HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1) \textit{GAG-PRO} GENETIC VARIABILITY IS RELATED TO THERAPY RESPONSE AND IMPACTS VIRAL FITNESS IN SPECIFIC CELL TYPES

By

STEPHANIE LEIGH ROSE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2002
This dissertation is dedicated to my parents, Rose and Russell Rose, who have made countless sacrifices so I could pursue my dreams.
ACKNOWLEDGMENTS

I would first like to thank my mentor, Dr. Maureen Goodenow, for her continued guidance and encouragement in the completion of my studies. Dr. Goodenow is an outstanding role model for women in science and I am honored to have been given the opportunity to work with her. I would like to thank Dr. John Sleasman for his clinical perspective on my research and for providing access to a critical set of samples important for this project. I would also like to thank the other members of my advisory committee (Dr. Ben Dunn, Dr. Jeff Harrison, Dr. Ayalew Mergia, and Dr. Keith Peden) for their contributions to this work. I am grateful for the members of the Goodenow lab, past and present (Dr. Brant Burkhardt, Steve Pomeroy, Susanna Lamar, Dr. Guity Ghaffari, Gregg Taylor, Dr. Zhong Kou, Carter Coberley, Brain Peyser, Dr. Dan Tuttle, Rose Mills and Sarah Koch) who provided technical and personal support for me over the past four years.

On a personal note, there are many people to whom I am grateful or indebted. Dr. James Kohler has provided immeasurable moral and scientific support during this endeavor, enabling me to keep life and science in perspective. I acknowledge the support of my friends, especially Melissa Fleegal, who took the challenging graduate school journey alongside me.

Most importantly I want to thank my parents for their unwavering support of this work and for believing in me even when I doubted my own capabilities. Without their love and encouragement it would not have been possible to achieve my goals.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION AND BACKGROUND</td>
<td>1</td>
</tr>
<tr>
<td>Specific Aims</td>
<td>3</td>
</tr>
<tr>
<td>Significance</td>
<td>5</td>
</tr>
<tr>
<td>Human Immunodeficiency Virus</td>
<td>5</td>
</tr>
<tr>
<td>HIV-1 Life Cycle</td>
<td>5</td>
</tr>
<tr>
<td>Targets of HIV-1 Infection</td>
<td>6</td>
</tr>
<tr>
<td>CD4+ Cell Subsets are Classified by CD45 Isoform Expression</td>
<td>7</td>
</tr>
<tr>
<td>Envelope</td>
<td>8</td>
</tr>
<tr>
<td>HIV-1 gag-pol</td>
<td>9</td>
</tr>
<tr>
<td>Reverse Transcriptase</td>
<td>9</td>
</tr>
<tr>
<td>Protease</td>
<td>10</td>
</tr>
<tr>
<td>Gag Cleavage Sites</td>
<td>11</td>
</tr>
<tr>
<td>Natural Genetic Variation of gag-pol</td>
<td>13</td>
</tr>
<tr>
<td>Processing of Natural Polymorphisms in the Bacterial Expression System</td>
<td>13</td>
</tr>
<tr>
<td>Protease and the Selective Pressure of Drug Therapy</td>
<td>14</td>
</tr>
<tr>
<td>Implication of Gag Cleavage Sites Polymorphisms</td>
<td>15</td>
</tr>
<tr>
<td>Response to Therapy</td>
<td>16</td>
</tr>
<tr>
<td>Compartmentalization of Drug Resistant Viruses</td>
<td>17</td>
</tr>
<tr>
<td>Viral Fitness and the Dynamics of Diversity</td>
<td>18</td>
</tr>
<tr>
<td>Viral Reservoirs and Latency</td>
<td>18</td>
</tr>
<tr>
<td>Phylogenetic Analysis</td>
<td>19</td>
</tr>
<tr>
<td>2 HIV-1 PROTEASE GENOTYPE PREDICTS THERAPY RESPONSE</td>
<td>29</td>
</tr>
<tr>
<td>Introduction</td>
<td>29</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>31</td>
</tr>
<tr>
<td>Study Patients</td>
<td>31</td>
</tr>
<tr>
<td>Clinical Monitoring, Sample Collection, and Processing</td>
<td>32</td>
</tr>
<tr>
<td>Response to Therapy</td>
<td>32</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Sequence Analysis</td>
<td>33</td>
</tr>
<tr>
<td>Genetic Analysis</td>
<td>35</td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>36</td>
</tr>
<tr>
<td>Results</td>
<td>37</td>
</tr>
<tr>
<td>Viral and Immune Response to Therapy</td>
<td>37</td>
</tr>
<tr>
<td>Pretherapy Clinical Characteristics among the Therapy Response Groups</td>
<td>39</td>
</tr>
<tr>
<td>Genetic Relationship Among Viruses</td>
<td>40</td>
</tr>
<tr>
<td>Amino Acid Polymorphisms within RT</td>
<td>40</td>
</tr>
<tr>
<td>Amino Acid Polymorphisms within Pretherapy PR</td>
<td>41</td>
</tr>
<tr>
<td>Therapy-Specific Amino Acid Substitutions in Pretreatment PR Genotypes</td>
<td>42</td>
</tr>
<tr>
<td>Clinical and Genetic Variables That Predict Viral and Immune Outcome</td>
<td>44</td>
</tr>
<tr>
<td>Discussion</td>
<td>45</td>
</tr>
</tbody>
</table>

3 GENETIC ANALYSIS OF $GAG-PRO$ FOLLOWING 24 WEEKS OF COMBINATION THERAPY

Introduction .................................................................................................................. 56
Materials and Methods ................................................................................................. 57
Study Patients ............................................................................................................. 57
Clinical Monitoring, Sample Collection, and Processing ....................................... 58
Sequence Analysis .................................................................................................... 58
Genetic Analysis .................................................................................................... 60
Results .......................................................................................................................... 60
Viral and Immune Response to Therapy ....................................................................... 60
Evaluation of Clinical Variables ............................................................................ 61
Amino Acid Polymorphisms in Protease [PR] ............................................................ 61
Supplemental Polymorphisms in Protease ................................................................ 62
Distance Analysis ........................................................................................................ 63
Genetic Relationship among Viruses ........................................................................... 65
Analysis of Gag ........................................................................................................... 66
Discussion .................................................................................................................... 68

4 SEGREGATION OF HIV-1 PROTEASE GENOTYPIC RESISTANCE WITHIN BLOOD COMPARTMENTS IDENTIFIES CD4$^+$ CD45RA T LYMPHOCYTES AS A LONG LIVED VIRAL RESERVOIR

Introduction .................................................................................................................. 87
Materials and Methods ................................................................................................. 89
Study Patients and Samples ........................................................................................ 89
Separation of CD4$^+$ T Cell Subsets ......................................................................... 89
Generation of Protease Sequence ............................................................................... 89
Generation of Envelope Sequence ................................................................................ 90
Genetic Analysis ........................................................................................................... 90
Culture of CD4$^+$ T Cell Subsets ............................................................................. 91
Results .......................................................................................................................... 91
Identification of Patients for Separation of CD4$^-$ T Cells Subsets ....................... 91
5 REPLICATION OF HIV-1 IS INFLUENCED BY GAG-PRO AND TARGET CELL TYPE

Introduction ................................................................. 108
Materials and Methods .................................................. 110
Recombinant Viruses ...................................................... 110
Generation of pLAI.4 ...................................................... 111
Generation of Patient Gag-pro Replication Competent Recombinant Viruses ... 112
Strategy 1 ........................................................................ 112
Strategy 2 ........................................................................ 112
Generation of Envelope Recombinant Viruses ........................ 113
Production of Recombinant Virus Stocks ............................ 114
Isolation of Target PBMC and Macrophages for Infections .... 115
Virus Replication Assays .................................................. 116
Infection of MDM .............................................................. 116
Gag-PR Processing in Jurkat Cells ..................................... 117
Results ............................................................................. 118
Processing Phenotype is Reflected by Virus Replication .......... 118
Gag Processing in Jurkat Cells .......................................... 119
Pre and Post Therapy Alleles from a Discordant Patient ....... 119
Replication of Recombinant Virus in PBMC ......................... 120
Replication of R5 Virus in PBMC ....................................... 121
Replication of Recombinant Virus in Macrophages ............... 121
Discussion ........................................................................ 122

6 DISCUSSION ..................................................................... 139
Protease Genotypic Resistance as a Novel Marker ............... 139
Response to Protease Inhibitor Therapy, Implications of Viral Genetics .......... 140

