of DNA is added, forming a double-stranded DNA (dsDNA) molecule. Conversion to dsDNA is followed by integration into the DNA of the host cell, forming the provirus (18). Once formed, the provirus directs the synthesis of retroviral proteins through the usual transcription process by which DNA is transcribed into RNA. The newly transcribed RNA then directs the translation of the transcribed sequence into proteins. Figure 1 illustrates this process.

The model for this thesis was the Journal of Virology.
FIG. 1. General Mechanism of Retroviral Genome Replication. (A) Single-stranded RNA (ssRNA) of retrovirus upon entry into host cell. The entire genome of the retrovirus is located on this ssRNA. Transcription of the ssRNA begins (hatched marks) in a 5' to 3' direction. (B) As transcription continues, Rnase H degrades the ssRNA template (dashed line), leaving only single-stranded DNA. (C) Transcription of a complimentary DNA strand occurs in the 3' to 5' direction. (D) Upon completion, double-stranded DNA results, forming the provirus. (E) The provirus is integrated into the DNA of the host cell (blackened areas).
does not occur spontaneously. Rather, these reactions are catalyzed by two enzymes, reverse transcriptase and integrase (39). The reverse transcriptase catalyzes the formation of the RNA-DNA hybrid, while integration of the dsDNA is catalyzed by the integrase. But, these are not the only two enzymes involved in the life cycle of HTLV-1. Another enzyme, the protease, plays an important role in the maturation of the virus (18). In fact, without the activity of the protease, the enzymatic activity of the reverse transcriptase and the integrase is never established (36).

In the absence of its protease, the retroviral particles produced upon transcription and translation of the retroviral genome are defective and incapable of inducing continued infection within the host (15, 21, 27, 30, 36). Thus, diminishing the activity of the protease leads to a decrease in the likelihood of continued infection, which makes the protease an attractive candidate for protease-inhibiting drug therapies (7, 39). However, before adequate compounds can be designed to inhibit protease activity, the enzyme itself must first be characterized. Part of the process of enzyme characterization requires that a method be devised which will provide researchers with adequate amounts of purified, activated enzyme. The development of expression and activation parameters, and the assessment of activity for several clones of the HTLV-1 protease are the focus of the current project.

THE RETROVIRAL LIFE CYCLE

The genetic information for a retrovirus is contained within genomic RNA and housed within a capsid envelope composed of various lipids, glycoproteins, and polypeptides (40). Introduction of the virion into the cytoplasm of the host cell is initiated by the binding of the envelope to the plasma membrane of the soon-to-be infected host cell. Upon binding, the virion is absorbed by the host cell. Absorption of the virion by the host leads to significant degradation of the envelope, thus exposing the ssRNA to the cytoplasm of the host environment. At this point, the reverse transcriptase and integrase enzymes quickly go to work converting the ssRNA of the retrovirus into dsDNA, which is then integrated into the DNA of the host cell. Integration of the retroviral dsDNA is irreversible, leading to the permanent synthesis of the proteins the retrovirus needs to survive and infect other cells. Once integrated, the provirus is formed,
and synthesis of retroviral proteins begins (40).

Transcription of the provirus and translation of the transcribed sequence results in the production of retroviral proteins. These proteins are expressed as polyprotein precursors, which, upon maturation, will form viral core proteins, the replication enzymes, protease, reverse transcriptase, and integrase, and envelope proteins, as well as other proteins involved in disease progression (36, 40, 41). These other proteins appear to be involved in host cell mutagenesis (40). Maturation is achieved when specific amino acid sequences are enzymatically cleaved, yielding active proteins (40). The enzyme responsible for promoting maturation of the precursor proteins is the protease (2, 37).

In HTLV-1, the precursor proteins for the viral core proteins, envelope proteins, and the replication enzymes are targeted to the plasma membrane of the host cell and an immature capsid shell is formed (1). Once assembled, the capsid buds from the membrane (1), and the precursor proteins are protolytically cleaved to yield mature proteins, which, upon proper assemblage, form the mature virion (36).

Prior to the proteolytic cleavage of the precursor proteins found within the immature virion, the protease is activated. It is unclear whether this even occurs before, during, or after the budding of the virion from the plasma membrane (1, 13). Although the timeline for this process is not completely clear (1), it is certain that two events need to occur to promote protease activity. First, the protease, which is expressed in a precursor form, must cleave the sequence of amino acids located N-terminally to the sequence encoding for the mature protease. This is commonly referred to as autoprocessing and is autocatalytic in nature (13, 26, 34, 41). Second, dimerization of the protease is required for activity (15, 37). The requirement of dimerization for activity is a common one among retroviral proteases, although all of the factors promoting dimerization have not been clearly established. However, it appears that dimerization may be related to enzyme concentration (1, 37) and the pH of the environment (13), which would explain why the protease is not activated in the cytoplasm of the cell. Prior to aggregation at the plasma membrane, the concentration of protease throughout the cytoplasm is too dilute to promote dimerization. However, once at the plasma
membrane, sufficient concentrations of protease are present to induce the formation of active, dimerized enzyme. Additionally, retroviral proteases are maximally active between pH 4.5-6.5 (19). The pH of the cytoplasm is generally higher than the pH required for maximal protease activity (13) which would have an adverse influence on the level of enzymatic activity.

Once matured, the virion is prepared to infect a new cell, and the process of retroviral replication and maturation is repeated. At the same time new cells are being infected, new virions continue to be produced from the originally infected cells (40).

THE GENOME OF HTLV-1

The genome of all retroviruses contains at least three genes, which are arranged 5' to 3' in the following order: 5'-gag-pol-env-3' (2). In HTLV-1, two additional genes are present. According to Ding et al, a gene encoding for the protease, known as pro, is located 3' to the gag gene (7). The other gene, known as pX, is located 3' to the env gene (11, 25). Thus, the genome of HTLV-1 is arranged as follows: 5'-gag-pro-pol-env-pX-3'. Open reading frames (ORFs) for the synthesis of gag, pro, pol, env and pX proteins are present within each gene (7, 24).

The proteins encoded by the first three genes of the genome are required for the replication and continued viability of the retrovirus. The gag gene encodes the gag proteins, p15, p19 and p24, which are viral core proteins, as well as matrix, capsid, and nucleocapsid proteins (9, 36). The pol gene encodes the reverse transcriptase and integrase enzymes (24). The env gene encodes the capsid proteins (9, 14), which assist in the mediation of virion binding and facilitate entry of the virion into the host cell.

Of these three genes, only the gag and env genes have the capacity to be expressed as independent gene products. In fact, the products of transcription and translation of the env gene are expressed exclusively as independent products. Transcription of the gag, pro, and pol genes are more complicated.

The gag gene can be expressed in three possible forms, independently, resulting
in the production of only gag proteins, or in conjunction with the pro and pol genes, resulting in the production of the fusion proteins, gag-pro or gag-pro-pol (24). If only the ORF for the gag proteins is transcribed, then only gag proteins are produced. Gag-pro protein is produced when both the gag and the pro genes are transcribed and translated. Transcription of the pro gene occurs when the amber termination codon located at the end of the gag gene is suppressed (38), or when the gene is transcribed by means of a frameshift mechanism (16, 17, 24). Upon cleavage of the primary translational product which results, both gag proteins and mature protease are produced.

Terminator codon suppression and frameshifting are also involved in the regulation of transcription of the pol gene, leading to the production of gag-pol and gag-pro-pol proteins (24). As with the gag and gag-pro proteins, cleavage of the primary translational products results in the production of mature enzymatic proteins.

Obviously, the gag gene can be expressed in many different forms. Although it is not known why this occurs, it has been hypothesized that this ability may be related to the increased need of the mature virion for the proteins encoded by the gag gene to ensure the continued viability of the retrovirus (2).

The role, as well as the translation products, of the terminal gene of the HTLV-1 genome, pX, has not been established. However, genes located in this region of the genome in other retroviruses have been implicated in the transformation of infected cells and the induction of disease within infected organisms (40). The role the pX gene may play in disease progression remains to be evaluated.

**HTLV-1 PROTEASE**

Although a tremendous amount of research has been performed on the protease of other retroviruses, research on the protease of HTLV-1 has been minimal. The lack of research is primarily due to the difficulties associated with obtaining active enzyme. However, some aspects of HTLV-1 protease have been very clearly established.

Like other retroviral proteases (2), the protease of HTLV-1 is an aspartic acid
protease. HTLV-1 protease functions as a dimer, composed of two identical monomeric subunits. Each subunit contains a highly conserved triplet sequence of Asp-Thr-Gly which form the active site for the enzyme (8, 21, 29). Dimerization is essential for enzymatic activity (4). All of the factors influencing dimerization have not been established, although there seems to be a link between pH, enzyme concentration, and dimerization. The HTLV-1 protease dimers are stabilized by a four-stranded, antiparallel, β sheet interaction between the two subunits of the protease. Experimental evidence using X-ray crystallography indicates that the N- and C-terminal regions of each monomeric unit, as well as conserved amino acid sequences found within the protease are primarily responsible for the stabilization of the dimer (10). The importance of dimerization in the activation of HTLV-1 protease cannot be underestimated, and, in fact, destabilization of the dimer is currently being evaluated for its ability to inhibit protease activity (3).

The Mechanism of HTLV-1

A formal mechanism specifically for the protease of HTLV-1 has not been established. However, all of the aspartic acid proteases share the same sequence of amino acid residues in their active sites, and a general mechanism of action has been proposed for this class of enzymes (12). Figure 2 illustrates the mechanism of action for the aspartic acid proteases.

Each monomeric unit of the HTLV-1 protease contains an active site containing an aspartic acid residue. During the activation of protease, individual monomers combine, yielding a dimer containing two active sites, each with one aspartic acid residue. The aspartic acid residues are identified in Figure 2 as Asp and Asp’. In the free state, the two residues share a proton and a water molecule, with the oxygen of the water molecule bearing a negative charge. This negative charge participates in the nucleophilic attack of the carbonyl carbon present in the substrate. The carbonyl oxygen, with its slight negative charge, then interacts with the proton contributed by the water molecule found on the Asp residue, forming an asymmetrical transition state intermediate. In the intermediate, the carbonyl carbon adopts an oxyanion form due to the nucleophilic attack of the water and the interaction between the proton on Asp’. Once this intermediate is
FIG. 2. Mechanism of Retroviral Proteases. (A) Enzyme in the free state. In the free state, the two aspartic acid residues in the active site of the enzyme share a water molecule and a proton. There is a slightly negative charge on the oxygen atom of the water molecule. (B) Enzyme-substrate complex. In the presence of substrate, the oxygen participates in a nucleophilic attack on the carbonyl carbon, which has a slightly positive charge. The carbonyl oxygen carries a slightly negative charge and is free to interact with the proton, H_1, located on Asp. These interactions result in the formation of an asymmetrical transition state intermediate. (C) Transition state intermediate. In the intermediate, the oxygen adopts an oxyanion form and the proton, H_3, is donated to the amino group of the substrate. (D) Formation of products. Donation of H_3 to the amino group results in the formation of two catalysis products. Upon the release of the two products, the free enzyme is regenerated (A).
formed, the other proton of the water molecule, which is located on the Asp' residue is donated to the amino group of the substrate. Upon donation, the products of catalysis are released, and the enzyme is regenerated in its free form.

PREVIOUS RESEARCH

This project was developed in response to previous work completed in two different laboratory environments.

In 1998, it was reported by Ding et al. (6) that a construct of the HTLV-1 protease gene inserted into a pET19b vector at the BamHI site had been made, and the gene had been expressed in an insoluble form in E. coli BL21(DE3)pLysS cells at 37°C. The resulting protease was purified and activated. The protease, as it was expressed, was fused to a molecular tag consisting of a sequence of ten histidine residues, and, upon denaturation, the protease was purified using a His-Bind column from Novagen. The protease was refolded using sodium acetate and sodium citrate buffers. Incubation of the protease with substrate resulted in the cleavage of substrate, indicating that the protease was active. According to this report, three mg of protease per liter of culture was obtained.