APPENDIX

A ALIGNMENT OF PROTEASE SEQUENCES .................................... 145
B ALIGNMENT OF GAG SEQUENCES ........................................... 169
REFERENCES ........................................................................................................................................175
BIOGRAPHICAL SKETCH ..................................................................................................................194
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1. Clinical characteristics of study population at entry</td>
<td>50</td>
</tr>
<tr>
<td>2-2. A combination of clinical and genetic variables predict therapy outcome</td>
<td>55</td>
</tr>
<tr>
<td>3-1. Clinical characteristics of study population at entry and 24 weeks posttherapy</td>
<td>71</td>
</tr>
<tr>
<td>3-2. Statistical analysis of amino acid substitutions in HIV-1 protease associated with reduced sensitivity to protease inhibitors</td>
<td>73</td>
</tr>
<tr>
<td>3-3. Sequence of dominant C cleavage site detected pre- and posttherapy</td>
<td>85</td>
</tr>
<tr>
<td>3-4. Percent of patients displaying polymorphism at the respective amino acid positions in the C Cleavage site</td>
<td>86</td>
</tr>
<tr>
<td>5-1. Protease recombinant viruses</td>
<td>130</td>
</tr>
<tr>
<td>5-2. Fold difference in supernatant p24 production relative to pretherapy</td>
<td>138</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1.</td>
<td>HIV-1 proviral genome.</td>
<td>22</td>
</tr>
<tr>
<td>1-2.</td>
<td>Life cycle of HIV-1.</td>
<td>23</td>
</tr>
<tr>
<td>1-3.</td>
<td>Ribbon diagram of the three-dimensional structure of HIV-1 protease bound to a protease inhibitor</td>
<td>24</td>
</tr>
<tr>
<td>1-4.</td>
<td>Electron micrograph of HIV-1 particle assembly and release</td>
<td>25</td>
</tr>
<tr>
<td>1-5.</td>
<td>Models for retroviral structure.</td>
<td>26</td>
</tr>
<tr>
<td>1-6.</td>
<td>Processing phenotypes of Gag-Pol polyprotein.</td>
<td>27</td>
</tr>
<tr>
<td>1-7.</td>
<td>Amino acid mutations in protease associated with reduced sensitivity to protease inhibitors</td>
<td>28</td>
</tr>
<tr>
<td>2-1.</td>
<td>Viral and immune responses during 44 weeks of combination for human immunodeficiency virus (HIV) type 1 infection</td>
<td>49</td>
</tr>
<tr>
<td>2-2.</td>
<td>Phylogenetic relationships of reverse transcriptase and protease sequences among the cohort of subjects with human immunodeficiency virus type 1 (HIV-1) infection</td>
<td>51</td>
</tr>
<tr>
<td>2-3.</td>
<td>Reverse transcriptase (RT) genotypes before initiation of combination therapy for HIV-1 infection</td>
<td>52</td>
</tr>
<tr>
<td>2-4.</td>
<td>Genotypes of protease (PR) before PR inhibitor (PI) combination therapy for HIV-1 infection</td>
<td>53</td>
</tr>
<tr>
<td>3-1.</td>
<td>Genotype of HIV-1 protease following 24 weeks of combination therapy</td>
<td>72</td>
</tr>
<tr>
<td>3-2.</td>
<td>Number of new amino acid substitutions in protease after 24 weeks of therapy</td>
<td>74</td>
</tr>
<tr>
<td>3-3.</td>
<td>Supplemental amino acid mutations in HIV-1 protease</td>
<td>75</td>
</tr>
<tr>
<td>3-4.</td>
<td>Total intrapopulation genetic distance (diversity) in HIV-1 protease at baseline and after 24 weeks of combination therapy including a protease inhibitor</td>
<td>76</td>
</tr>
</tbody>
</table>
3-5. Total divergence in HIV-1 protease following 24 weeks of combination therapy including a protease inhibitor. ........................................................................77

3-6. Synonymous divergence in HIV-1 protease following 24 weeks of combination therapy including a protease inhibitor. ..............................................78

3-7. Nonsynonymous divergence in HIV-1 protease following 24 weeks of combination therapy including a protease inhibitor. .............................................79

3-8. Contribution of nonsynonymous mutations relative to synonymous mutations in HIV-1 protease. ........................................................................80

3-9. Median genetic divergence following 24 weeks of combination therapy including a protease inhibitor. ........................................................................81

3-10. Parsimony analysis of HIV-1 protease isolates prior to and following combination therapy ..........................................................................................82

4-1. Viral and immune response to combination therapy for HIV-1 infection. ..........101

4-2. Phylogenetic analysis of envelope V3 ................................................................102

4-3. HIV-1 protease genotype prior to combination therapy including a protease inhibitor. ..............................................................................................103

4-4. Presence of alleles with therapy specific mutations following 24 weeks of combination therapy. ..............................................................................104

4-5. Protease genotypic resistance segregates within CD4+ CD45 T cell subsets ......105

4-6. Distinctive PR alleles were identified in cultured cell subsets .......................106

4-7. Phylogenetic tree of protease alleles obtained from patient VFIS15. .............107

5-1. Strategy 1 for the production of gag-pro recombinant viruses .......................126

5-2. Linker Insertion Mutagenesis ........................................................................127

5-3. Map of recombinant viruses ..........................................................................128

5-4. Strategy 2 for the production of gag-pro recombinant virus .......................129

5-5. Replication kinetics by recombinant viruses containing natural gag-pro variant.131

5-6. Immunoblot analysis of Gag-Pol processing in Jurkat cells by recombinant viruses ..........................................................................................132

5-7. Amino acid sequence of Gag and PR alleles isolated from patient D1 prior to and following therapy that were used to construct recombinant viruses.........133
5-8  Supernatant p24 relative to LAI following infection by recombinant viruses.....134

5-9.  Median supernatant p24 following infection of X4 PBMC by gag-pro recombinant viruses.................................................................135

5-10. Median supernatant p24 following infection of R5 PBMC by gag-pro recombinant viruses........................................................................................................136

5-11. Median supernatant p24 following infection of macrophages by R5 gag-pro recombinant viruses.................................................................137
Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1) GAG-PRO GENETIC VARIABILITY IS RELATED TO THERAPY RESPONSE AND IMPACTS VIRAL FITNESS IN SPECIFIC CELL TYPES

By
Stephanie Leigh Rose

August 2002

Chair: Maureen M. Goodenow
Department: Pathology, Immunology, and Laboratory Medicine

Human immunodeficiency virus type 1 (HIV-1) is the etiological agent of AIDS. HIV-1 infection is treated with a combination of inhibitors against the viral enzymes reverse transcriptase and protease. Viral rebound during combination therapy may result from the accumulation of mutations in gag-pro. Drug resistant mutations can result in reduced replicative ability or fitness of the virus. A number of mutations in Gag and Protease [PR] may restore the replicative ability of mutant viruses.

The specific aims of this dissertation were 1) to evaluate the relationship of pretherapy protease polymorphisms with therapy response, 2) to analyze gag-pro sequences following 24 weeks of combination therapy for patterns indicative of therapy response, 3) to evaluate viral trafficking within the blood compartments using protease inhibitor resistance as a genotypic marker, and 4) to determine the functional impact of gag-pro polymorphisms on replication capacity in T cells and macrophages.
Therapy response by pediatric patients following 24 weeks of combination therapy was classified based on CD4$^+$ T cell percent and plasma viral load. Patients displayed viral and immune success; viral and immune failure; or discordant, viral failure and immune success, responses. Naturally occurring polymorphisms in HIV-1 pro that could diminish sensitivity to inhibitors were detected in 94% of this pediatric cohort. Pretherapy protease genotype was a significant variable in predicting outcome. The discordant response group relative to other response groups had a significant increase in the number of amino acid substitutions in PR and Gag following therapy. Both naturally occurring and therapy-associated genotypic gag-pro mutations modulated viral fitness. Recombinant virus with protease resistant gag-pro demonstrated reduced replicative ability relative to pretherapy gag-pro virus in peripheral blood mononuclear cells expressing CXCR4. In contrast, replication by protease resistant viruses in macrophages was unaltered. Using protease resistance as a genotypic marker, it was determined that HIV-1 trafficked from the plasma to the CD4$^+$ CD45RO cells and then to the CD4$^+$ CD45RA cell subset. The longevity of CD4$^+$ CD45RO cells and the persistence of early viral genotypes identified these cells as an important reservoir for drug resistant HIV in patients who achieve reconstitution of CD4$^+$ lymphocytes.
CHAPTER 1
INTRODUCTION AND BACKGROUND

An estimated 40 million people (2.7 million children) were infected with the human immunodeficiency virus type 1 (HIV-1), the agent known to cause AIDS (acquired immunodeficiency syndrome) in 2001 (102). HIV infection is characterized by a progressive loss of CD4\(^+\) and CD8\(^+\) T cells which may result from an accelerated destruction of mature T cells in the periphery as well as inhibition of de novo T cell production. The alarming rate of infection (about 14,000 new HIV-1 infections a day) demonstrates the need to identify ways to prevent and combat HIV-1 infection.

Antiretroviral therapy has dramatically impacted the natural history of HIV-1 disease progression (172). Treatment protocols including a combination of reverse transcriptase inhibitors and protease inhibitors have transformed HIV-1 infection from a progressive disease causing death to a chronic condition requiring life-long therapy. The advent of protease inhibitors in the mid 1990s brought hope that antiretroviral therapy would clear the viral infection. Unfortunately, these expectations were unfulfilled as it became evident that a latent reservoir of virus can persist in long-lived cells for many years (29,73,181). Replication of virus in the presence of suboptimal drugs leads to the emergence of viruses with a high level of resistance to the administered drugs and cross-resistance to related drugs. Suboptimal therapy in patients can be attributed to a number of factors including viral genetics, bioavailability of the drug in different tissues, individual pharmacokinetics, drug metabolism (52) and patient compliance.
Additional strategies for fighting HIV-1 infection include development of inhibitors to viral fusion and integrase, as well as vaccines (40). There has been considerable interest in the use of fusion inhibitors to prohibit virus from interacting with and infecting cells. Competitors for CD4, CCR5 and CXCR4 receptors have been tested in animal models but are currently unsuitable for use in humans (40). T-20 (pentafuside), an inhibitor that interferes with the viral gp41 to prevent fusion with the target cell, has been proven effective and Phase III clinical trials are in progress (112). As with other antiretrovirals, amino acid changes in the HIV-1 genome result in mutant viruses that escape T-20 (197). Thus the search continues for fusion inhibitors that have a broad spectrum of anti–HIV-1 activity coupled with a reduced propensity to develop resistance (22). Clinical development of integrase inhibitors has been delayed by both the limited knowledge of the integration process and by the requirements of the inhibitor to act on a restricted number of integrase molecules that catalyze only four chemical steps (37). Recent studies have identified compounds that may translate to clinically useful therapies, which could enhance the effectiveness of the inhibitor cocktail currently used (159).

Natural history studies and experiments in animal models have shown that specific immunity to HIV can be induced. The natural variability of HIV-1 has prevented the induction of broadly cross-reactive neutralizing antibody against primary isolates. A vaccine strategy that elicits CTL (cytotoxic T lymphocyte) response to diverse primary isolates has been shown to control virus infection (80). Nevertheless an effective preventive vaccine has yet to be designed (80). These novel chemo-
immunotherapeutic strategies are limited and will require continued exploration and understanding of the currently used therapies.

The standard of care in this country includes a combination of inhibitors to the viral enzymes reverse transcriptase and protease. The aims of this study were to evaluate 
\textit{gag-pol} polymorphisms genetically and functionally to elucidate the impact of mutations on clinical response and viral replication. My hypothesis was that amino acid polymorphisms in \textit{gag-pol} influence therapy response by impacting the virus’s replicative ability. One possibility is that modulation of replication may be cell-type specific. Understanding the functional impact of mutations in \textit{gag-pol} will lead to the design of more effective therapies. My dissertation consists of four specific aims that examine my hypothesis.

\textbf{Specific Aims}

\textbf{Specific Aim 1: Evaluate the Relationship of Pretherapy Protease Polymorphisms with Therapy Response}

HIV-1 displays considerable genetic diversity resulting from the high mutation rate of reverse transcriptase, rapid viral turnover, and the selective pressure of the host immune response (35). Polymorphisms associated with reduced sensitivity to protease inhibitors have been identified in therapy naïve patients. However, the clinical relevance of such mutations was inconclusive (21,125,183,199). This study evaluated pretherapy protease alleles in a cohort of 26 patients for polymorphisms associated with decreased sensitivity to the administered protease inhibitor. Results from Specific Aim 1 are presented in Chapter 2.
Specific Aim 2: Analyze Gag-pro Sequences Following 24 Weeks of Combination Therapy for Patterns Indicative of Therapy Response

Escape of HIV-1 from protease inhibitor therapy results from a combination of variation within protease and the cleavage sites. The genetic makeup of the virus with respect to therapy response has not been studied. The goal of this specific aim was to establish whether the accumulation of gag-pro mutations are distinguishable among each of the three therapy response groups. The results from Specific Aim 2 are presented in Chapter 3.

Specific Aim 3: Evaluate Viral Trafficking within the Blood Compartments Using Protease Inhibitor Resistance as a Genotypic Marker

The trafficking of virus within the blood compartments has been virtually unstudied. Studies in the Goodenow lab established that alleles within peripheral blood mononuclear cells (PBMC) are heterogeneous (183). Analysis of additional patients in this dissertation confirmed that, after 24 weeks of therapy, a diverse population of protease alleles, ranging from inhibitor sensitive to highly resistant, existed within the PBMC. One possible explanation for this variability is segregation of protease alleles within the CD4$^+$ T cell subsets. To test the hypothesis that protease alleles segregate by blood compartment, viral genotypes in the plasma along with the CD4$^+$ CD45RA and CD4$^+$ CD45RO T cell subsets were evaluated. This study provides a unique perspective on the half-life of cells and the identification of a viral reservoir. Results for Specific Aim 3 are presented in Chapter 4.

Specific Aim 4: Determine the Functional Impact of Gag-pro Polymorphisms on Replication Capacity in T Cells and Macrophages

Natural polymorphisms have a profound effect on proteolytic processing in a bacterial expression system(77). Differences in processing phenotypes might be
expected to impact viral replication. Genetic analysis was used to identify alleles with polymorphisms/mutations for evaluation of their functional impact on viral replication. To directly evaluate the impact by genetic polymorphisms in protease and Gag on viral replication, recombinant viruses were constructed by replacing \textit{gag-pro} in HIV\textsubscript{LAI} with the \textit{gag-pro} regions from patient isolates. The results for Specific Aim 4 are presented in Chapter 5.

\textbf{Significance}

The best treatment available involves use of inhibitors against two enzymes, reverse transcriptase and protease, which are both vital to the viral life cycle. Treatment of HIV-1 infection is currently limited to 15 approved drugs that, while effective, have significant limitations (40). This study will contribute knowledge about the interaction of the viral genetic characteristics and protease inhibitors.

\textbf{Human Immunodeficiency Virus}

The human immunodeficiency virus type 1 (HIV-1) was first isolated in 1983 by Barre-Sinoussi \textit{et al.} at the Pasteur Institute in Paris (8). HIV-1 is a single strand positive sense virus belonging to the Retroviridae family. The 9.2 kb viral genome consists of three main genes; \textit{gag}, \textit{pol}, and \textit{env}, as well as a number of accessory genes (Figure 1-1).

\textbf{HIV-1 Life Cycle}

A virion must interact with the target cell receptor complex via the viral envelope proteins to gain access to the cell (Figure 1-2). Following entry, the viral RNA genome is reverse transcribed to produce double-stranded linear DNA, which is translocated to the nucleus where it is integrated into the host genome. The viral genome is transcribed and mRNAs are translocated to the cytoplasm where translation of the viral proteins required for progeny production occurs. Virions assemble at the plasma membrane and bud from
the host cell. Particle release alone does not render the virion infectious; protease activity is essential to render viral infectivity (117).

**Targets of HIV-1 Infection**

Host genetic factors in combination with viral properties determine the susceptibility of an appropriate target cell for HIV-1 infection as well as the replicative potential of the virus in the cell resulting in an overall productive infection. Cells that express CD4: T cells, macrophages, monocytes and other phagocytic cells are targets for HIV-1 infection. There is evidence to suggest that CD8 T cells can also be infected, especially in the late stage of disease.

The majority of HIV-1 infected cells are CD4\(^+\) T lymphocytes (4,5,114,177). CD4\(^+\) T lymphocytes expressing CD45RO are infected to greater levels than those expressing CD45RA (205,217). Although macrophages are a target for HIV-1, only small portions of tissue macrophages are infected (29,140,214).

The life of target cells impacts the disease progression of HIV-1 infection. Monocytes circulate in the periphery for less than a day before differentiating into macrophages in tissues (144). The half-life of macrophages has been estimated at approximately 2 weeks (235). The short half-life of these cells reduces the likelihood that they serve as a reservoir for latent HIV-1.

The half-life of the CD4\(^+\) CD45RA and CD45RO cells in human has been shown to vary dramatically. Resting memory CD4\(^+\) CD45RO T cells have an estimated half-life of between 6 to 44 months, while the half-life of recently infected resting CD4\(^+\) T cells is less than 1 week (15). In contrast, CD45RA cells may persist for up to 10 years (97,146). The goal of the immune system is to maintain the ability to mount responses against new as well as past threats to the body over the course of a lifetime. Because HIV-1 is stably
integrated into T cells with extended half-lives, the virus may be carried in these often quiescent cells for an extended period of time.

**CD4⁺ Cell Subsets are Classified by CD45 Isoform Expression**

The common leukocyte antigen, CD45, is expressed on the surface of T cells in different molecular weight isoforms, generated by alternative splicing of three exons (A, B, C) encoded by the leukocyte common antigen gene. The low molecular weight isoform, CD45RO, and the high molecular weight isoform, CD45RA, can be used to identify memory and naïve cells, respectively. A shift in expression from CD45RA to CD45RO has been considered an irreversible postthymic differentiation event (31,32,206). Studies challenging this position have established in animal models that both subsets of cells generate cells of the opposite phenotype (11,12,85), demonstrating reversion from a CD45RO phenotype to CD45RA is possible. These data suggest that CD45 molecule expression reflects the state of activation, rather than differentiation. It is unclear whether data from the animal model are equivalent phenotypically or functionally to reversion in humans. Although there are limitations to studies in humans, a growing amount of evidence suggests that memory cells may indeed re-express CD45RA (2,146,200,241)

Reversion would indicate that two categories of memory cells exist. The first category expresses CD45RO, have a relatively rapid turnover time, and respond easily to recall antigens. The second category expresses CD45RA and survive for long periods without cell division (13). Additional studies support the idea of immune memory residing in CD4⁺ CD45RA cells as well as the CD45RO cells (79,196).

Two subsets of CD45RO⁺/CD45RA⁻ memory cells were identified by the expression of chemokine receptor CCR7. Central memory cells identified by CCR7
expression are a clonally expanded antigen-primed population that lack inflammatory and cytotoxic function. Central memory cells travel to secondary lymphoid organs and upon a secondary challenge can efficiently stimulate dendritic cells, help B cells and generate a new wave of effector cells. Effector memory cells (CCR7\(^{-}\)) are endowed with various effector functions and represent a readily available pool of antigen primed cells, which can enter peripheral tissues to mediate inflammatory reactions or cytotoxicity (201). Both memory subsets persist in vivo for up to ten years and their relative proportions do not change after a booster immunization (201).

**Envelope**

The *env* gene encodes for a 160kDa glycoprotein precursor, gp160, which is cleaved by cellular enzymes to yield gp120 (SU), an external glycosylated hydrophilic polypeptide, and gp41 (TM) a membrane-spanning protein (7,74,138). Together gp120 and gp41 form an oligomeric knobbed spike on the surface of the virion (37), which mediates viral binding to the CD4 receptor located on CD4\(^{+}\) T-lymphocytes, macrophages, and dendritic cells (45,133). The gp120 protein consists of 5 variable regions (V1-V5) and four constant regions (C1-C4) (150,245). Viral gp120 contains binding sites for both CD4 and either or both CXCR4 (X4) and CCR5 (R5) chemokine receptors (34,119,133,139,221). The variable regions do not play a direct role in binding with CD4\(^{+}\) (61,192), while mutations within the constant region influence the CD4 binding ability (119,165). The third hypervariable loop of envelope, referred to as V3, impacts cellular tropism and coreceptor use (28,98,210). Single amino acid polymorphisms within V3 can alter the tropism of HIV-1 (213).