At about the same time, Moen et al. inserted a clone of the protease gene into the plasmid vector, pET29. The protease was inserted into the plasmid at the EcoRV site, and the protease expressed from this plasmid contained a molecular tag known as an “S-tag” (S-tag construct). In this vector, the protease was expressed in a soluble form at 25°C. However, the protease could not be activated.

The sequence for the protease gene was the same in each of these experiments; yet, in one case, active protease was produced, while in the other, the protease was inactive. This led to the hypothesis that the plasmid vector into which the protease gene was inserted may exert an influence over whether or not active protease could be obtained.

With this hypothesis in mind, analyses of the constructs used in the previous
research by Ding et al and Moen et al were begun. The most significant difference between these two constructs was the sequence of amino acids, which, upon expression of the protease, would be located N-terminally to the sequence of amino acids which would require cleavage to produce the autoprocessed, active protease. The sequence of amino acids was contributed primarily by the plasmid into which the protease gene was inserted. This observation led to a new, refined hypothesis that the sequence of amino acids contributed by the plasmid vector into which the protease gene was inserted may influence the ability of the protease to autoprocess. To evaluate this new, modified hypothesis, surface probability analyses (SPAs) were performed on the amino acid sequences located N-terminally to the protease gene in the pET19b and pET29 vectors. SPAs were also conducted on native cleavage sequences for the protease, notably, the sequences which, when cleaved, lead to the production of mature protease, as well as the viral core proteins, p15 and p24, and p19 and p24.

The results of these analyses on the native cleavage sites are presented in Figure 3. Figure 3A is the SPA for the sequence, which, upon cleavage, produces the p15 and p24 proteins. In this SPA, the amino acid sequence located N-terminally to the protease recognition (underlined in the figure) sequence is likely to be found on the surface of the protein, while the sequence of amino acids at which cleavage occurs, VL/VV, is more hydrophobic. The hydrophobicity of the cleavage sequence is evidenced by its lower surface probability. The cleavage site is followed by another sequence of amino acids possessing a more hydrophilic character, as evidenced by its higher degree of surface probability. The same general pattern was seen in Figure 3B and 3C, in which the SPA of the amino acid sequences which, upon cleavage, produced the viral proteins, p19 and p24, and the mature protease, are presented. However, in these two figures, the hydrophilic character of the amino acids located N-terminally to the protease is much less pronounced.

In the work reported by Ding et al (6), the protease cleavage site had the sequence ED/PV. The SPA for this construct is presented in Figure 4A. SPA of the pET19b vector containing the protease gene in this construct displayed some striking similarities between the surface probabilities of the pET19 vector and the surface probabilities of the
FIG. 3. Surface Probability Analyses for HTLV-1 Native Cleavage Sequences. The underlined amino acid sequences comprise the sequence of amino acids recognized by the protease for cleavage. (A) Cleavage at the leucine and valine residues (residues 14 and 15) of the underlined amino acid sequence produces the viral core proteins, p15 and p24. (B) Cleavage at the proline and valine residues (residues 20 and 21) of the underlined amino acid sequence produces the viral core proteins, p19 and p24. (C) Cleavage at the leucine and proline residues (residues 17 and 18) of the underlined amino acid sequence produces the mature protease.
FIG. 4. Surface Probability Analyses of Protease Constructs Used in Previous Research Experiments. The underlined amino acid sequences comprise the sequence of amino acids recognized by the protease for cleavage. (A) Surface probability analysis of protease construct used by Ding et al, consisting of a clone of the HTLV-1 protease gene inserted into a pET19b plasmid vector at the BamHI site. Cleavage occurs between aspartic acid and proline residues (residues 20 and 21) of the underlined amino acid sequence. (B) Surface probability analysis of protease construct used by Moen et al, consisting of a clone of the HTLV-1 protease gene inserted into a pET29 plasmid vector at the EcoRV site. In this construct, cleavage occurs between the leucine and proline residues (residues 17 and 18) of the underlined amino acid sequence.
native cleavage sequence. Specifically, the sequence of amino acids preceding and following the cleavage site were more likely to be found on the surface of the protein, while the sequence in which the cleavage sequence was located was not.

In contrast, there were marked differences between the SPA of the native sequence and the SPA of the S-tag construct, shown in Figure 4B. The protease cleavage sequence in this construct was AD/PV. As in the SPA of the native sequence, the cleavage sequence for the protease in the pET29 vector displayed low surface probability. However, in this vector, the amino acids located N-terminally to the protease cleavage sequence also exhibited low surface probability.

These observations, in conjunction with the experimental results obtained by both research groups, seemed to lend credence to the hypothesis that the amino acid residues contributed by the plasmid vector had an effect on the ability of the protease to autoprocess, perhaps by helping or hindering recognition of the protease cleavage sequence.

THE DEVELOPMENT OF AN EXPERIMENTAL METHODOLOGY

The goal of the current study is to evaluate expression and activation parameters for the protease of HTLV-1 produced from two different plasmid expression vectors. Using the information obtained from the SPA for the native cleavage sequence, pET19b and pET29, a methodology was developed to assess the effects of the N-terminal amino acid sequence on the ability of the protease to autoprocess.

Development of an experimental methodology involved careful consideration of all of the variables present in the previous experiments. Essentially, there were three variables. The first, and, probably, the most influential, was the different surface probabilities of the amino acid sequences located N-terminally to the protease cleavage site, which were primarily contributed by the plasmid vector into which the protease gene was inserted. The second was the type of molecular tag located on the expressed protein, also contributed by the plasmid vector. The last was the type of restriction enzyme used for the insertion of the protease gene into the plasmid vector.
Selection of Plasmid Vectors

In developing the experimental methodology, it was important to eliminate as many variables as possible. To achieve this goal, two plasmid vectors were selected for use in the experimental protocol. Each contained the same molecular tag, and the protease gene was inserted into each at the same restriction enzyme site.

The two plasmid vectors selected for use in the current study were pET15b and pET19b. The SPAs for the sequence of amino acids leading up to the insertion site for the protease gene for these two plasmids are presented in Figure 5. The pET19b vector was selected because of the success in obtaining active enzyme when the gene was expressed in this vector. The second plasmid, pET15b, was selected due to the similarities between its surface probability analysis, shown in Figure 5A, and the SPA of the S-tag construct. In both the pET15b and the S-tag construct, the sequence of amino acids leading up to the site at which the protease gene was inserted was hydrophobic. The pET15b and pET19b vectors each contained a molecular tag consisting of a sequence of histidine residues, known as a His-tag. Although the identity of the molecular tag was the same, the length of the tag was different. In the pET19b vector, the tag consisted of ten histidine residues, while, in the pET15b vector, it was six histidine residues in length. Each plasmid contained a restriction enzyme site for Nde I, and it was at this site that the protease gene was inserted.

The Development of Protease Clones

Three clones of the protease gene were produced. The actual sequence for the protease gene in each of the clones was the same. However, the sequence of amino acids comprising the protease recognition site was different in each clone.

The first clone utilizes the same sequence, LED/PVI, as published by Ding et al (6). The cleavage site for protease autoprocessing occurs between the aspartic acid and the proline residues. Since this clone was designed by the research group led by Dr. Richard Ikeda, this clone was known as the Ikeda (Ik) clone.

The second clone has the same sequence as the clone used in the pET29 study, with the
FIG. 5. Surface Probability Analysis of the N-terminal Amino Acid Sequences of Plasmid Vectors pET15b and pET19b. The sequences shown are the amino acids located N-terminally to the insertion site for the protease gene in the final constructs. The underlined sequences compose the six amino acids immediately preceding the site at which the protease gene was inserted. (A) Surface probability analysis of the pET15b vector. (B) Surface probability analysis of the pET19b vector.
exception of the addition of a Nde cleavage site, to facilitate insertion of the gene into the plasmids. This second clone uses a slightly more hydrophobic sequence of MAD/PVI. Again, the protease cleavage site for autoprocessing is between the aspartic acid and the proline residues. This clone was known as the Nde clone.

The third clone was designed with an optimized cleavage sequence. Substrate specificity studies have shown that the HTLV-1 protease favors a cleavage site at which cleavage occurs between leucine and proline residues (28, 31). This clone was constructed with a sequence KVL/PVI, with cleavage occurring between the leucine and proline residues. This clone was known as the Opt clone.

The Development of Protease Constructs
Each of the protease clones was inserted into each of the two plasmid vectors, pET15b and pET19, producing a series of six protease constructs. These were known as Ik15, Ik19, Nde15, Nde19, Opt 15, and Opt 19. The number indicates the identity of the plasmid vector into which the clone was inserted. As an example, the Ik15 construct is composed of the Ikeda clone inserted in the pET15b vector.

PROJECT OVERVIEW
The goal of the current project is to evaluate a series of protease constructs for their ability to produce protease capable of autoprocessing, resulting in enzymatic activity. By utilizing a series of constructs which differ in the amino acid sequences located N-terminally to the protease cleavage site, it may be possible to determine what factors influence the ability of the protease to autoprocess. Since one of the primary difficulties associated with HTLV-1 protease research is the difficulty which with active protease is obtained, it is hoped that this study will assist in the elucidation of the processes by which active protease is generated.
CHAPTER II
MATERIALS AND METHODS

PREPARATION OF COMPETENT CELLS

A stock of competent cells was prepared using commercially available competent cells from Novagen. Novagen #69451-4 *E. coli* BL21(DE3)pLysS competent cells were applied to a supporting media composed of 2.5% Luria Bertani (LB) broth (Difco Laboratories) and 1.5% agar (Sigma) which had been autoclaved. The autoclaved solution was cooled, and 34 µg/mL chloramphenicol (Fisher) was added to the solution. Distribution of the antibiotic throughout the solution was accomplished by swirling the flask containing the solution. The solution was poured into petri dishes (Falcon, model #1029) and solidified at room temperature. Once the supporting media had set, the cells were applied to the media using a sterilized loop. Sterilization was achieved by heating the loop in a flame until it was red-hot, then cooling it at ambient temperatures. The sample was dried at room temperature, then placed sample-side down in a 37°C incubator for 14-16 hours, resulting in the growth of numerous cell colonies.

One of the colonies was inoculated into five mL of sterile LB media containing 34 µg/mL chloramphenicol. The culture was grown for 14-16 hours at 37°C with shaking at a speed of 250 rpm. Following this initial incubation, one mL of the culture was inoculated into 100 mL sterile LB media containing a chloramphenicol concentration of 34 µg/mL. This culture was incubated at 37°C with shaking at 250 rpm until the optical density of the culture reached 0.375 at a wavelength of 595 nm.

Once an optical density of 0.375 had been achieved, the culture was aliquoted into two sterile, chilled fifty mL round bottomed centrifuge tubes and placed on ice for 5-10 minutes. Chilling was followed by centrifugation at 4°C using a Beckmann GPR centrifuge and rotor model #345371 for seven minutes at 3000 rpm. No brake was applied during deceleration. The supernatant was removed, and each pellet was resuspended in ten mL cold CaCl₂ solution by gently rubbing the tubes on ice. Following resuspension, the cells were centrifuged, without braking, at 4°C for five minutes at 2500
rpm. Again, the supernatant was discarded and the pellet was resuspended in ten mL cold CaCl₂. Once resuspension of the pellet was complete, the solution was kept on ice for thirty minutes, then centrifuged, without braking, for five minutes at 2500 rpm. The resulting supernatant was discarded, and each pellet was resuspended in two mL of cold CaCl₂. Once resuspended, the solution was placed in sterile, chilled, 2.0 mL capacity cryogenic vials (Corning) in 250 µL aliquots and stored immediately at -70°C.

PREPARATION OF PLASMIDS

The gene that encodes for the HTLV-1 protease was inserted into plasmid vectors, pET15b and pET19b. Insertion of the gene into each plasmid was performed by Ms. Barbara Conyers. Accuracy of insertion was ascertained by sequencing the DNA of the plasmid subsequent to insertion of the gene.