Tropism is the ability of an HIV isolate to replicate in either T-cell lines (MT-2) or macrophages. Isolates that replicate to high levels in CD4\(^{+}\)T cells, induce syncytium,
utilize X4, and typically display a high V3 net charge ($\geq +5$) are frequently referred to as T-cell tropic (56). Isolates that replicate well in macrophages (or M) tropic isolates are generally unable to replicate in CD4$^+$T cell lines. M-tropic isolates utilize R5, fail to induce syncytia, and display a low V3 net charge ($\leq +4$). Isolates that are capable of replicating in both T-cell lines and macrophages are referred to as dual tropic (38,212). Generally NSI (non syncitia inducing) variants are associated with acute infection and display low V3 net charge ($\leq +4$), while SI (syncytium inducing) variants emerge late in disease and display high V3 net charges ($\geq +5$) (42,56).

A coreceptor is required for infection by HIV-1. The CXCR4 coreceptor (72) promotes fusion of HIV with a cell. M-tropic strains of HIV-1 use the CCR5 coreceptor to infect cells (1,53,59,63,233). HIV-1 phenotype can be classified by the ability to use either or both CXCR4 (X4) and CCR5 (R5) coreceptors. Although CXCR4 and CCR5 are the dominant coreceptors used by HIV-1, other coreceptors have been identified but their physiological relevance is unclear.

**HIV-1 gag-pol**

**Reverse Transcriptase**

The HIV-1 pol gene encodes for reverse transcriptase (RT), an RNA and DNA directed DNA polymerase. RT functions early in the viral life cycle to produce a DNA copy of the RNA genome, which must be integrated into the host genome for productive infection. Unlike other DNA polymerases, RT lacks a proofreading or 3’ to 5’ exonuclease, activity (10). The rate for misincorporation of bases is 1 in $10^4$, which equates to one misincorporated base per genome (227). The low fidelity of RT contributes to the high genetic variability of the virus.
RT was the first target of antiretroviral drugs. This class of drug includes nucleoside, nucleotide and non-nucleoside inhibitors. Nucleoside and nucleotide inhibitors compete with the naturally occurring nucleosides for incorporation into the DNA genome. Incorporation of these inhibitors causes chain termination, preventing the production of a complete provirus. Zidovudine, didanosine, zalcitabine, stavudine, abacavir, and lamivudine are nucleoside inhibitors and adefovir is a nucleotide inhibitor. Non-nucleoside inhibitors noncompetitively inhibit reverse transcriptase by binding to the enzyme’s catalytic site. The non-nucleoside RT inhibitors include nevirapine, delavirdine, and efavirenz (248). Mutations that confer resistance to RT inhibitors impact the usefulness of this class of drug in the battle against HIV-1 (204).

Protease

The HIV-1 protease is classified as an aspartic protease because of the presence of the highly conserved catalytic site motif, Asp-Thr-Gly (117,230). Protease is translated as part of the p160 Gag-Pol polyprotein resulting from a -1 frameshift near the end of the \textit{gag} open reading frame (99). Two 99 amino acid protease subunits must dimerize to produce a proteolytically active protease (124,160,242) (Figure 1-3).

The protease subunits are linked by a four-stranded antiparallel $\beta$ sheet, which includes both the amino and carboxyl termini of each subunit. A cleft between the two subunits forms the substrate binding domain. Loops from each monomer interdigitate to form the hydrophobic core of the enzyme, where the catalytically important aspartic acids are located (225). A flexible flap, consisting of an antiparallel $\beta$-hairpin with a $\beta$-turn, extends over the substrate binding site opening to allow the substrate or inhibitor access to the active site and closing down when catalysis or inhibition occurs (83,83).
Protease activity is dependent on the amino acid sequence of the enzyme itself (131), in addition to determinants that map to the Gag-Pol polyprotein (62,134,254). Mutation or inhibition of protease abolishes production of infectious viruses (106,188,244). Protease functions late in the life cycle by processing the Gag and Gag-Pol polyproteins into a number of functional and structural units (Figure 1-1) (168,216). Whether protease becomes active immediately following or just before viral particles separate from infected cells is still debated (237). Proteolytic cleavage of the Gag polyprotein allows for assembly of the structural proteins resulting in the phenotypic differences between a mature and immature virion. Noninfectious (immature) viral particles, which have a spherical structure and are characterized by an electron-lucent center, are released from the cell (Figure 1-4 C). Mature (infectious) virions have the same spherical structure with an electron dense core (Figure 1-4 D)(225), which results from the organization of the structural proteins (Figure 1-5)(238). Changes in the Gag-Pol protein have a significant affect on the assembly and stability of the particle (174). It is possible that maintenance of an immature phenotype results from mutations in protease or Gag that modulate the enzyme's ability to process.

Protease’s role early in the viral life cycle is debatable. Evidence from several groups suggest that protease cleaves nucleocapsid post-infection, influencing the formation or transport of the preintegration complex (3,109,141,157). Yet, others maintain that protease is not involved in the early stages of the viral life cycle (100,105,234).

**Gag Cleavage Sites**

The *gag* open reading frame is located downstream of the 5' LTR and encodes all of the required structural proteins. The translation products are polyproteins produced as
either p55 Gag or p160 Gag-Pol (128). Approximately 20 copies of p55 Gag are produced for every p160 Gag-Pol to ensure the proper quantity of the structural proteins relative to enzymatic proteins (99). The structural proteins encoded by gag include p17 (matrix), p24 (capsid), p7 (nucleocapsid [NC]), and p6(238), which are proteolytically cleaved from the polyprotein into the respective polyproteins.

The cleavage sites recognized by protease are categorized into three groups: Class I, Class II and other. Class I sites (E, F, and A) have an aromatic-Pro at the scissile bond (88,178). Class II which includes sites B, C, and D is characterized by the presence of hydrophobic-hydrophobic bond at the cleavage site and a glutamine or glutamate at position p2' (189). The "other" group (sites G and H) does not have features in common with the first two groups. An additional cleavage site has also been postulated to occur immediately adjacent to the gag-pol frameshift locus in HIV-1. (81). Cleavage sites recognized by protease consist of three to four amino acids surrounding the scissile bond (104,175). Amino acids surrounding the bond are named as follows: P1, P2, … from the scissile bond to the N terminus and P1', P2',… from the scissile bond to the C terminus (203).

The order of processing at each site (first to last) E, C, D, A, B was determined using synthetic substrates (121) and confirmed thru additional experiments (18,188). Maturation of HIV-1 is a sequential process controlled by the cleavage at individual sites (244). Such a mechanism could imply that, even if specific mutations do not alter protein folding or interaction domains, a change in the rate of processing at individual cleavage sites could disrupt the regular virion maturation process (244). Accurate and complete processing is essential to the formation of infectious, morphologically normal virions.
Natural Genetic Variation of \textit{gag-pol}

A cross-sectional analysis of natural variants in Gag cleavage sites indicated that cleavage sites B, D', and F were predominantly identical to the HIV\textsubscript{LAI} (≥ 73%) (9). Although the D site showed less identity to HIV\textsubscript{LAI} (43%), it was similar to molecular clone HIV\textsubscript{NL4-3}. In contrast, the C cleavage site was unique in that a dominant sequence was not identified (only 19% were identical to HIV\textsubscript{LAI}) and it displayed the most variability. The C site sequence was specific to individuals, thus intrapatient heterogeneity is low while interpatient heterogeneity was the rule (9).

The role of natural polymorphisms in \textit{gag-pol} on response to combination therapy has recently received a great deal of attention. Pretherapy polymorphisms have been identified in protease amino acid positions 10, 15, 13, 12, 20, 33, 35, 36, 37, 41, 46, 62, 63, 77, and 93 (207),(226). Several studies have reported the pre-existence of polymorphisms associated with resistance to protease inhibitors in therapy naïve patients (9,120,125,184). Baseline mutations with as little as fourfold increase in baseline resistance were sufficient to compromise treatment with two protease inhibitors (87). Natural mutations in HIV-1 subtype O have been shown to decrease protease inhibitor effectiveness (47).

\textbf{Processing of Natural Polymorphisms in the Bacterial Expression System}

A series of experiments conducted in the laboratory of Dr. Ben Dunn, in collaboration with Dr. Maureen Goodenow, evaluated the role of natural polymorphisms in protease and identified determinants in Gag that modulate protease processing activity in a bacterial expression system(77). Three phenotypes were identified based on the appearance of Gag intermediates and accumulation of mature p24\textsuperscript{CA} (Figure 1-6 A). Rapid, complete processing with virtually no accumulation of Gag intermediates, was
termed a Type I processing phenotype. D1.10 Gag-Pol displays a Type I processing phenotype. Type II processing (intermediate/complete) phenotype was displayed by D2.21 and was characterized by transient accumulation of Gag intermediates with the appearance of p24 delayed for at least 20 minutes post induction. HIV\textsubscript{LAI} processing of Gag-Pol polyprotein displayed a type II processing phenotype. Type III processing was exhibited by MD.03 and was characterized as incomplete based on a failure to produce detectable p24. Results from this series of experiments suggested that processing at the C-cleavage site and accumulation of p40 had an impact on subsequent PR cleavage at A- or B-sites (77).

The direct impact of Gag sequences on processing by protease was assessed by site-directed mutagenesis and construction of chimeric Gag-Pols, in which the Gag region of one construct was paired with the Pol region of a different construct. Optimal protease activity occurred when Gag and Pol regions were derived from the same gag-pol allele (Figure 1-6 B). Heterologous Gag regions generally diminished rates and extent of protease processing. Natural polymorphisms in novel positions in p7\textsuperscript{NC} and the C-cleavage site have a dominant effect on protease processing activity (77).

Gag-Pol alleles tested in the bacterial expression system were used for construction of replication competent recombinant viruses. Data from such experiments are discussed in Chapter 5 of this dissertation.

**Protease and the Selective Pressure of Drug Therapy**

Since protease activity is crucial to the viral life cycle, this enzyme became an obvious target for therapeutic intervention. Ritonavir, indinavir, saquinavir, and nelfinavir, all used in the clinical setting, interact with protease’s active site, blocking access for the peptide and thus preventing polyprotein processing (249). One
problematic outcome with current enzymatic inhibitor therapies is the emergence of drug-resistant variants. Virus mutants resistant to protease inhibitors used in the clinical setting involve an accumulation of multiple amino acid substitutions primarily in the protease region (Figure 1-7) (39,41,82,130,152,202,204). Because many of these mutations are selected for the ability to decrease inhibitor binding in the active site, they also affect substrate binding and potentially substrate specificity (67).

Since these inhibitors have identical function, mutations within the binding sites confer cross-resistance to multiple inhibitors (41,229). Mutations outside of the substrate binding domain and active site have also been associated with decreased susceptibility to protease inhibitors.

Changes located close to the enzyme active site can decrease the protease catalytic activity and replicative capacity (19,82,130). In vivo, selection for resistance to protease inhibitors results in a significant impairment of HIV replicative capacity due to a decrease in protease cleavage efficiency (252). Primary mutations, which cause decreased inhibitor binding, are frequently the first mutations selected (92). Single cycle virus assays demonstrate that viruses present following therapy, which contain primary mutations, have a lower infectivity relative to protease alleles present at the initiation of therapy (166). Secondary or compensatory mutations in protease contribute to drug resistance by restoring enzymatic activity (94).

Implication of Gag Cleavage Sites Polymorphisms

Since mutations in protease affect processing of Gag proteins, it is plausible that changes in the target protein could assist the mutant protease. Mutations identified in the D’ and D cleavage sites serve in a compensatory manner to restore (at least partially) the processing of the Gag precursor in vitro (62). Cleavage site mutations similar to those
described by Doyon et al. were subsequently identified in patients on protease inhibitor therapy (254).

Substitutions of alanine to valine at position p2 (amino acid position 431) in the D’ (p7/p1) cleavage site and leucine to phenylalanine at position p1’ (amino acid 449) in p1/p6 (D) cleavage site have been associated with viral failure to protease inhibitors (6), while patients with the proline rich motif in the p6 Gag protein were less likely to experience virological failure (108). The A431V mutation is reported to occur in up to 29% of patients with resistant proteases (43,116). Both A431V and L449F have been associated with increased viral fitness for viruses with protease inhibitor resistance (134,252).

A number of mutations in protease have been associated with the p7/p1 (position 431) cleavage site mutation, including M46I, V47A, and V82A/F/T (6,43,116). Proteases with multiple mutations associated with protease inhibitor resistance demonstrated up to a fourfold change in the preference for a valine relative to alanine at position p2 of the D’ cleavage site when tested on individual peptide substrates (46). Gag mutations H219Q and R409K have been associated with multi-drug resistant HIV-1 variants and although these changes can restore replicative capacity, they do not correct for altered Gag cleavage patterns (76). Mutations in Gag may represent a mechanism by which severely compromised, drug-resistant viral strains can increase fitness. In addition to contributing to efficient proteolytic processing Gag mutations play a role in escape from cytotoxic T lymphocyte responses (111).

**Response to Therapy**

Response to antiretroviral therapy can be classified based on viral and immune outcomes. Viral and immune successes [VSIS], which occurs in about 30% to 40% of
treated pediatric patients, are characterized by a reduction of viral load in combination with an increase in CD4\(^+\) T cells. Viral failure immune failure [VFIF] occurs in about 15% to 20% of pediatric patients (69,184). Since the inception of this study we and others have identified an unexpected discordant response to therapy which occurs in about 25% to 40% of pediatric patients. This response group is characterized by a paradoxical increase in CD4\(^+\) T cell count in combination with a high viral load (69,101,184). The discordant response can persist for years in spite of high viral burdens that would predict disease progression. (101). Discordant responses have also been identified in variety of HIV-1 infected adults cohorts (50,51,107,126,129,145,190,195).

Despite treatment with inhibitors viral replication may occur in body sites lacking adequate exposure to antiretrovirals (136) such as the brain or male genital track (68). Additionally residual virus replication has been detected in patients who display a successful response to highly active antiretroviral therapy (60,75,136,194), therefore all patients receiving therapy are at risk for developing mutations associated with reduced sensitivity to the administered protease inhibitor

**Compartmentalization of Drug Resistant Viruses**

HIV-1 variants with protease resistance mutations are detectable in the plasma before PBMC (110,115,183,198). Following viral rebound divergence at both the nucleotide and amino acid levels was higher in plasma than in PBMC (240). However, mutations in plasma and PBMC showed a substantial correlation in highly treated patients (55). Despite the correlation of protease alleles between these compartments a minority of mutations associated with previous therapy can be detected in the PBMC and not the plasma (55). The source of virus genotypically analyzed may dramatically impact the results of a study. This is most relevant in a clinical setting, if plasma genotype is
used for screening before therapy switch critical mutations from previous therapies may be missed.

**Viral Fitness and the Dynamics of Diversity**

Fitness can be defined in strict Darwinian terms as the selective advantage for replication among viral variants competing in a particular environment. The definition of fitness is less strict in the current literature surrounding HIV-1, and can be defined as the competitive advantage or disadvantage relative to wild-type or other mutants in the presence or absence of drug (135). Mutations associated with protease inhibitors reduce viral fitness (44,137,252). A number of mutations in protease and gag have been associated with restoring fitness in drug resistant variants (134,252).

High rates of replication have significant implications for the rapid appearance and accumulation of therapy-resistant virions. HIV infection is a dynamic process in which the production of newly infected cells and free virions serve as replacements for the degradation of present virions and infected cells (35,93,93,243). Because HIV-1 exists as a quasispecies, the most fit variant becomes dominant (95). This means that the frequency of a specific set of mutations, which may otherwise be detrimental to viral replication, will increase provided the effect of the mutations is advantageous (increased fitness) under the specific conditions of that replication cycle (36). Since complete suppression of viral replication is not always possible due to drug bioavailability or patient compliance, there is a low level of viral replication in which resistant virions can arise.

**Viral Reservoirs and Latency**

A viral reservoir can be defined as a cell type or anatomical location in which a replication competent form of HIV-1 accumulates and persists with more stable kinetics
than the main pool of actively replicating virus (15). As a consequence of reservoir stability, viral species arising at various times in the infection should be present in the reservoir. Recent as well as archival sequences should not only be detected but both should be capable of being released (15). A number of cell types have been implicated as viral reservoirs, including macrophages, resting memory CD4\(^+\) T cells, resting naïve CD4\(^+\) T cells, recently infected resting CD4\(^+\) T cells, follicular dendritic cells, B cells, and monocytes.

Latency describes infected cells that are not producing virus but retain the capacity to do so under appropriate conditions (15). Current modeling suggests that most latently infected cells belong to the memory subset of resting CD4\(^+\) T cells (15). It is estimated that less than 0.01% of resting memory cells carry integrated HIV-1 DNA (29). The latent reservoir of resting CD4\(^+\) cells detected in both children and adult has been shown to harbor drug resistant protease alleles (186). The latent reservoir is extremely stable with a mean half-life of over 43 months (73), which would then require an average of more than 60 years to eradicate this compartment. The stability of the latent reservoir in resting memory T cells translates to a lifetime persistence of virus for most patients receiving current combination therapy.

**Phylogenetic Analysis**

The goal of phylogenetic analysis is to determine the genetic relationship among a set of given sequences. Sequence analysis includes a measure of nucleotide changes that result in amino acid changes (non-synonymous substitutions [NS]) and of nucleotide changes that do not result in amino acid changes (silent or synonymous [S]). Computer based algorithms can translate the relative occurrence of NS and S substitutions into genetic distance (162). Methods of analyzing genetic distance include Kimura-2-
parameter, which assumes that the rate of transitions will outnumber transversions and the Jukes-Cantor, which infers that the rate of nucleotide substitution is equal for all four nucleotides (103,113).

Diversity is the genetic difference within a specific group of sequences. For example, if you were comparing protease among all sequences obtained from the CD45RO cells at a given time that would be considered diversity. Divergence is the genetic distance between two distinct groups of sequences. For example, if you were comparing protease sequences obtained from the CD45RA cells with those from the CD45RO cells that would be considered divergence.

Phylogenetic trees are models of evolutionary history for a group of sequences. Internal nodes represent hypothetical ancestors, and branch lengths represent the amount of evolution that occurs. Rooted trees have direction, since a node is identified as the root from which all other nodes descend. The closer a node is to the root of the tree the older it is in time, implying that the node closer to the root of the tree is the ancestor of the node further away from the root (171).

Two methods of constructing phylogenetic trees are discussed and utilized in this dissertation. Each method can be used to create a tree based on the nucleotide or amino acid sequence. The neighbor joining method of tree building, a distance based method, uses the overall genetic distance between pairs of sequences without incorporating the qualitative information about the differences (90,91). A limitation to neighbor joining analysis is that the fit between the tree and data is not optimized, however it is a good method for estimating the minimum evolution tree (171).
Parsimony analysis, a character based method, looks at each site and uses the pattern of variation to produce a plausible sequence of changes that relates the sequence to a common ancestor (223). The goal of this method is to reconstruct the evolution of a site with the constraint of invoking the fewest possible evolutionary changes. One advantage of the parsimony method is that few assumptions are made about the evolutionary process. In the absence of strong selective pressure, the rate of synonymous substitutions in \textit{gag} and \textit{pol} exceed that of nonsynonymous substitutions, indicating that these genes are subject to purifying selection (96).