Plasmid DNA Preparation
Plasmid DNA into which the protease gene had been inserted was prepared by Ms. Barbara Conyers using either a BIGGER™ prep (5’ to 3’) or a Perfect Prep™ (5’ to 3’) DNA preparative kit. The DNA was prepared according to the manufacturer's directions.

Plasmid DNA lacking the protease gene was obtained from Novagen for both the pET15b and pET19b vectors for use as negative controls.

Transformation of Plasmid DNA into E. coli Competent Cells
A sterile, polystyrene tube (Falcon brand, model #2059) was placed on ice and chilled. While the tube was chilling, a vial of competent E. coli BL21(DE3)pLysS cells and a vial of DNA (of either one of the protease constructs or one of the negative control plasmid) was placed on ice and thawed.

While the competent cells and DNA thawed, two LB/agar plates were prepared for each transformation. The media was prepared using LB, agar and chloramphenicol as explained previously. In addition, 100 µg/mL ampicillin (Stratagene) was added to the solution prior to pouring the solution into the petri dishes.
Once the competent cells and DNA had thawed and the tube had chilled thoroughly, 100 µL of competent cells were aliquoted into the chilled tube, and 1 µg of DNA was added to the cells. Following the addition of DNA, the tube was kept on ice for thirty minutes. After the competent cells and DNA had incubated, the cells were heat shocked by placing the tube in a 42°C water bath for forty-five seconds. The heat shocking process allowed the \textit{E. coli} cells to become permeable to the plasmid DNA. The tube was placed on ice for two minutes, thereby enabling the cells to recover from the heat shock process. Then, 450 µL of sterile LB media was added to the tube, and the culture was incubated at 37°C with shaking at 220 rpm for 1-2 hours.

Once the incubation was complete and the LB/agar plates had set, 100 µL of transformed cells was applied to one of the LB/agar plates. The sample was spread across the supporting media using a glass spreader which had been dipped in 95% ethanol (Aaper Alcohol), then ignited until all of the ethanol on the surface had been removed, resulting in its sterilization. The remains of the incubated culture were transferred to a sterile, siliconized, 1.5 mL volume, microcentrifuge tube (Fisher) and centrifuged using a Savant Speedfuge HSC10AC at the 1000 rpm for ten minutes at room temperature. The majority of the supernatant was removed and discarded. The pellet resulting from centrifugation was gently resuspended in the residual supernatant using a pipette tip. Once resuspended, all of the solution was applied to the remaining LB/agar plate. The sample was distributed across the supporting media as previously described. The samples were applied to both plates at room temperature, then placed sample-side down in a 37°C incubator, and incubated for 14-16 hours.

\textbf{EXPRESSION OF PROTEASE GENE}

\textit{Inoculation of Culture and Initiation of Cell Growth}

The inoculation of media and initiation of cell growth was conducted in the same manner regardless of whether a control or a protease construct was induced.

Cell growth was initiated by inoculating a single colony of transformed \textit{E. coli} cells into a sterile, 100 mL culture of 2.5% LB media containing 34 µg/mL chloramphenicol and 100 µg/mL ampicillin. The culture was incubated at 37°C with shaking at 320-350 rpm. The
density of the solution was periodically monitored by removing one mL aliquots from the culture. The aliquot was placed in a disposable, one cm cuvette (Fisher), and the absorbance of the solution was determined using a CARY 3 Bio UV-Visible Spectrophotometer set at a wavelength of 600 nm which had been zeroed against a sample of sterile LB media. When the absorbance of the culture reached a value between 0.6-0.8, the culture was removed from the incubator.

**Inductions at 25°C**

Upon removal of the culture from the incubator, the culture was placed on ice for thirty minutes. During this time, the heating element of the incubator was turned off, and the incubator was cooled to room temperature.

Upon removal of the culture from the incubator, four-one mL aliquots were removed from the culture and placed in sterile, siliconized, 1.5 mL capacity, microcentrifuge tubes (Fisher). These samples were the zero hour induction samples. The cells were harvested by centrifugation at room temperature using a Savant Speedfuge HSC10AC centrifuge set at maximal speed for one minute. The supernatant was discarded, and the pellet was washed with 500 µL of 1 M Tris Hydroxymethyl Aminomethane Hydrochloride (Tris, Fisher), pH 8.0. The wash was decanted, and the pellets were stored immediately at -70°C.

Once the thirty minutes had elapsed, expression of the protease gene was induced via the addition of 126 µL of an 800 mM solution of isopropyl-β-D-thiogalactoside (IPTG, Promega) for a final concentration of 1 mM IPTG. The flask was placed back in the incubator and shaken at 320-350-rpm for 14-18 hours.

For the first six hours of the induction, four-one mL aliquots were removed from the culture at one-hour increments, then the inductions continued overnight. Another set of four-one mL samples were collected at the end of the 14-18 hour induction period. The cells in each of these samples were harvested as described previously.

In addition, the absorbance of the culture during the induction was monitored at a
wavelength of 600 nm. A zero absorbance reading against a sample of sterile LB media preceded each spectrophotometric analysis of sample. The absorbance of the solution was monitored each time aliquots of the culture were removed for harvest. As the induction proceeded, the absorbance of the samples increased, until the absorbance of the solution exceeded the maximum absorbance capacity of the spectrophotometer. To compensate for this, samples were diluted. The dilution factors for the samples used to determine absorbance of the solution are provided in Table 1. Once the final set of four-one mL samples had been procured and harvested, and the absorbance of the culture had been determined, the remaining culture was discarded.

Inductions at 37°C
Inductions were also conducted at a temperature of 37°C. Once the culture had reached the appropriate density, it was removed from the incubator, and a set of four-one mL aliquots were removed from the culture. These were the zero hour induction samples. The cells were harvested and stored as previously described. Following the removal of the samples, expression of the gene was immediately induced by the addition of 126 µL of 800 mM IPTG for a final concentration of 1 mM IPTG. The flask was placed in the incubator, and the induction continued at 37°C with shaking at 320-350 rpm.

As in the 25°C induction, four-one mL samples were taken at one-hour intervals for six consecutive hours, then the inductions continued overnight, with another set of four-one mL samples collected at the end of the 14-18 hour induction period. The cells were harvested as previously described and stored at -70°C. Absorbance of the culture was monitored using a Cary 3 Bio UV-Visible Spectrophotometer set at a wavelength of 600 nm, which had been zeroed against a sample of sterile LB media. The dilution scheme shown in Table 1 was used to prepare the samples for spectrophotometric analysis.

SDS-PAGE Analysis of 25°C and 37°C Induction Samples
Samples removed from the cultures during induction were analyzed for gene expression, as well as to assess the solubility of the expressed protein.
TABLE 1. Dilutions of Samples for 14-18 Hour Inductions

<table>
<thead>
<tr>
<th>Sample</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 hour</td>
<td>1:2</td>
</tr>
<tr>
<td>2 hour</td>
<td>1:4</td>
</tr>
<tr>
<td>3 hour</td>
<td>1:8</td>
</tr>
<tr>
<td>4 hour</td>
<td>1:10</td>
</tr>
<tr>
<td>5 hour</td>
<td>1:12</td>
</tr>
<tr>
<td>6 hour</td>
<td>1:14</td>
</tr>
<tr>
<td>14-18 hours</td>
<td>1:20</td>
</tr>
</tbody>
</table>

*aOne mL samples were removed from the cell cultures at one hour increments for six consecutive hours. One additional sample was removed from the cell cultures after the induction had continued for 14-18 hours.*

*bSamples removed from the cell cultures were diluted with sterile LB media. The dilution represents the ratio of sample to LB media in the samples analyzed.*
Preparation of 16% SDS-PAGE Gels

Throughout the development of the experimental protocols, 16% polyacrylamide gels were used for analysis. The gels were prepared in the same way for each stage of analysis and protocol development. Each gel contained approximately four mL of a 16% polyacrylamide separating gel and one mL of a 4% polyacrylamide stacking gel. Sample and running buffers containing Tricine were used throughout the analyses.

The separating gel contained 37.6% 1 M Tris, pH 8.8, 26.8% MilliQ water, 16% of a 50% acrylamide/bis-acrylamide solution (acrylamide, Fisher; bis-acrylamide, Fisher), 1% sodium dodecyl sulfate (SDS, Fisher), 2.5% ammonium persulfate (Fisher), and 0.025% N, N, N, N-Tetramethylethylenediamine (TEMED; Bio Rad Laboratories). The Tris, MilliQ water, acrylamide/bis-acrylamide solution, and SDS were placed in a beaker and mixed by swirling. The ammonium persulfate was added, followed by TEMED, and the solution was mixed by swirling. Once mixed, the solution was loaded into a gel cassette (Novex™). The polymeric mixture was overlaid by approximately one mL MilliQ water and allowed to polymerize at room temperature for 30-60 minutes. Once polymerized, the water overlay was poured off, and the stacking gel was applied.

The stacking gel contained 51.4% MilliQ water, 33.6% 0.375 M Tris, pH 6.8, 8% ammonium persulfate, 4% acrylamide/bis-acrylamide solution, and 0.04% TEMED. As in the preparation of the separating gel, the water, Tris and acrylamide/bis-acrylamide solution were mixed together, then the ammonium persulfate and TEMED were added to the solution. The stacking gel was layered onto the separating gel in the cassette. Once the stacking gel had been applied, well combs (Novex™) were inserted into the stacking gel. The stacking gel polymerized at room temperature for 30-60 minutes.

Once the polymerization of both the separating and stacking gels was complete, the gel was wrapped in plastic wrap and stored at 4°C in a hydrated environment until needed. Gel analyses were conducted using an X Cell II gel box (Novex™). To prepare the gel for use in analysis, the gel was removed from storage, and the tape located at the bottom of the gel was removed. The gel was rinsed thoroughly with deionized water and placed in the gel box. The apparatus was assembled and 800 mL of a 1X solution of Tricine
running buffer (100 mM Tris triple crystallized freebase, Fisher; 100 mM Tricine, Fisher; 0.1\% SDS; MilliQ water to appropriate volume) was poured into the apparatus. Once each of these steps was completed, gel analysis could begin.

**Evaluation of Gene Expression Using Whole Cell Samples**

To evaluate expression of the protease gene in a specific construct, a set of samples from both the 25°C and 37°C inductions were removed from -70°C storage and thawed on ice. A set of samples was comprised of one cell pellet removed from the culture prior to induction, six cell pellets representing samples removed at one hour increments during the first six hours of induction, and one cell pellet removed once the induction had continued for 14-18 hours.

Once the samples had thawed, 2X Tricine sample buffer (Novex™) containing 5\% 2-mercaptoethanol (Fisher) was added to each of the samples. The volume of sample buffer added to each sample was dependent upon the absorbance reading of the sample. For every 0.6 absorbance unit, fifty µL of sample buffer was added to the sample.

Following the addition of sample buffer, the cell pellets were resuspended in the buffer by vortexing, then the samples were sonicated using a Fisher Sonic Dismembrator 60 at a setting of three for three-one second pulses. The samples were heated at 85°C for 10 minutes, then the samples were centrifuged using a Savant Speedfuge HSC10AC for 10-15 seconds at room temperature at 10,000 rpm. The samples were stored at -20°C until needed.

The samples, as well as molecular weight markers (supplied by either Bio Rad Laboratories or Benchmark) were removed from -20°C storage and thawed at room temperature immediately prior to gel analysis. The samples were loaded into the wells of the gel using a Gilson Pipettman Model O61573L 0.5-10 µL pipette. Standard molecular weight markers were loaded into the outside lanes of the gel. The gel was run at 125 constant voltage for 90-120 minutes, or until the colored band of indicator dye had reached the bottom of the gel.
Once the gel run was complete, the gel was removed from the gel box and plastic cassette. The gel was placed in MilliQ water and washed at room temperature with shaking on a Red Rocker Model #PR50 from Hoefer Scientific Instruments shaker at a setting of three for one hour. Using Gel Code Blue Stain Reagent (Pierce), the gel was stained at room temperature with shaking for one hour. Destaining was accomplished by placing the gel back in MilliQ water and allowing the gel to wash with shaking overnight at room temperature. Photos of the gel were taken with a Hoefer’s Photoman Direct Screen camera, model DS34.