**Bootstrapping.** Bootstrap analysis verifies the reliability of the inferred tree by taking multiple samples from the population being studied and comparing the estimates obtained from the different samples. This process is repeated many times and the proportion of replications in which a given sequence cluster appears is calculated (161). The bootstrap value represents how often each branch was represented with the same topology on the re-sampled trees. Bootstrap values are typically a conservative measurement (161). Bootstraps values of 70 or higher may correspond to a 95% confidence level and is typically considered the cut-off for a significant branch (90).
Figure 1-1. HIV-1 proviral genome. The main genes, *gag*, *pol*, and *env*, as well as the accessory genes are represented above. Two identical long terminal repeats (LTR) flank the viral genome. Translation of *gag-pol* yields a polyprotein that is proteolytically processed at specific cleavage sites (designated with letters) to produce the structural [p17 (matrix), p24 (capsid), p7 (nucleocapsid), p6 and p6*], and enzymatic [PR (protease), RT (reverse transcriptase) and IN (integrase, not shown)] proteins.
Figure 1-2. Lifecycle of HIV-1. The main steps of the viral lifecycle are outlined. 1, Binding of HIV-1 gp120 to CD4, and a coreceptor (either CCR5 or CXCR4) followed by fusion with the cell and release of HIV-1 RNA from the protein core to the cytoplasm. 2, Reverse transcription of viral RNA into double-stranded DNA. 3, Translocation of viral DNA to the nucleus, integration into the host genome and transcription of viral RNA. 4, Translation of viral gene products. 5, Viral assembly at the plasma membrane and budding of the HIV-1 virion. 6, Protease processing to produce structural and enzymatic viral proteins resulting in mature virions.
Figure 1-3. Ribbon diagram of the three-dimensional structure of HIV-1 protease bound to a protease inhibitor [PI]. The flaps (navy blue) form a roof over the active site, which is bound by a PI. Aspartyl groups are indicated by stars (*). The location of amino acids positions 10, 36, 63, 71, 77 and 82, identified as pre-existing polymorphisms, are designated. Coordinates for protease bound to inhibitor from Kempf et al. (184), figure designed by Jose Clemente. The amino acid sequence of HIV$_{LAI}$ protease is shown with bars indicating the amino acid positions that comprise the catalytic domain, flap region and psi loop.
Figure 1-4. Electron micrograph of HIV-1 particle assembly and release. A, Assembly of viral proteins at the plasma membrane. B, HIV-1 particle budding from the cell membrane. C, Immature virions characterized by a spherical shape and an electron lucent center are released from the cell. D, Mature electron dense HIV-1 particle. (225)
Figure 1-5. Models for retroviral structure. Top, The immature HIV-1 virion. Gag and Gag-Pol proteins are shown in different colors to suggest the domains corresponding to the mature proteins formed from the precursors. The SU (gp120) and TM (gp41) envelope proteins are shown jutting out of the lipid membrane. Bottom, The mature HIV-1 virion. Protease has proteolytically cleaved the Gag and Gag-Pol proteins, releasing the individual proteins and allowing for condensation of the viral core into a cone shape (238).
Figure 1-6. Processing phenotypes of Gag-Pol polyprotein. A, Immunoblot analysis of natural variant polyprotein processing in a bacterial expression system. Lanes A p54^Gag^, B, purified p24^{CA}; C, no induction control. Type I, rapid kinetics and complete processing to p24^{CA} displayed by D1.10 Gag-Pol; Type II, intermediate kinetics and complete processing displayed by D2.21 Gag-Pol; Type III, intermediate kinetics but incomplete processing with no detectable p24^{CA} after 60 minutes displayed by MD.01 Gag-Pol. B, Processing of chimeric gag-pol constructs. Lane A and B, p54^Gag^ and p24^{CA}, respectively. Blot 1, D1.10 Gag with D1.10 protease; blot 2, D2.21 Gag chimera; blot 3, MD.03 Gag chimera (77).
Figure 1-7. Amino acid mutations in protease associated with reduced sensitivity to protease inhibitors. Numbers and letters at the top of the figure are amino acid positions in the reference strain HIV\textsubscript{LAI}. Protease inhibitors currently approved for use in the adult population: AMP (amprenavir); IDV (indinavir); NFV (nelfinavir); RTV (ritonavir); and SQV (saquinavir). Highlighted positions indicate that the presence of a mutation in this position would reduce sensitivity to the given drug. Note that mutations in some positions such as I84 reduce viral sensitivity to multiple inhibitors.
CHAPTER 2
HIV-1 PROTEASE GENOTYPE PREDICTS THERAPY RESPONSE

Introduction

Antiretroviral therapy using protease inhibitors (PIs) in combination with reverse transcriptase inhibitors (RTIs) has produced a dramatic impact on the natural history of HIV-1 disease by delaying progression to AIDS through the control of viral replication and by reversing or preventing immunodeficiency (54,172). Plasma virus levels and CD4 T cell counts are intermediate markers for disease progression that are generally applied as prognostic measures of therapy outcome and provide the basis of current guidelines for use of antiretroviral agents in HIV-infected adults and children (24,26,142,149,151,173). Unfortunately, a number of HIV-infected patients who receive PI combination therapy fail to achieve or sustain undetectable plasma viral levels despite prognostic markers that would predict successful therapy response (23,24,50,84,149,164,190). In addition, both adults and children can reconstitute and sustain improved CD4 T cells counts when treated with combination therapy, even though viral levels rebound (69,127,190,219).

Designing effective combination therapy from the multiple choices of PI and RTIs that are available necessitates a more diverse repertoire of prognostic markers for successful therapy outcome. Genetic characteristics of protease (PR) and reverse transcriptase (RT), which are targets in the viral genome for drug action, provide an additional level of prognostic value in predicting treatment outcomes (30,185,199). Switching antiretroviral drugs based on RT and PR genotype increases the likelihood of
successful response for treatment-experienced patients (64,132). In contrast, little data to access the value of RT and PR genotyping for treatment of therapy naïve patients are available (92).

Theoretical calculations based on rate of virus mutation, amount of replication, and levels of virus in vivo estimate that HIV-1 quasispecies with at least one, or even two, amino acid substitutions, which could reduce drug efficacy, exist in treatment naïve individuals (35,36). Assessments of therapy-naïve patient populations indicate the occurrence of natural polymorphisms in either RT or PR that could reduce drug sensitivity when therapy is initiated (9,120,125). Consequences of low levels of resistance may be difficult to detect in in vitro assays that measure biochemical activity of PR (21). Natural polymorphisms in PR might reduce sensitivity to drugs and impact viral responses to combination therapy setting the stage for rapid emergence of viruses with multiple drug-resistance mutations (41). Consequently, pretherapy resistance may provide an additional factor to account for failure to sustain suppression of virus in patients who appear compliant with medications (52).

A prospective study was designed to evaluate if genotypic determinants in PR or RT would improve the prognostic value of pretherapy viral and immune variables in predicting therapy outcome. The study focused on a cohort of HIV-infected children and adolescents, who were naïve to PI treatment, but were at greatest risk for combination therapy failure due to advanced disease and high viral burden. This work has been published in The Journal of Infectious Diseases, volume183 in 2001 (184).
Materials and Methods

Study Patients

The study cohort was comprised of HIV-infected children and adolescents who were enrolled prospectively between January 1996 and October 1999 in a treatment protocol that involved combination therapy with one PI and one or two RTIs (26). Eligible patients were between one and eighteen years of age, naïve to PI therapy, and immune compromised (CDC Immune Stage 2 or 3) with plasma viral levels greater than 4.0 log10 copies per ml as measured by the Amplicor assay with 400 copies/ml as the lower limit of detection (Roche Molecular Systems, Somerville, New Jersey). Prior therapy with RTIs was permitted if patients were naïve to at least one RTI in their new combination therapy protocol. Among a population of about 100 infected children, twenty-six patients fulfilled the inclusion criteria and also demonstrated compliance with combination therapy. Seventeen patients (65%) received combination therapy with ritonavir (RTV), three (12%) with nelfinavir (NFV), and six (23%) with indinavir (IDV). Selection of PI was based on availability of pediatric formulations, as well as ability of the patient to swallow capsules or tolerate liquid formulations and to adhere to the treatment regimen. Optimal drug dosing was based on pharmacokinetic studies of PIs in children (122,154,155). Adherence was carefully monitored at each study visit by measurement of returned medications and interviews with the patient and/or the family. Adherence to therapy was defined as evidence that a patient received correctly over 90% of the prescribed doses of medications (147,246).

The cohort had a median age of eight years (range one to seventeen years) and included eight adolescents. Thirteen patients were females and thirteen were males. Thirteen children (50%) were African Americans, twelve (46%) were Caucasian, and one
(4%) was Asian American. A majority of the cohort (24/26, 92%) was infected perinatally either through maternal transmission (20/24) or HIV contaminated blood products (74). Two adolescents were infected by sexual transmission approximately two years before initiation of therapy. All patients were infected prior to introduction of PI therapy into the adult population. Most patients displayed symptoms of HIV infection (CDC clinical stage B or C) (20/26, 77%) and immune suppression (CDC immune stage 2 or 3) (25/26, 96%) (25). Pretherapy median CD4 T cell counts were 147 cells per ul (25th to 75th quartile ranges from 41 to 309 cells/ul), median CD4/CD8 ratio was 0.15, and median HIV plasma RNA levels were \(5.15 \log_{10}\) (25th to 75th quartile ranges from 4.5 to 5.7 \(\log_{10}\)) copies per ml.