_Evaluation of Solubility of Expressed Enzyme_

To determine the solubility of the expressed protease, a set of gel analyses was undertaken using sets of samples from both the 25°C and 37°C inductions for each protease construct. As in the whole cell analyses, a sample set consisted of a cell pellet removed from the culture prior to the addition of IPTG, a series of cell pellets removed from the culture at six-one hour increments following the addition of IPTG, and a cell pellet removed from the culture once the induction had proceeded for 14-18 hours.

The cell pellets were removed from storage at -70°C and thawed on ice. Once thawed, the pellet was resuspended by vigorous vortexing in 1 M Tris, pH 8.0. The volume of Tris used in the resuspension was related to the absorbance of the culture at the time the sample was removed from the cell culture. For each 0.6 absorbance unit, fifty µL of Tris was added to the cell pellet. Following resuspension, each sample was sonicated for three-one second pulses using a Fisher Sonic Dismembrator 60 at a setting of three, then centrifuged in a Savant Speedfuge HSC10AC at room temperature and 10,000 rpm for five minutes.

Centrifugation facilitated the separation of the soluble proteins produced during induction from those expressed in an insoluble form. Soluble proteins were found in the supernatant following centrifugation, whereas the insoluble proteins were located in the cell pellet.
Fifty µL of the supernatant from each sample was removed and placed in a labeled, siliconized, 1.5 mL capacity, microcentrifuge tube (Fisher). The remaining supernatant was removed from the pellet. The pellet was washed with one mL of 1 M Tris, pH 8.0, and the pellet was resuspended by vortexing. The resuspended pellet was centrifuged at room temperature and 10,000 rpm for 5 minutes using the Savant Speedfuge HSC10AC, and the supernatant was removed. This process was repeated for a total of four washings.

To prepare the samples for gel analysis, fifty µL of 2X Tricine sample buffer (Novex) containing 5% 2-mercaptoethanol (Fisher) was added to each of the microcentrifuge tubes containing the soluble samples. A volume of 2X Tricine sample buffer containing 5% 2-mercaptoethanol equal to the volume of Tris used in the original resuspension was added to each of the washed cell pellets. Addition of sample buffer was followed by vortexing. Only a brief vortexing was needed to mix the soluble samples with the sample buffer. In contrast, a more vigorous vortexing was required to resuspend the insoluble cell pellets in the sample buffer.

The insoluble samples were sonicated for twenty-one second pulses using a Fisher Sonic Dismembrator 60 at a setting of three. Using the same sonication setting, the soluble samples were sonicated for three-one second pulses. Sonication was followed by heating at 85°C for ten minutes, then the samples were centrifuged at room temperature using the Savant Speedfuge HSC10AC for 10-15 seconds at maximum speed. The samples were stored at −20°C until used in gel analysis.

The samples and molecular weight standards markers (supplied by either Bio Rad Laboratories or Benchmark) were thawed and loaded onto a 16% polyacrylamide gel as described previously. Gel analysis, as well as the washing, staining, destaining, and photographing procedures were conducted as previously described.

**Inductions of 100 mL cultures at 25°C**

Sterile, 100 hundred mL cultures of LB media containing 34 µg/mL chloramphenicol and 100 µg/mL ampicillin were inoculated with one colony of transformed *E. coli* BL21(DE3)pLysS cells. The induction procedure was conducted as described
previously. The length of induction varied according to which construct was being induced. Some constructs were induced for six hours, whereas others were induced for 14-18 hours. However, in contrast to the previous inductions, no one mL samples were removed during these inductions.

At the end of the induction, the absorbance of the solution was determined using the Cary 3 Bio UV-Visible Spectrophotometer at a wavelength of 600 nm and a disposable, one cm cuvette (Fisher). A zero absorbance reading was obtained against a sample of LB media, and the sample was diluted 1:20 with sterile LB media prior to the determination of absorbance.

The remaining culture was transferred into a preweighed, 100 mL centrifuge bottle and centrifuged at 25°C and 10,000 x g for 10 minutes using a Beckman J2-21 centrifuge and a JA-14 rotor. The supernatant was decanted, the bottle was reweighed, and the weight of the pellet calculated. The pellet was resuspended in 500 µL of lysis buffer (50 mM Tris, pH 8.0; 300 mM NaCl, Fisher; 5 mM imidazole, Aldrich Chemical Company; MilliQ water), then transferred to a sterile, cryogenic, 2.0 mL capacity cryogenic vial (Corning). The pellet was frozen in liquid nitrogen and stored immediately at -70°C.

PURIFICATION OF PROTEASE

_E. coli_ produce a variety of proteins during induction. In order to separate the protease from the majority of the other cellular proteins, it was necessary to establish a purification protocol by which the further analysis of the protease would be possible without significant interference from other proteins in the solution. Gel electrophoresis was used to analyze the overall effectiveness of the purification protocol.

Purification Protocol

A pellet from one of the 100 mL inductions described previously was removed from storage and kept at 0°C. Five hundred µL of lysis buffer was added to each pellet. Once the lysis buffer was added, the pellet was kept on ice until thawed.

Once thawed, the pellet was transferred to a fifty mL round bottomed centrifuge tube.
The cryogenic vial originally containing the pellet was rinsed with one mL of lysis buffer, and the rinsing was added to the centrifuge tube. Using a Fisher Sonic Dismembrator 60 at a setting of three, the pellet was sonicated on ice for a total of six-ten second pulses. Each set of ten second pulses was followed by a ten second "cooling down" period. This process completed the lysis of the pellet, producing a crude lysate.

The crude lysate was centrifuged for thirty minutes at 4°C and 10,000 rpm using a Beckmann J2-21 centrifuge and a JA-20 rotor. The supernatant resulting from the centrifugation was decanted into a 14 mL polystyrene tube (Falcon, model #2057). One hundred µL of the supernatant was removed, placed in a siliconized, 1.5 mL capacity, microcentrifuge tube (Fisher) and placed on ice. One mL of Nickel-NTA resin (Qiagen) was added to the tube containing the supernatant. The supernatant/resin mixture was incubated at room temperature with shaking using a Red Rocker model #PR50 from Hoefer Scientific Instruments at a setting of three for one hour.

The incubated supernatant/resin was then loaded onto a glass column (Bio Rad Laboratories). The tube containing the supernatant/resin mixture was rinsed with one mL lysis buffer, and the rinsing was loaded onto the column. The stopcock was opened, and the flow-through was collected in a waste container. The column was then washed with two mL lysis buffer, which was collected as waste, followed by two washings with four mL wash buffer (50 mM Tris, pH 8.0; 300 mM NaCl; 20 mm imidazole; MilliQ water). These washings were also collected in the waste container. Each washing was collected prior to the application of the next wash to the column. Finally, the protease was eluted from the column using two mL of elution buffer (50 mM Tris, pH 8.0; 300 mm NaCl; 1 M imidazole; MilliQ water). Two hundred µL of the eluate fraction was placed in a siliconized, 1.5 mL capacity, microcentrifuge tube (Fisher) and stored placed on ice.

The remaining eluate was injected into a Slide-A-Lyzer® dialysis cassette (Pierce) with a molecular weight cut off of 7000 kD. Injection of the sample into the cassette was facilitated by the use of a 20½ gauge needle (Becton Dickinson) and a three mL syringe (Becton Dickinson). A Slide-A-Lyzer® buoy (Pierce) was attached to the dialysis cassette, and dialysis of the eluate against a series of buffers was begun.
Analysis of Efficiency of Purification Protocol Using Gel Electrophoresis

The efficiency of the purification process was analyzed using gel electrophoresis. Precipitation of the proteins contained in the eluate fraction was facilitated by the addition of trichloroacetic acid (TCA, Fisher) to a final volume of 10%. Once the TCA was added to the sample, the sample was vortexed briefly. The tube was capped and centrifuged for five minutes at 14,000 rpm and 4°C using a Beckmann Eppendorf 5415C centrifuge. The supernatant was discarded. The pellet resulting from the TCA precipitation procedure was washed with 500 µL 95% ethanol (Aaper Alcohol) and centrifuged at room temperature 10,000 rpm for five minutes using the Savant Speedfuge HSC10AC. The supernatant was removed, and the washing repeated. The excess ethanol was removed from the sample by placing the microcentrifuge tube containing the eluate into a Speed Vac Concentrator SVC100H. The vacuum provided by the Speed Vac, coupled with the centrifugal force provided by the rotor removed the excess liquid from the sample. The amount of time the sample spent in the Speed Vac varied. Once the sample was dry, it was removed from the Speed Vac.

The pellet resulting from the TCA precipitation was resuspended in 40 µL 2X Tricine sample buffer (Novex™) containing 5% 2-mercaptoethanol (Fisher). Ten µL of lysate was placed in a siliconized, 1.5 mL capacity, microcentrifuge tube (Fisher), and an equal volume of 2X Tricine sample buffer (Novex™) containing 5% 2-mercaptoethanol (Fisher) was added. The remainder of the lysate was stored at −20°C. Mixing of the samples and resuspension of the TCA precipitated protein pellet was accomplished using vortexing, followed by sonication for three-one second pulses using a Fisher Sonic Dismembrator 60 at a setting of three. The samples were heated for ten minutes at 85°C, followed by 10-15 seconds of centrifugation at room temperature and 1,000 rpm using the Savant Speedfuge HSC10AC. Once prepared, the samples were stored at −20°C until needed.

The samples and molecular weight standards markers (supplied by either Bio Rad Laboratories or Benchmark) were thawed and loaded onto a 16% polyacrylamide gel as described previously. Gel analysis, as well as the washing, staining, destaining, and photographing procedures were conducted as previously described.
ACTIVATION OF PROTEASE

Upon completion of the purification phase, the eluate from the purification was dialyzed at room temperature against 500 mL of buffer containing 50 mM Tris, pH 8.0, and 300 mM NaCl. Dialysis continued against this buffer overnight, approximately 18-20 hours. The buffer was replaced with fresh buffer in the same volume, and dialysis continued for another 2-4 hours. In the event that a precipitate formed when dialyzing against the first volume of buffer, 8 M urea (Fisher) was added to the second volume of buffer to facilitate resolubilization of the precipitate.

The eluate was then dialyzed at room temperature against 500 mL of buffer containing 10 mM sodium acetate, pH 3.5, 1 mM dithiothreitol (DTT, Fisher), and MilliQ water ("acetate buffer"). Dialysis continued for two hours. Then, the buffer was replaced with the same volume of fresh acetate buffer. Dialysis continued for an additional two hours.

Dialyses to promote autoprocessing of the protease were performed against two different activation buffers. Each buffer contained 5 mM ethylenediamine tetraacetic acid (EDTA, Acros Organics), 1 M NaCl (Fisher), 1 mM DTT (Fisher), and MilliQ water. In addition, one buffer contained 100 mM 2-(4-morpholino)ethanesulfonic acid (MES, Fisher) at a pH of 5.5. The other buffer contained 100 mM sodium citrate dihydrate (Mallinckrodt) at a pH of 5.3. These buffers were known as MES buffer and citrate buffer, respectively.

Once the eluate sample had dialyzed against the acetate buffer, it was dialyzed against 500 mL of either MES or citrate buffer for 12-18 hours.

Analysis of Efficiency of Activation Protocol
The efficiency of the activation protocol was analyzed using gel electrophoresis.

Samples of lysate, eluate, and activated protease were used in the analysis. Samples of protease produced from a specific construct were activated against both MES and citrate
buffers and analyzed on a gel to examine the effectiveness of enzyme activation using the different buffer systems.