**Clinical Monitoring, Sample Collection, and Processing**

At least two blood samples and a complete clinical evaluation were obtained from each patient within eight weeks prior to initiation of antiretroviral therapy. Additional clinical exams were performed and blood samples were collected at 4, 12, 16, 24, 32, 44, and 48 weeks after initiation of therapy. Plasma HIV levels, complete blood counts, and T cell subsets were evaluated at each time point. T cell subsets were determined by flow cytometry analysis. Plasma was separated and stored at -80°C within two hours of collection. PBMC were obtained by ficoll-hypaque density centrifugation, cyropreserved using a liquid nitrogen step freezer, and stored in liquid nitrogen (218). All samples were processed and stored in a BSL-2 facility that was free from HIV-1 cultures and amplified or plasmid HIV-1 DNA.

**Response to Therapy**

Study endpoints were defined as any new AIDS defining illness or failure to improve CD4\(^+\) T cell counts by twenty-four weeks of therapy (156). Response to therapy
was classified by plasma virus levels and CD4$^+$ T cell numbers. Each patient was either a viral success or failure based on the extent and duration of viral suppression following initiation of therapy. Viral success involved a decline in plasma viral RNA by greater than 1.5 log$_{10}$ per ml during the first four weeks of therapy with sustained suppression below 400 copies per ml for at least sixteen weeks. Plasma viral levels below the limits of detection for the assay were calculated as 399 copies per ml. Patients whose plasma viral RNA levels declined less than 1 log$_{10}$ per ml or rebounded to detectable levels by sixteen weeks of therapy were defined as viral failures.

Immune success was defined as an increase in CD4$^+$ T cells by at least one CDC stage by twenty-four weeks of therapy. Children who were immune suppressed and showed no improvement in CD$^+$4 T cell counts were classified as immune failures. Five patients who were immune and viral failures were discontinued from the study at 24 weeks and changed to a new antiretroviral therapy. Children, who experienced increased CD4$^+$ T cell counts, but no clinical disease progression despite high viral levels, were continued on study for an additional twenty-four weeks.

**Sequence Analysis**

Plasma samples collected prior to the initiation of combination therapy were used to evaluate RT and PR genotype. Viral RNA was extracted with the QIAamp™ Viral RNA kit (QIAGEN Inc., Valencia, CA) and included controls to rule out carry over. Population sequencing of RT and PR was performed using the TruGene™ HIV-1 Genotyping Assay (Visible Genetics Inc., Toronto, Ontario, Canada) as described by the manufacturer. Briefly, amplification of cDNA by RT-PCR was followed by combination PCR/sequencing reactions (CLIPTM sequencing) using forward and reverse primers
labeled with two different dyes. Reaction products were analyzed using the OpenGeneTM automated DNA sequencing system (Visible Genetics, Inc.).

Sensitivity of population sequencing allowed detection of polymorphisms in RT or PR that occurred with a frequency greater than 20%. To verify population sequence of PR, allele-specific PR sequences were generated using reverse primer p2 (5’ CTTTTGGGCCATCCATTCCCTGGC 3’ from nucleotides 2170-2193 in the HIVLAI genome) [40] and SuperscriptTM II RNase H- Reverse Transcriptase (Gibco, Gaithersburg, MD) to synthesize cDNA. First-round amplification was performed with forward primer p1 (5’ CAGAGCCAACAGCCCCACCAG 3’, nucleotides 1724-1740) and reverse primer p2, followed by second-round amplification with nested primers p3 (5’ ACTGTATCCTTTAAACTTCCC 3’, nucleotides 1817-1836) and p4 (5’ AGTTTCAATAGGACTAATGGG 3’, nucleotides 2132-2152) to yield a 335-bp product that included the entire PR open-reading frame. Amplifications were performed in a Perkin-Elmer 9600 thermocycler using a protocol provided by J. Condra (Merck) that included denaturation at 95°C for 5 minutes, followed by 35 cycles with each cycle consisting of 94°C for 15 seconds, 56°C for 1 minute, and 72°C for 2 minutes, and a final elongation cycle at 72°C for 7 minutes. Amplified products were cloned using pGEM-TåEasy (Promega, Madison, WI) vector system and competent DH5a strain of Escherichia coli. Recombinant plasmids were purified by QIAprep™ Miniprep Kit (QIAGEN), and the presence of an insert was verified by digestion with restriction enzymes. Approximately ten (range 5 to 12) plasmids were sequenced per amplification reaction. Sequences were generated with primers p3 and p4 using the ABI PRISM™
BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer Corp., Foster City, CA) and ABI 377 automated sequence instrumentation.

Nucleic acid sequences of HIV-1 RT and PR were edited, translated, and analyzed with GeneObjects™ v3.1 (Visible Genetics, Inc.). Any detectable amino acid polymorphisms at positions in RT or PR that are associated with reduced sensitivity to drugs were scored as resistant. The GenBank accession numbers for the nucleotide sequences are AF320515-AF320559.

**Genetic Analysis**

Phylogenetic analysis was performed using neighbor-joining and maximum parsimony in the PHYLIP software package (239). Bootstrap values were based on generation of 100 replicate trees. Integrity of genetic data was verified by using amino acid alignments and construction of phylogenetic trees to compare each sequence with all pol sequences generated in our laboratory. Distances in RT and PR nucleic acid sequences within or between therapy response groups were determined by Jukes-Cantor algorithm.

drug administered are considered therapy specific, while therapy non-specific substitutions would confer decreased sensitivity to a drug not administered to the patient. Analysis of PR polymorphisms included primary positions, where initial mutations develop in drug-treated patients that have discernable effects on HIV-1 drug resistance in vitro, and secondary changes that usually arise after primary mutations in drug-treated patients and have more subtle effects on decreased sensitivity to inhibitors in in vitro assays (92,153).

**Statistical Analysis**

Clinical variables were summarized among therapy response as median and range or quartiles. Comparisons between RT or PR genotype and therapy outcome were performed using Fisher’s exact test (Sigma Stat software package, Jandel Scientific, San Rafael, CA). Statistical significance was set as a P value less than 0.05.

Univariate ordinal regression models (SAS Institute, Inc., Cary, NC) were used to examine relationship between each variable (clinical or genetic) and therapy response by combining viral and immune response into a single ordered outcome variable with three levels VFIF, VFIS, or VSIS. Individual odds ratios for viral or immune response were estimated from the regression coefficients of this model.

An exploratory and descriptive receiver operating characteristic (ROC) analysis was performed using SAS to examine if viral and/or immune outcome could be predicted as a function of clinical variables, pretherapy PR genotype, or a combination of clinical and genetic parameters (143,180). A least-squares model was employed to carry out the ROC analysis to avoid numerical fitting problems typically associated with the logistic regression approach in smaller samples (180). Area under the ROC curve (AUC) was calculated by the trapezoidal rule and used as the summary measure of fit for each model.
An AUC value of 1.0 corresponds to a perfectly fitting predictive model, while an AUC value of 0.5 corresponds to a poor predictive model (180). The best one, two, and three variable models were developed from the univariate information using a forward stepwise selection procedure. One variable models for immune or viral outcomes were determined independently using the best individual predictors of success; i.e., the variable with the highest AUC. Two-variable models were developed from one-variable models by adding the variable that provided the greatest percent increase in AUC. The procedure was repeated for development of three-variable models.

The outcome obtained from the fitted model for each patient consisted of two probability estimates (34), which corresponded to viral or immune responses following initiation of therapy. Optimal classification of actual versus predicted outcomes using viral and immune criteria was generated from P1 and P2 based on the sensitivity and specificity given by the two ROC curves.

Results

Viral and Immune Response to Therapy

Immune success (IS) in response to combination therapy was achieved by 80% (21/26) of the cohort, while viral success (VS) occurred in only 35% (9/26) of the patients. When changes in viral levels were evaluated in combination with CD4\(^+\) T cell counts, patients displayed one of three therapy responses (Figure 2-1). Nine patients (35%), who experienced sustained suppression of viral replication and significant improvement in CD4\(^+\) T cell counts, were viral and immune successes (VSIS). Five children (19%), who failed to maintain suppression of viral replication or demonstrate improvement in CD4\(^+\) T cells counts, were viral and immune failures (VFIF). Twelve children (46%), who demonstrated only transient suppression of viral replication, but
marked improvement in CD4\(^+\) T cell counts, were classified as viral failures and immune successes (VFIS). No pediatric patient displayed sustained viral suppression in the absence of increased CD4\(^+\) T cells (VSIF), which can develop in adults and most likely reflects an age-related potential for immune cell reconstitution (219).

VSIS children demonstrated a dramatic reduction in viral burden over the first four weeks of therapy with a median decline of 1.82 \(\log_{10}\) copies per ml (quartile ranges -2.84 to -1.80 \(\log_{10}\)) (Figure 2-1, upper panel). A second-phase decline in viral levels to -2.35 \(\log_{10}\) copies/ml (quartile ranges -3.02 to -1.83 \(\log_{10}\)) occurred by sixteen weeks of therapy followed by sustained suppression of plasma virus levels below 400 copies per ml. CD4\(^+\) T cell counts in the VSIS group rose by a median of 110 cells per \(\mu\)l at four weeks and continued to increase during twenty-four weeks of treatment (Figure 2-1, upper panel). Median increases of 186 CD4\(^+\) T cells per \(\mu\)l above baseline were sustained through forty-four weeks of therapy.

The VFIF group displayed an initial decline in median plasma viral levels of 1.27 \(\log_{10}\) (quartile ranges -1.86 to -0.68 \(\log_{10}\)), which was transient and rebounded to baseline levels by sixteen weeks of therapy (Figure 2-1, middle panel). CD4\(^+\) T cell numbers remained essentially unchanged during therapy.