Upon completion of the activation dialyses, lysate and eluate samples were removed from storage at −20°C and thawed on ice. The precipitate which formed when purified protease samples were dialyzed against either citrate or MES buffers was separated from the soluble components of the solution by centrifugation at 4°C and 14,000 rpm for twenty minutes using a Brinkmann Eppendorf 5414C centrifuge. The soluble fraction was transferred to a siliconized, 1.5 mL capacity, microcentrifuge tube (Fisher). The insoluble fraction was placed on ice.

The eluate sample and the proteins remaining in the soluble fraction were precipitated from solution using TCA, as previously described.

Ten µL of lysate was mixed with ten µL 2X Tricine sample buffer buffer (Novex™) containing 5% 2-mercaptoethanol (Fisher). The pellets resulting from the TCA precipitation of the eluate and soluble fractions were resuspended in forty µL and twenty µL of 2X Tricine sample buffer containing 5% 2-mercaptoethanol, respectively. The insoluble fraction was resuspended in sixty µL of 2X Tricine sample buffer containing 5% 2-mercaptoethanol.

The pellets were resuspended by vortexing, then sonicated. The soluble fractions were sonicated for three-one second pulses using a Fisher Sonic Dismembrator 60 at a setting of three. The insoluble fractions were sonicated for twenty-one second pulses. The samples were heated at 85°C for ten minutes, then stored at −20°C until used in gel analysis.

The samples and molecular weight standards markers (supplied by either Bio Rad Laboratories or Benchmark) were thawed and loaded onto a 16% polyacrylamide gel as described previously. Gel analysis, as well as the washing, staining, destaining, and photographing procedures were conducted as previously described.
ACTIVITY ASSAYS USING FLUOROMETRY

For the activity assays, a substrate was utilized which had the sequence, Abz-KTKVLYVQPK-(3NO2)A, with a corresponding molecular weight of 1556 g/mol. This substrate had been custom-made at the University of Michigan for use in these assays. Cleavage of the substrate by the protease occurred at the underlined leucine and valine residues. The substrate had a concentration of 15 mM in a 0.1% acetic acid solution. Immediately prior to initiation of the activity assay, a vial of the substrate solution was removed from storage at -20°C and thawed on ice.

While the substrate thawed, a 5X stock solution of assay buffer was prepared. The assay buffer contained 0.5 M MES or sodium citrate and 2.5 M NaCl. The stock solution of assay buffer was stored in a siliconized, 1.5 mL capacity, microcentrifuge tube and diluted to a 1X concentration prior to its use in activity assays. The type of assay buffer used in the activity assay was dependent on the buffer type against which activation of the protease occurred. If activation was performed by dialysis against MES buffer, then the assay buffer containing MES was the one used in the activity assay, and vice versa.

Protease activated using the established protocol was transferred from the dialysis cassette using a 20 ½ gauge needle (Becton Dickson) and a three mL syringe (Becton Dickson) to a siliconized, 1.5 mL capacity, microcentrifuge tube. Transfer was followed by vortexing to thoroughly mix the sample. Once mixed, 500 µL was removed from the microcentrifuge tube and placed in a ten mm quartz spectrophotometer cell (Starna Cells). The cell was placed in a Perkin Elmer Luminescence Spectrometer Model LS50B which had been thermostatically adjusted to 37°C using a Fisher Model 10L water bath and a Bio Rad Laboratories Econopump Model #EP1. The sample was warmed at 37°C for five minutes. Instrumental settings for the experimental conditions for the activity assay are shown in Table 2.

While the protease sample warmed in the fluorometer, 180 µL of 1X assay buffer and five µL of substrate were placed in a siliconized, 1.5 mL capacity, microcentrifuge tube and mixed together by vortexing. The assay buffer and substrate were mixed
TABLE 2. Experimental Parameters for Activity Assay on Activated Protease

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Setting&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excitation wavelength</td>
<td>320 nm</td>
</tr>
<tr>
<td>Emission wavelength</td>
<td>450 nm</td>
</tr>
<tr>
<td>Excitation slit width</td>
<td>15.0 nm</td>
</tr>
<tr>
<td>Emission slit width</td>
<td>2.5 nm</td>
</tr>
<tr>
<td>Duration of data collection</td>
<td>900 seconds</td>
</tr>
<tr>
<td>Interval of data collection</td>
<td>1 second</td>
</tr>
<tr>
<td>Delay in data acquisition</td>
<td>None</td>
</tr>
</tbody>
</table>

<sup>a</sup>The instrumental settings on the Perkin Elmer Luminescence Spectrometer Model LS50B could be changed to provide different experimental conditions. The values in this column represent the instrumental settings used to assay the activated protease for activity.
together immediately prior to use in the activity assay. During the brief period between mixing of the assay buffer and the substrate together and the actual use of the solution in the assay, the solution was kept at room temperature.

The warmed sample was removed from the cell using a pipette and mixed with the assay buffer/substrate solution by pipetting. The reaction mixture was placed back into the cell, and the increase of fluorescence over time was monitored for a period of 900 seconds.

**Determination of Protein Concentration**

The protein concentration of each of the protease samples assayed using fluorometry was determined using a Bradford protein assay. A calibration curve was generated using bovine serum albumin fraction V (BSA, Pierce) containing two mg/mL BSA. A total of five concentrations of BSA were assayed for protein content. The volumes of BSA and water used in each sample prepared for the assay are shown in Table 3.

Each sample was prepared by pipetting the appropriate volume of BSA into a borosilicate, disposable, glass tube (Fisher) using a Gilson Model O61573L 0.5-10 µL pipettor. This was followed by the addition of the specified volume of water using a Gilson Model P58265E 100-1000 µL pipettor. One mL of Coomassie Plus Protein Assay Reagent (Pierce) was added to each sample using an automatic reagent dispenser (Oxford™ Pipettor Model R). Each sample was mixed by gentle swirling. Samples used in the generation of the calibration curve were prepared and assayed in duplicate.

The absorbance of each sample was determined using the Cary 3 Bio UV-Visible Spectrophotometer set at a wavelength of 595 nm. One mL of sterilized, MilliQ water was placed in a disposable, one cm cuvette (Fisher), and the cuvette was placed in the reference cell of the spectrophotometer. One mL of sterilized, MilliQ water was placed in another disposable cuvette, placed in the sample cell of the spectrophotometer, and a zero absorbance reading was procured. The water in the sample cell was removed and discarded. Any residual water was removed by inserting the tip of a Kimwipe® (Kimberly-Clark®) into the cuvette, resulting in the absorbance of any residual liquid.
TABLE 3. Volumes of Bovine Serum Albumin (BSA), MilliQ Water, and Bradford Reagent Used to Generate a Standard Curve

<table>
<thead>
<tr>
<th>Concentration of BSA (µg)$^a$</th>
<th>Volume BSA (µL)$^b$</th>
<th>Volume MilliQ Water (µL)$^c$</th>
<th>Volume of Bradford Reagent (µL)$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>998</td>
<td>1000</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>996</td>
<td>1000</td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>992</td>
<td>1000</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>990</td>
<td>1000</td>
</tr>
</tbody>
</table>

$^a$A 2 mg/mL solution of BSA was used as the standard in the generation of a calibration curve. The concentrations of BSA used to generate the standard curve are shown here.

$^b$The values in this column represent the volume of BSA added to each of the samples used in the generation of the standard curve.

$^c$The values in this column represent the volume of MilliQ water added to each of the samples used in the generation of the standard curve.

$^d$The values in this column represent the volume of Bradford Reagent added to each of the samples used in the generation of the standard curve.
The cuvette in the reference cell was not removed from the cell until each of the samples had been analyzed.

Once the zero absorbance reading had been obtained, the samples prepared using the BSA, water, and protein assay reagent were analyzed. Each sample was mixed using a pipette, then pipetted into the disposable cuvette located in the sample cell. The absorbance of the solution was determined, then the sample was removed from the cuvette and placed in the borosilicate tube from which it was originally taken. Residual sample was removed using a Kimwipe® (Kimberly-Clark®), as previously described, and a new sample was added to the cuvette. This procedure was repeated until the absorbance of each of the ten samples prepared for use in the generation of the calibration curve had been assayed.

Using the absorbance data, a calibration curve was produced using Cricket Graph III. The average absorbance of the solutions containing 0 µg BSA was subtracted from the absorbance of each of the samples containing 4-20 µg BSA. The resulting values were entered into a spreadsheet, and a graph of µg BSA in solution vs. Absorbance of the solution was produced. Linear regression analysis of the data was performed, and the equation of the line, as well as a correlation coefficient was produced. Using the equation of the linear regression line, it was possible to calculate the protein concentrations in samples containing an unknown quantity of protein.

To assay the protein concentration in the protease samples used in the activity assay, twenty µl of protease, 980 µL of sterile, MilliQ water, and one mL of protein reagent (Pierce) were placed in a borosilicate glass tube (Fisher). The same volumes were used to prepare negative controls, which were used to evaluate the absorbance of the assay buffer. In place of the protease, each negative control contained twenty µL of the buffer against which activation took place. As in the activity assay, a 1X concentration of 5X assay buffer was used. Each set of samples and controls were prepared and analyzed in duplicate. The zero absorbance reading was obtained as previously described. Absorbance of the samples and negative controls was monitored at 595 nm. The average
absorbance reading for the controls was calculated, and this value was subtracted from the absorbance readings obtained for each of the samples.

Using the equation of the linear regression line, the amount of protein in the sample was calculated. The protein concentration in µg/µL was determined by dividing the amount of protein in the sample by the volume of sample used in the assay.

**Calculation of Specific Activity**
Before the specific activity of the protease could be calculated, the activity and the protein concentration from each sample must be calculated.

*Calculation of Activity of Protease Samples*
The activity of each protease sample was calculated using the following equation:

\[
\text{Activity} = \Delta \text{Integrated Fluorescence Units/Seconds} \quad (\text{Equation 2.1})
\]

*Calculation of Protein Concentration*
The protein concentration for each protease sample was calculated by using the absorbance of the protease sample to solve the equation of the linear regression line to yield the µg of protein contained in the sample. Then, the protein concentration in units of µg/µL was determined by using the following equation:

\[
[\text{Protein}] = \frac{\text{µg protein in protease sample}}{20 \ \text{µL}} \quad (\text{Equation 2.2})
\]

*Calculation of Specific Activity*
The specific activity of each protease sample was obtained using the following equation:

\[
\text{Specific Activity} = \frac{\text{Activity}}{[\text{Protein}]} \quad (\text{Equation 2.3})
\]

Specific activity has units of Integrated Fluorescence Units•µL/Seconds•µg.
CHAPTER III
RESULTS

EXPRESSION OF PROTEASE GENE AT 25°C AND 37°C

Each of the constructs, as well as the negative control plasmids, pET15b and pET19b, were induced at 25°C and 37°C. Results of gel analysis on whole cell samples and on the soluble and insoluble fractions of the samples are shown in Figures 6-9. Figures 6 and 7 show the results of the gel analyses when inductions were conducted at 37°C. Figures 8 and 9 show the results of the gel analyses when inductions were conducted at 25°C. The samples used in these analyses were harvested six hours after the introduction of IPTG to the cell culture.

In Figure 6, expression of the protease was evidenced by the presence of a 25 kDa band in lanes 3-5 and 7-9. The protease band was not present in the lanes containing samples of the negative control plasmid vectors. However, expression of the protease gene in the Opt19 construct was very low. The apparent molecular weight of the expressed protease was higher than anticipated; the expected weight was approximately 20 kDa. However, since gel electrophoresis provided a means by which the molecular weight of a protein could be estimated, the apparent molecular weight of a protein could vary from gel to gel. Other gels run with different markers supported a molecular weight for the protease of about 20 kDa.

Analysis of the soluble and insoluble fractions of the proteins produced by the Ikeda and Nde constructs indicated that the protease was expressed as an insoluble protein in each of the constructs when the inductions were performed at 37°C, as shown in Figure 7. The soluble and insoluble fractions of the proteins produced by the Opt19 construct were not analyzed. Analysis of the soluble and insoluble fractions of the Opt15 construct (data not shown) indicated that the protease was expressed insolubly.

In an attempt to generate the production of soluble protease, all six of the constructs were induced at 25°C. As shown in Figure 8, gel analysis of whole cell samples indicated that all of the constructs, with the exception of the Opt19 construct,
FIG. 6. Gel Analysis of Whole Cell Samples from Inductions at 37°C. Expression of the protease is evidenced by the presence of a 25 kDa molecular weight band in lanes 3-5 and 7-8. Lane 1, broad range molecular weight markers; lane 2, pET15b negative control sample; lane 3, Ik15; lane 4, Nde15; lane 5, Opt15; lane 6, pET19b negative control sample; lane 7, Ik19; lane 8, Nde19; lane 9, Opt19; lane 10, low range molecular weight markers.
FIG. 7. Gel Analysis of Soluble and Insoluble Fractions from Samples Induced at 37°C. At 37°C, the protease is expressed as an insoluble protein. Lane 1, broad range molecular weight markers; lane 2, Ik15 soluble fraction; lane 3, Ik15 insoluble fraction; lane 4, Ik19 soluble fraction; lane 5, Ik19 insoluble fraction; lane 6, Nde15 soluble fraction; lane 7, Nde15 insoluble fraction; lane 8, Nde19 soluble fraction; lane 9, Nde19 insoluble fraction, lane 10, low range molecular weight markers.
FIG. 8. Gel Analysis of Whole Cell Samples from Inductions at 25°C. As in the gel shown in Figure 6, the protease is expressed as a 25 kDa molecular weight protein. Lane 1, broad range molecular weight markers; lane 2, pET15b negative control sample; lane 3, Ikl5; lane 4, Nde15; lane 5, Opt15; lane 6, pET19b negative control sample; lane 7, Ikl9; lane 8, Nde19; lane 9, Opt19; lane 10, low range molecular weight markers.
FIG. 9. Gel Analysis of Soluble and Insoluble Fractions from Samples Induced at 25°C. At 25°C, the protease is expressed both solubly and insolubly. Lane 1, Benchmark molecular weight markers; lane 2, Ik15 soluble fraction; lane 3, Ik15 insoluble fraction; lane 4, Ik19 soluble fraction; lane 5, Ik19 insoluble fraction; lane 6, Nde15 soluble fraction; lane 7, Nde15 insoluble fraction; lane 8, Nde19 soluble fraction; lane 9, Nde19 insoluble fraction.
produced protease when expression of the gene was induced at 25°C. Again, no band corresponding to the molecular weight of the protease was apparent in the lanes containing samples of the negative control plasmids.

As before, the soluble and insoluble fractions of the induced samples were analyzed for the Ikeda and Nde constructs. The results for this analysis are shown in Figure 9. At the 25°C induction temperature, the protease was still expressed primarily in an insoluble form, however, some protease was expressed solubly. Soluble and insoluble fractions for the Opt19 construct were not evaluated. Analysis of the soluble and insoluble fractions of proteins produced by the Opt15 construct (data not shown) indicated that some of the protease was expressed solubly. Based on these results, subsequent inductions were conducted at 25°C. Due to the difficulties associated with expression of the Opt19 construct at both 25°C and 37°C, no further studies were performed on either the Opt15 or the Opt19 constructs.

The length of the induction for each construct was determined by evaluating the absorbance of the solution when inductions were conducted for 14-18 hours. The absorbance of each of the constructs increased during the first six hours of the induction. The level of absorbance continued to increase during the 14-18 hour induction period for the Ik15, Ik19, and Nde15 constructs. Subsequent inductions of these constructs were performed for 14-18 hours. The Nde19 construct exhibited a decrease in absorbance during the 14-18 hour induction, and subsequent inductions on this construct were performed for six hours.

RESULTS OF PURIFICATION

Initially, the protease was expressed as a fusion protein of approximately 20 kDa molecular weight. Upon autoprocessing, the molecular weight of the protease dropped to approximately 14 kDa. In order to use gel analysis to evaluate whether autoprocessing occurred, most of the proteins produced by the E. coli cells corresponding to a molecular weight of 20 kDa or less needed to be removed from the sample. The results of the purification procedure are shown in Figure 10.
FIG. 10. Evaluation of Purification Protocol Using Gel Electrophoresis. Samples of crude lysate from 100 mL pellets for each of the protease constructs was applied to lanes 2, 4, 6, and 8. Each of these lanes contains many proteins of varying molecular weight. Once purified, samples of the eluate fraction from the construct was applied to lanes 3, 5, 7, and 9. With the exception of the Ik19 construct (eluate fraction in lane 5), the purification protocol was successful in removing many of the lower molecular weight proteins which were present in the lysate samples. Lane 1, Benchmark molecular weight markers; lane 2, Ik15 lysate; lane 3, Ik15 eluate; lane 4, Ik19 lysate; lane 5, Ik19 eluate; lane 6, Nde15 lysate; lane 7, Nde15 eluate; lane 8, Nde19 lysate; lane 9, Nde19 eluate.
Although the protease was not completely purified from the lower molecular weight proteins, the majority of the proteins corresponding to a molecular weight of 20 kDa or less were removed from the purified samples of Ik15, Nde15 and Nde19. The samples for these constructs are found in lanes 3, 7 and 9, respectively. Purification of the Ik19 construct, shown in lane 5, was not as complete.

RESULTS OF PROTEASE ACTIVATION

Protease activation was accomplished by a two-step dialysis protocol. First, the protease sample was dialyzed against sodium acetate buffer at a pH of 3.5. Dialysis against sodium acetate was followed by dialysis against either sodium citrate buffer or MES buffer for each of the constructs. The pH of the sodium citrate and MES buffers was 5.3 and 5.5, respectively.

Figure 11 shows the results of gel analysis of Ik15 dialyzed against both citrate and MES buffers. Curiously, even though each of the constructs was subjected to the same experimental protocol, only the gel analysis of the activated Ik15 construct conclusively demonstrated that autoprocessing of the protease had occurred. In this figure, a 25 kDa protein was present in both the lysate and eluate fractions, shown in lanes 2 and 3. In lane 4, most of the 25 kDa band had vanished, but a new 14 kDa band was present. The presence of the 14 kDa band indicated the occurrence of autoprocessing in this construct when the sample was dialyzed against citrate buffer. Lanes 6-9 of this same figure contain samples of the Ik15 construct dialyzed against MES buffer. As before, the presence of a 25 kDa band was present in lanes 6 and 7, indicating the presence of the His-tagged protease in both the lysate and eluate fractions. However, no 14 kDa band was visible in the samples which had been dialyzed against MES buffer, shown in lanes 8 and 9. In fact, the 25 kDa band was still clearly evident in lanes 8 and 9, suggesting the autoprocessing did not occur when the sample was dialyzed against MES buffer.

RESULTS OF ACTIVITY ASSAYS

The graphical representations of the increase in fluorescence during the course of the activity assay exhibited by each of the constructs are presented in Figures 12-19.
FIG. 11. Evaluation of Ik15 Protease Autoprocessing Using Gel Electrophoresis. Lane 1, broad range molecular weight markers; lane 2, Ik15 lysate; lane 3, Ik15 eluate; lane 4, Ik15 eluate dialyzed against citrate buffer, insoluble fraction; lane 5, Ik15 eluate dialyzed against citrate buffer, soluble fraction; lane 6, Ik15 lysate; lane 7, Ik15 eluate; lane 8, Ik15 eluate dialyzed against MES buffer, insoluble fraction; lane 9, Ik15 eluate dialyzed against MES buffer, soluble fraction.
FIG. 12. Activity Assay of Protease Produced from the Ik15 Construct and Activated Against Citrate Buffer. 500 µL of enzyme was added to the reaction mixture and assayed as described in methods.
FIG. 13. Activity Assay of Protease Produced from the Ik15 Construct and Activated Against MES Buffer. 500 µL of enzyme was added to the reaction mixture and assayed as described in methods.
FIG. 14. Activity Assay of Protease Produced from the Ik19 Construct and Activated Against Citrate Buffer. 500 µL of enzyme was added to the reaction mixture and assayed as described in methods.
FIG. 15. Activity Assay of Protease Produced from the Ik19 Construct and Activated Against MES Buffer. 500 µL of enzyme was added to the reaction mixture and assayed as described in methods.
FIG. 16. Activity Assay of Protease Produced from the Nde15 Construct and Activated Against Citrate Buffer. 500 µL of enzyme was added to the reaction mixture and assayed as described in methods.
FIG. 17. Activity Assay of Protease Produced from the Nde15 Construct and Activated Against MES Buffer. 500 µL of enzyme was added to the reaction mixture and assayed as described in methods.
FIG. 18. Activity Assay of Protease Produced from the Nde19 Construct and Activated Against Citrate Buffer. 500 µL of enzyme was added to the reaction mixture and assayed as described in methods.
FIG. 19. Activity Assay of Protease Produced from the Nde19 Construct and Activated Against MES Buffer. 500 µL of enzyme was added to the reaction mixture and assayed as described in methods.
Protein concentrations, specific activity calculations, and ratio of activity levels for the assayed samples are presented in Table 4. Units for specific activity were Integrated Fluorescence Units·µL/Seconds·µg.

Positive and negative controls were also performed and are presented in Appendices A-D. The negative controls did not contain any enzyme. The positive control contained trypsin, an enzyme known to cleave the substrate sequence used in the assay.

**Results of Activity Assays of the Ik15 and Ik19 Constructs**
The increase in fluorescence in the Ik15 constructs dialyzed against citrate and MES buffers are shown in Figures 12 and 13. Figures 14 and 15 show the fluorescence assays for the Ik19 construct dialyzed against citrate and MES buffers, respectively. A summary of these results is shown in Table 4.

The results of the activity assays for these two constructs indicate that approximately the same level of activation was achieved in both the Ik15 and Ik19 constructs when dialyzed against citrate buffer. When the same constructs are dialyzed against MES buffer, the Ik15 construct produces approximately twice as much activated enzyme as the Ik19 construct.

**Results of Activity Assays of the Nde15 and Nde19 Constructs**
The results of the fluorometric analysis for the Nde15 construct dialyzed against citrate and MES buffers are shown in Figures 16 and 17. Figures 18 and 19 contain the results of the activity assay for the Nde19 construct dialyzed against citrate and MES buffers. This data is summarized in Table 4.

The results of these assays indicate that the enzyme produced by the Nde15 construct when dialyzed against citrate buffer was approximately four times as active as that produced by the Nde19 construct when it was dialyzed against citrate buffer. When the Nde15 construct was dialyzed against MES buffer, it was still more active than the enzyme produced by the Nde19 construct that was dialyzed against MES buffer.
TABLE 4. Comparison of Activity Levels of Protease Produced by Constructs

<table>
<thead>
<tr>
<th>Construct</th>
<th>Change in Fluorescence/Time&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Protein (mg/mL)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Specific Activity&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ik15</td>
<td>0.0758</td>
<td>0.3045</td>
<td>0.2489</td>
</tr>
<tr>
<td>Ik19</td>
<td>0.0513</td>
<td>0.2285</td>
<td>0.2247</td>
</tr>
<tr>
<td>Nde15</td>
<td>0.0241</td>
<td>0.1969</td>
<td>0.1225</td>
</tr>
<tr>
<td>Nde19</td>
<td>0.0068</td>
<td>0.2057</td>
<td>0.0329</td>
</tr>
</tbody>
</table>

<sup>a</sup>This value represents the overall increase in the level of fluorescence emitted during the activity assay, divided by the length of time the sample was assayed (900 seconds).

<sup>b</sup>The protein concentration was determined using a Bradford Assay, as described in methods.

<sup>c</sup>The units for specific activity are Integrated Fluorescence Units·µL/Seconds·µg.
TABLE 4. Continued

<table>
<thead>
<tr>
<th>Samples Activated Against MES Buffer</th>
<th>Citrate:MES Activity Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Fluorescence/Time&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Protein (mg/mL)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>0.0354</td>
<td>0.4071</td>
</tr>
<tr>
<td>0.0094</td>
<td>0.1973</td>
</tr>
<tr>
<td>0.0129</td>
<td>0.1810</td>
</tr>
<tr>
<td>0.0072</td>
<td>0.2643</td>
</tr>
</tbody>
</table>

<sup>a</sup>This value represents the overall increase in the level of fluorescence emitted during the activity assay, divided by the length of time the sample was assayed (900 seconds).

<sup>b</sup>The protein concentration was determined using a Bradford Assay, as described in methods.

<sup>c</sup>The units for specific activity are Integrated Fluorescence Units·µL/Seconds·µg.

<sup>d</sup>Once the specific activities and protein concentrations for each sample were calculated, a comparison of activity levels was made by taking a ratio of the specific activity of the protease activated against citrate buffer and the specific activity of the protease activated against MES buffer.
However, in this case, the level of activity exhibited by the Nde15 construct was approximately 2½ times higher than the activity exhibited by the Nde19 construct.
CHAPTER IV
DISCUSSION OF RESULTS

This study was divided into four main phases. In the first phase, expression of the protease gene and the solubility of the expressed protein were evaluated. In the second, the protease was purified from most of the other proteins produced by *E. coli* cells during the process of induction. Purification was followed by activation of the protease against a two-step dialysis procedure utilizing buffers of varying pH to promote enzyme activation through autoprocessing and dimerization. The last step of the analysis process provided a quantitative measurement of enzyme activation utilizing fluorometry.

Expression of the protease gene in *E. coli* BL21(DE3)pLysS cells was attempted on each of the protease constructs. Analysis of whole cell samples removed during the induction process indicated that expression was successful at induction temperatures of 37°C in all of the constructs, with the exception of the Opt19 construct. Analysis of the soluble and insoluble fractions of proteins produced by each construct, with the exception of the Opt19 construct, indicated that the protease was expressed in an insoluble form at 37°C. These results indicated the probable formation of inclusion bodies, which often result during the expression of recombinant fusion proteins in *E. coli* cells (23, 35). However, by changing the temperature at which inductions are performed, it is sometimes possible to obtain recombinant protein from *E. coli* cells in a soluble form. Because of the relative ease with which soluble proteins can be analyzed, expression of the protease as a soluble protein was highly sought after, and experiments were undertaken to assess the feasibility of obtaining soluble protein.

Whole cell analyses had indicated that the protease gene would express at 25°C in all of the constructs, again with the exception of the Opt19 construct. Analysis of the soluble and insoluble fractions of the expressed proteins indicated that the protease was expressed as both a soluble and insoluble protein at 25°C. Since the protease was expressed solubly at 25°C, all subsequent inductions were performed at 25°C for lengths of time ranging from six hours (for the Nde19 construct) to 14-18 hours (for the Ik15, Ik19, and Nde15 constructs).
Although six constructs of the protease were initially produced, only four were analyzed beyond phase one, i.e., the phase involving analysis of protease expression. In the Opt15 construct, the protease was expressed as a soluble protein at 25°C, as well as in an insoluble form. However, no expression was evident in the Opt19 construct, regardless of whether inductions were performed at 25°C or 37°C. Since the constructs were designed in sets of two, the difficulties associated with the expression of the Opt19 construct led to the abandonment of both Opt constructs in subsequent analyses.

Upon optimization of the expression parameters for the Ik and Nde constructs, a purification methodology was developed to purify the protease from most of the other proteins produced by *E. coli* cells during induction. Purification of the protease utilized the presence of the molecular tag composed of histidine residues. This tag was contributed by the plasmid vector into which the protease gene had been inserted and was fused to the N-terminal end of the expressed protein. A resin was used which contained a nickel atom bound to trinitriloacetic acid (Ni-NTA). In this resin, nickel had the capacity to bind to six other atoms. Four of these binding sites are occupied by atoms contributed by the trinitriloacetic acid molecule. The remaining two sites were available to bind to nitrogen atoms located in the imidazole ring of the histidine residues fused to the protease.

Upon preparation of a crude lysate solution, incubation with the Ni-NTA resin, and introduction of the mixture to a small chromatographic column, the resin was washed with a series of buffers containing gradually increasing concentrations of imidazole. The washing of the resin by these buffers resulted in the removal of proteins non-specifically bound to the resin. The protease was eluted from the column using a buffer containing a one molar concentration of imidazole. The high concentration of imidazole displaced the histidine tagged protease, resulting in its removal from the column.

The purification protocol was reasonably effective for the Ik15, Nde15, and Nde19 constructs. It was not as effective for the Ik19 construct. Early in the investigation, a methodology was developed to enhance the level of purification of this construct by purifying it a second time using the Ni-NTA resin. However, as the
investigation progressed, it was noted that, oftentimes, a precipitate would form in the sample when dialyzed against the lysis buffer. To resolubilize the precipitate, the sample was dialyzed against a solution of lysis buffer containing an eight molar concentration of urea. Subsequent experiments indicated that when the sample had been dialyzed against the buffer containing urea, the successful completion of a second purification was not possible. Consequently, the protease from each of the constructs was purified only once.

Although the purification scheme employed was useful, it had its limitations. The primary limitation stemmed from the inability to design a purification methodology which would take into account the varying concentration of proteins within any given sample. For this reason, the success of the purification protocol varied according to the individual sample. However, the purification procedure developed was successful in removing most of the excess proteins of 20 kDa molecular weight and less in most of the samples analyzed.

Purification of the protease was followed by activation. When the protease is expressed, it is inactive. Although all of the factors involved in protease activation in vivo have not been established, two aspects of the activation process are very clear. First, the amino acid sequence, which, upon cleavage, yields the HTLV-1 protease monomer must be cleaved. Second, dimerization of the monomeric units must be promoted. Activation in vitro was accomplished utilizing a two-step dialysis process with buffers of varying pH. In the first step, the purified enzyme samples were dialyzed against a sodium acetate buffer with a pH of 3.5. This dialysis was followed by dialysis against either sodium citrate (pH 5.3) or MES buffer (pH 5.5). Separate samples of each construct were dialyzed against citrate or MES buffer. Dialysis of the sample against either the citrate or MES buffers resulted in the formation of a white precipitate, although most of the samples remained soluble during dialysis. In order to facilitate the analysis of both the soluble and insoluble fractions, the two were separated via centrifugation. Samples of the precipitate and the soluble proteins were analyzed for the occurrence of autoprocessing using gel electrophoresis.
Only the Ik15 construct conclusively demonstrated protease autoprocessing. In this construct, autoprocessing was only seen when the sample was dialyzed against the citrate buffer. No autoprocessing was evident when the sample was dialyzed against MES buffer. Although these results seemed unusual at first, the differences in the level of autoprocessing occurring in the same construct when dialyzed against different activation buffers introduced a new aspect of analysis to this project. These results led to the idea that the buffer against which activation is promoted affected the level of protease activation. In fact, an evaluation of the effects of buffers on protease activation was not originally part of this project. However, during the course of the study, it became apparent that the buffers may play an important role in protease activation, and it was for this reason that dialyses were conducted against both citrate and MES buffers.

Although it would have been ideal to have qualitative evidence of autoprocessing prior to beginning the activity assays, the lack of qualitative data did not necessarily mean that autoprocessing did not occur within each of the constructs. To quantitatively evaluate enzymatic activity, activity assays using fluorometry were employed. Using the same activation protocol used previously to activate the Ik15 construct, samples of all four of the constructs were prepared and dialyzed against either citrate or MES buffer prior to use in the activity assay.

Quantitative measurement of the activity level of the enzyme produced by each construct could only be obtained by measuring the activity of the activated enzyme. A substrate with the sequence, Abs-KTKVLVVQPK(3NO2Y)A, was used. This sequence corresponded to the native sequence, which, upon cleavage at the leucine and valine residues (residues 5 and 6 of the sequence), yields the mature core proteins p15 and p24. In this substrate, fluorescing groups were present at either end of the molecule. When the fluorescing groups were close to one another, as when the substrate molecule is intact, fluorescence was quenched. However, when the substrate was cut and the groups could move away from one another, fluorescence increased. Monitoring the increased level of fluorescence over time provided a means by which the enzyme produced by each construct could be evaluated for its ability to cleave the substrate.
Although activity assays were important in the evaluation of enzymatic activity, the activity assay alone was insufficient to allow conclusions to be drawn with regards to the ability of each construct to produce active enzyme. In order to make accurate comparisons of activity levels from one construct to another, the protein concentration of each sample was determined using a Bradford protein assay. Using the data from the protein assay, in conjunction with the level of increased fluorescence recorded during the activity assay, the specific activity of the enzyme produced from each construct was calculated. The determination of specific activity for the enzyme produced by each construct enabled accurate conclusions to be drawn with regards to the actual enzymatic activity of each sample.

The data from the activity assays was evaluated from two different perspectives. First, any differences in the level of activation achieved by a construct when dialyzed against either citrate or MES buffer needed to be established. Please refer back to Table 4 for these data. In each of the constructs, dialysis against citrate buffer led to a higher level of enzymatic activity when compared to the activity level obtained in the same construct when dialyzed against MES buffer. In the Ik15 construct, dialysis against citrate buffer produced enzyme almost three times as active as the Ik15 sample dialyzed against MES buffer. Similarly, in the Ik19 construct, dialysis against citrate buffer produced enzyme that had almost five times as much activity as the Ik19 sample dialyzed against MES buffer. The differences in the level of activation evidenced in the Nde15 constructs were not as drastic. In the Nde15 construct, dialysis against citrate produced enzyme approximately 1½ times as active as the Nde15 construct dialyzed against the MES buffer. The least difference in the level of activation was seen in the Nde19 constructs. When dialyzed against citrate, the enzyme produced by this construct was approximately 1½ times as active as that produced by the construct dialyzed against MES buffer. These results clearly indicate that the type of buffer used in the activation step has an effect on the level of activity exhibited by the enzyme, regardless of which construct is used.

These differences in the level of activation due to the influences of buffer can be explained, at least in part, by analyzing the two buffers used to promote activation. Both
buffers contained the same amount of sodium chloride, EDTA, DTT, and water. The only difference between them lay in the presence of sodium citrate or MES within the buffer system. The concentrations of sodium citrate and MES in each of the buffer solutions were also the same. However, sodium citrate is an anti-chaotroph, while MES is not. As an anti-chaotroph, sodium citrate has the ability to bring molecules together within a solution. Since dimerization is an integral part of protease activation, it is suspected that the sodium citrate buffer encourages dimerization, resulting in an increased level of enzyme activation.

Data from the activity assays also needed to be evaluated to determine the levels of enzyme activity seen in each set of protease constructs, dialyzed against each buffer system. Since the same gene was put into both the pET15b and pET19b vectors, it was possible to evaluate the effects of the N-terminal amino acid sequence on the ability of the protease to autoprocess.

In the Ikeda constructs dialyzed against citrate buffer, the level of activation was about the same in both the Ikl5 and Ikl9 construct. However, when the Ikeda constructs were dialyzed against the MES buffer, the Ikl5 construct exhibited approximately twice as much activity as the Ikl9 construct. In the Nde constructs dialyzed against citrate buffer, the Nde15 construct possessed about four times as much enzymatic activity as the Nde19 construct. And, when the Nde constructs were dialyzed against MES buffer, the Nde15 construct was about two and half times as active as the Nde19 construct.

Analysis of these data shows that the construct composed of a protease gene inserted into the pET15b vector produces a higher level of enzymatic activity than the same gene inserted into the pET19b vector. However, the difference in the level of enzymatic activity in the Ikeda constructs dialyzed against citrate buffer appears to be negligible. These differences in the activity levels of the enzymes produced from each construct may be related to the differences in the plasmids into which each gene has been inserted.
We referred to the original hypothesis upon which this study was based in order to explain the differences in the level of activity exhibited by specific constructs. In this hypothesis, it was assumed that the N-terminal amino acid sequence had some influence over the ability of the protease to autoprocess. To evaluate these effects, the SPAs of each of the constructs, as well as the SPAs of the native cleavage sequences, must be analyzed. These are presented in Appendices E-G and Figure 3, respectively.

Analysis of the SPA of the native cleavage sequence for the p15 and p24 proteins show that a hydrophilic region consisting of seven amino acid residues is located two amino acids N-terminally to the protease cleavage site. Analyses of the SPAs of the native sequences which produce the mature protease and the proteins, p19 and p24, indicate the presence of hydrophilic regions located five amino acids N-terminally to the protease autocatalytic/cleavage site. However, the N-terminal sequences found in these last two are not as hydrophilic as the N-terminal sequence which produced the p15 and p24 proteins. Upon evaluation of the SPAs of the Ikeda and Nde constructs, some similarities and differences were noted when compared to the SPA of the native cleavage sequences. In the SPAs for the Ik15 and Ik19 constructs, there was a small, hydrophilic region located approximately two amino acids N-terminally to the cleavage site. In the Nde15 construct, a small hydrophilic region was also noted, but it was located approximately seven amino acids N-terminal to the cleavage site. In the Nde19 construct, no small region of hydrophobic amino acids was noted. Instead, in this construct, a large hydrophilic region consisting of about eight amino acids was located six amino acids N-terminal to the cleavage site. An attempt was made to explain these results based on the original hypothesis.

The highest activity levels were seen in enzyme produced from the Ik15, Ik19, and Nde15 constructs. In the samples dialyzed against citrate buffer, the enzyme produced by the Ik15 and Ik19 constructs was almost twice as active as the enzyme produced by the Nde15 construct. In all three of these constructs, the cleavage site was preceded by a short hydrophilic sequence. In the two constructs exhibiting the highest levels of activity, the Ik15 and Ik19 constructs, this hydrophilic sequence was virtually adjacent to the cleavage site. The Nde15 construct also contained this hydrophilic
sequence N-terminally to the cleavage site. However, in this construct, it was located seven amino acids away from the sequence of amino acids at which cleavage occurred. This distance, coupled with the decreased levels of active enzyme produced by this construct in comparison to the levels produced by the Ik15 and Ik19 constructs, implied that the presence of a hydrophilic sequence immediately adjacent to the cleavage site was required for optimal enzymatic activity. Analysis of the native cleavage sequences also supports the presence of a hydrophilic sequence in close proximity to the protease cleavage site as a probable requirement for recognition of the cleavage site by the protease.

Further support for this conclusion was obtained by analyzing the data produced from the Nde19 construct. The enzyme produced by the Nde19 construct had the lowest activity levels, regardless of whether the enzyme activation was promoted by citrate buffer or MES buffer. The inability of the Nde19 construct to produce comparable levels of activated enzyme implied that autoprocessing and activation of enzyme produced by this construct was impaired in some way. Analysis of the SPA for this construct showed the presence of a hydrophilic sequence, approximately eight amino acids long, located six amino acids N-terminally to the cleavage site. In addition, as in all of the other constructs, the sequence located five amino acids C-terminally to the cleavage site, also exhibits a significant degree of hydrophilicity. The flanking of the hydrophobic cleavage site by two hydrophilic sequences seemed to result in a decreased ability for cleavage site recognition, producing the lowest levels of enzymatic activity, regardless of which buffer system was used to promote activation.
CHAPTER V
CONCLUSIONS

The current study was developed in response to experimental results obtained in two different laboratory environments. In the previous studies, active, insoluble protease of the HTLV-1 retrovirus was produced when the protease gene was inserted in the plasmid vector, pET19b. In contrast, when the same gene was inserted in the plasmid vector, pET29, active protease was not produced. Evaluation of the differences between these two plasmids led to the development of the hypothesis that the sequence of amino acids located N-terminally to the autoprocessing site in HTLV-1 protease may affect the ability of the protease to autoprocess and become active.

Although the exact reasons for these differences were not clear, it was thought that the lack of enzyme activation could be due, at least in part, to an inability of the enzyme to recognize the cleavage site. Analysis of the surface topology of native cleavage sites indicated that the cleavage sequence itself was located in a hydrophobic region. However, the regions located N-terminally to the cleavage site were hydrophilic in nature. By utilizing this information, a study was designed to assist in the evaluation of the role the N-terminal amino acids may play in the production of active protease. This was accomplished by evaluating expression, purification and activation parameters for a set of protease constructs.

Based on the results obtained during this study, it would appear that the findings supported the original premise, i.e., that the sequence of amino acids located N-terminally to the autoprocessing site within the expressed protease exerted an effect on the ability of a particular construct to produce active enzyme. In addition, during the course of this project, it became apparent that the type of buffer used to encourage enzymatic activation also an effect on the levels of activity achieved by the enzyme produced by a specific construct. The results obtained through this research indicated that a hydrophilic region located N-terminally and in close proximity to the protease cleavage site, as well as the
use of anti-chaotrophic buffers to promote dimerization and activation of the protease, led to higher levels of enzymatic activity.

In addition, under the current experimental conditions, it appeared that the presence of only one of these parameters was insufficient to promote maximal enzymatic activity. In the Ik15 and Ik19 constructs, the hydrophilic region was virtually adjacent to the cleavage site, and maximal activity was achieved when a buffer containing the anti-chaotroph, sodium citrate, was employed to promote activation. When MES buffer was used to promote activation in the same constructs, over half of the enzymatic activity was lost. The need for the presence of both parameters was also evident when analyzing the data resulting from analysis of the Nde15 construct. The hydrophilic sequence in this construct was located several amino acids away from the cleavage site. In this construct, the sequence was small, only a couple of amino acids in length, just as in the Ikeda constructs. Dialysis of the enzyme produced by this construct against citrate buffer produced nearly twice as much active enzyme when compared to the activity levels achieved by the enzyme produced by this same construct when dialyzed against MES buffer. However, the enzyme produced by this construct still exhibited only about half of the enzymatic activity exhibited by the Ikeda constructs when they were dialyzed against citrate buffer.

And, still, further evidence to support the need for both of these parameters was evident when analyzing the data from the Nde19 construct. In this construct, the hydrophilic region is large, approximately eight amino acids in length, and is located about six amino acids N-terminal to the cleavage site. The enzyme produced by this construct exhibited the lowest levels of active enzyme, regardless of which buffer was used to promote activation.

Based on these results, it would be ideal to be able to conclude that the N-terminal amino acid sequence made a difference in the level of enzymatic activity achieved by the protease produced by a specific construct. However, it is doubtful that such a firm conclusion can be drawn, based on the experimental data. Although it appears that there are differences in the level of activation exhibited by the protease, one has to determine
whether these differences are significant. In all likelihood, they probably aren’t. In order to be considered significant, there should be, at a minimum, a five-fold difference in activity levels. The protease produced by the Ik19 construct when dialyzed against the citrate buffer exhibits almost five times the level of enzymatic activity as protease produced by the same construct dialyzed against the MES buffer. But, this could be due to the anti-chaotrophic nature of the citrate buffer, rather than having anything to do with the amino acid sequence located N-terminally to the cleavage site. Protease produced from the other three constructs did not exhibit this degree of difference in the level of activation.

In conclusion, it is not possible to state with certainty that the sequence of amino acids located N-terminally to the cleavage site has an effect on the ability of the protease produced by a specific construct to autoprocess. In fact, since enzymatic activity was achieved by the protease produced from each construct, fine-tuning the experimental protocol may lead to an increased level of activity in those constructs which currently exhibit lower activity levels, namely, the Nde15 and Nde19 constructs. And, if this is the case, then it could be conclusively stated that the N-terminal amino acid sequence does not affect protease autoprocessing.

The goal of this project was to evaluate expression and activation parameters for HTLV-1 protease produced from a variety of different constructs. Previous research on the HTLV-1 protease has been hampered by the difficulties associated with the production of mature, active enzyme. Although the original hypothesis was not necessarily supported by the experimental data, active, soluble enzyme was produced during the course of this study with minimal difficulties. Hopefully, the results of this study will assist in overcoming some of the obstacles previously associated with research on this enzyme.
REFERENCES


REFERENCES


REFERENCES


REFERENCES


FIG. 20. Positive Control for the Activity Assay Using Citrate Buffer. Trypsin is known to cleave the substrate used in the activity assays, and a reaction mixture containing trypsin, substrate, and buffer was assayed for activity. The level of fluorescence emitted by the fluorescing groups attached to the ends of the substrate increased by 93.2 integrated fluorescence units during the assay.
FIG. 21. Positive Control for the Activity Assay Using MES Buffer. Trypsin is known to cleave the substrate used in the activity assays, and a reaction mixture containing trypsin, substrate, and buffer was assayed for activity. The level of fluorescence emitted by the fluorescing groups attached to the ends of the substrate increased by 98.1 integrated fluorescence units during the assay.
FIG. 22. Negative Control for the Activity Assay Using Citrate Buffer. A reaction mixture containing substrate and buffer was assayed for activity. No increase in fluorescence was noted.
FIG. 23. Negative Control for the Activity Assay Using MES Buffer. A reaction mixture containing substrate and buffer was assayed for activity. No increase in fluorescence was noted.
FIG. 24. Surface Probability Analyses for the Ikeda (Ik) Protease Constructs. The underlined amino acid sequences comprise the sequence of amino acids recognized by the protease for cleavage. Cleavage between the aspartic acid and proline residues of the underlined amino acid sequence results in the production of autoprocessed protease. (A) Surface probability analysis for the Ik15 construct. (B) Surface probability analysis for the Ik19 construct.
FIG. 25. Surface Probability Analyses for the Nde Protease Constructs. The underlined amino acid sequences comprise the sequence of amino acids recognized by the protease for cleavage. Cleavage between the aspartic acid and proline residues of the underlined amino acid sequence results in the production of autoprocessed protease. (A) Surface probability analysis for the Nde15 construct. (B) Surface probability analysis for the Nde19 construct.
APPENDIX G

SURFACE PROBABILITY ANALYSES FOR THE OPT PROTEASE CONSTRUCTS

FIG. 26. Surface Probability Analyses for the Opt Protease Constructs. The underlined amino acid sequences comprise the sequence of amino acids recognized by the protease for cleavage. Cleavage between the leucine and proline residues of the underlined amino acid sequence results in the production of autoprocessed protease. (A) Surface probability analysis for the Opt15 construct. (B) Surface probability analysis for the Opt19 construct.
VITA

Hidayah Muhammad Kendall

Education
Old Dominion University • Norfolk, Virginia
• Master of Science, Chemistry, 1999
• Bachelor of Science, 1997
  Biochemistry major, Biology minor

Experience
Old Dominion University
Department of Chemistry and Biochemistry
4541 Hampton Boulevard, Alfriend Chemistry Building
Norfolk, Virginia 23529

Graduate Researcher: Analyzed six HTLV-1 protease clones to determine the effects of N-terminal amino acid sequence on the expression of active recombinant protease

Graduate Teaching Assistant: Instructed undergraduate students in general and organic chemistry laboratories.

Undergraduate Research Assistant: Provided research support for analytical laboratory. Focused on the detection of cocaine and heroin, as well as their hydrochlorides and metabolites, using ion mobility spectrometry.

Presentations and Seminars
• "A Study of HTLV-1 Protease Clones". Departmental seminar, 1999.

Awards and Accomplishments
• John D. Van Norman Award for Outstanding Graduate Research, 1998-1999
• John D. Van Norman Award for Outstanding Undergraduate Research, 1996-1997
• Phi Theta Kappa Honor Society, 1993-1994
• Honor Roll and Dean’s List, various semesters