A discordant VFIS response to therapy was characterized by a decline in plasma viral levels of 2.0 \(\log_{10}\) copies (quartile ranges -3.17 to -0.76 \(\log_{10}\)) per ml by four weeks of therapy, which rebounded to pretherapy levels by sixteen weeks (Figure 2-1, lower panel). Viral responses among the individuals in the VFIS group were virtually identical to the VFIF group of patients, but clearly distinguishable by sixteen weeks of therapy from viral success patients. Despite poor viral responses to therapy, VFIS patients
displayed sustained increases in CD4\(^+\) T cells (Figure 2-1). Median CD4\(^+\) T cell counts increased above pretherapy levels by 198 cells per µl after four weeks, 296 cells per µl following twenty-four weeks, and 772 cells per µl after forty-four weeks of combination therapy. Increased CD4\(^+\) T cell numbers displayed by the VFIS group paralleled closely immune response by the VSIS group (Figure 2-1). The magnitude of CD4\(^+\) T cell reconstitution differed between immune success and immune failure groups and was evident after only four weeks of therapy.

Long-term clinical outcomes between 24 and 48 weeks of therapy differed significantly among the therapy response groups. All VFIF patients experienced clinical decline, including new AIDS defining diseases or death. In contrast, VSIS or VFIS patients experienced no significant decline in CD4\(^+\) T cells, new AIDS defining diseases, or death.

**Pretherapy Clinical Characteristics among the Therapy Response Groups**

Clinical variables were related to therapy response by ordinal regression models (Table 2-1). Although VFIF patients (median age of 10 years) were slightly older than VSIS or VFIS groups (median age of 8 or 6 years, respectively), age and years of infection among the response groups were similar (P=0.73). Median pretherapy viral levels of 4.7 log\(_{10}\) copies per ml in the VSIS group, did not differ statistically from 5.5 or 5.2 log\(_{10}\) copies per ml among VFIS or VFIF groups, respectively (P=0.07). A majority of patients (17/26) received combination treatments that included RTV as the PI, but distribution of RTV therapy among the response groups (67% of VSIS or VFIS and 60% of VFIF patients) was similar (data not shown). Therapy response was unrelated to either
CD4⁺ T cell number (P=0.46) or CD4/CD8 ratios (P=0.16). Disease stage was the only clinical variable that displayed an association with therapy outcome (P=0.02).

**Genetic Relationship Among Viruses**

Phylogenetic analysis of viral sequences in RT or PR prior to initiation of combination therapy was assessed for each patient. Nucleotide sequences from individual patients formed distinct monophyletic branches, which confirmed the integrity of both RT and PR data sets (Figure 2-2). Branches with sequences from patients with different responses were interspersed, rather than clustered together, in the trees indicating that similar therapy response were independent of pretherapy RT or PR quasispecies.

Genetic diversity (genetic distance among viruses) in either RT or PR within the three groups of patients was similar. For example, genetic distance in RT among VS patients (mean +/- SEM, 5.6 +/- 0.6%) was indistinguishable from mean distance among VF patients (4.4 +/- 0.5% in VFIS group or 6.0 +/- 0.9% in VFIS group). Diversity in PR within VSIS patients (5.3 +/-0.7%) was similar to diversity found within VFIS (4.6 +/- 0.6%) or VFIF patients (3.9 +/- 0.7%). Pairwise comparisons of RT and PR sequences between response groups revealed similar genetic distances (data not shown), which indicated that relationships in RT or PR quasispecies among the patients at the initiation of combination therapy were independent of outcome.

**Amino Acid Polymorphisms within RT**

A potential relationship between pretherapy genotypic resistance in RT, as measured by amino acid substitutions, and viral and immune outcomes was evaluated in twenty-three patients from the cohort (Figure 2-3). All patients had received nucleoside RTIs therapy prior to initiation of combination therapy, which was evident by mutations in RT amino acids positions 41, 62, 67, 69, 70, 74, 178, 184, 210, 215, or 219 in 20
patients. Three patients, VSIS 17, VSIS 21 and VFIF 09, displayed no RT resistance mutations despite a history of RTI medications, similar to patients in other studies (153,251).

Amino acid substitutions in RT were classified as therapy specific or therapy non-specific for the RTI(s) included in the new combination treatment for each patient (204). RT genotypes were defined as sensitive, if amino acid changes were non-specific for all RTIs administered; resistant, if mutations were specific for all RTIs received; or mixed, if RT mutations were sensitive for one, but resistant to the second RTI in the treatment combination (Figure 2-3 A). RT displayed sensitive amino acid profiles in ten (44%) patients, mixed genotypes in nine (39%) patients, or resistant genotypes in four (17%) patients.

RT genotypic profiles were evaluated as variables in outcome to combination therapy (Figure 2-3 B). A majority of patients (19/23) were sensitive to at least one, and in many cases both RTIs. Among nine VSIS patients, three had sensitive RT genotypes, while four had mixtures of sensitive and resistant RT genotypes. Among ten VFIS patients, six had RT sensitive genotypes and two had mixed genotypes. RT among VFIF patients displayed either sensitive or mixed genotypes. In contrast, resistant RT genotypes were detected in only four patients, including two VSIS and two VFIS patients. No significant relationship between the genotype of RT and response to combination therapy was apparent by these analyses.

**Amino Acid Polymorphisms within Pretherapy PR**

A contribution by pretherapy PR genotypes to combination therapy outcome was assessed. Analysis focused on amino acid positions in PR that are related to reduced sensitivity to PIs (204). Amino acid substitutions in PR at positions other than those
related to decreased sensitivity to inhibitors were detected, but were unrelated to therapy outcomes (data not shown). Predominant amino acid polymorphisms among the cohort were restricted to six positions: 10, 36, 63, 71, 77, and/or 82. Even though all patients were infected prior to widespread treatment of adults by PIs, pretherapy PR genotypes, which contained at least one amino acid substitution at a position in PR that could diminish sensitivity to inhibitors, appeared in 94% (24/26) of the cohort. PR genotypes with two or more substitutions were detected in more than half the patients (58%, 15/26).

The frequency of polymorphisms differed among the amino acid positions. L63P was the most frequent substitution and appeared in 68% (17/26) of patients, followed by M36I in 36% (9/26), V77I in 32% (8/26), L10I in 20% (208), and A71T or V82I each in 8% (123) of patients. Differences in mean numbers of substitutions at the four most frequent polymorphic amino acid positions in PR (10, 36, 63, and 77) between VF and VS groups were not significant.

**Therapy-Specific Amino Acid Substitutions in Pretreatment PR Genotypes**

Amino acid substitutions in pretreatment PR were characterized as sensitive if none of the polymorphisms would reduce sensitivity to the administered PI, or resistant if polymorphisms would reduce sensitivity to the administered PI (204) (Figure 2-4 A). PI sensitive or resistant genotypes were equally distributed among the twenty-six patients in the cohort. Amino acid polymorphisms appeared at no more than one position in PR in 9/13 (69%) patients with PI sensitive genotypes, but only 2 of 13 (15%) patients with PR resistant genotypes. In contrast, PR alleles with more than one amino acid polymorphism were found in 4/13 (65) of patients with PI sensitive genotypes, but in 11/13 (85%) of patients with PI resistant genotypes (P=0.015, Fisher’s exact test).
The frequency of PI sensitive or resistant genotypes differed among the clinical response groups (Figure 2-4 B). Among patients who had PI sensitive PR genotypes, 100% achieved immune success and more than half (61% or 8/13) achieved durable viral suppression. The reliability of PI-sensitive genotype as a marker for therapy success was diminished somewhat by five VFIS patients (19% of the cohort) whose viruses failed to respond to therapy, although their PR genotypes were sensitive to the PI administered. In contrast, PI-resistant genotypes were restricted almost exclusively to VF patients. Among patients with PI-resistant amino acid substitutions 12 (3) of 13 had viral failure, whereas only patient VSIS 04 achieved a successful viral response to combination ritonavir therapy, despite a pretherapy PR genotype that included M36I. PR genotype based on population sequencing was confirmed by allele-specific sequencing (data not shown).

Overall, PR genotype as a single variable was related correctly to viral outcome in 76% (20/26) of patients. Among patients who demonstrated immune reconstitution, 62% (13/21) had sensitive genotypes, while 100% (5/5) of immune failures had resistant genotypes. Sensitive or resistant PR genotype was associated with viral or immune outcome within the response groups (P = .01 or 0.04 respectively, Fisher’s exact test).

Viral and immune responses were combined into a single ordered outcome variable with three levels (VFIF, VFIS, and VSIS) for analysis by univariate ordinal regression models. PR-sensitive or -resistant genotype, as a variable, was related significantly to viral and immune response to therapy (P=0.005). Odds ratios for viral or immune response were estimated from the regression coefficients of the model. The odds for viral failure relative to success for PR resistant genotypes were 2.5 to 1. In contrast,
the odds for immune failure compared to immune success for PR resistant genotypes were 17.7 to 1, reflecting that all immune failures occurred in patients with PR resistant genotypes.

**Clinical and Genetic Variables That Predict Viral and Immune Outcome**

To develop a more sensitive and specific model to predict therapy response, pretherapy clinical and PR genetic variables were evaluated by multivariate analysis. The best single pretherapy clinical variable for predicting immune outcome by ROC analysis was CD4% (AUC=0.83), while disease stage (AUC=0.75) provided the best single clinical predictor of viral response. Other clinical variables, such as age, sex, or viral burden, had AUC values between 0.54 and 0.69 with limited predictive value. PR genotype as a single variable was similar for immune or viral outcome (AUC=0.81 or 0.80, respectively). When the two best variables, either CD4% plus PR genotype or disease stage plus PR genotype, were combined, the AUC increased to 0.91 for immune outcome, or 0.88 for viral outcome. The best predictive models for therapy outcome with AUC close to the maximum of 1.0 included three variables, two clinical and one genetic. PR genotype plus CD4% and age resulted in an AUC of 0.97 (confidence interval 0.89-0.99) for immune response. PR genotype combined with disease stage and CD4% resulted in an AUC of 0.92 (confidence interval 0.82-0.99) for viral outcome. CD4+ T cell number, plasma viral levels, or sex were also evaluated in three-variable models, but failed to improve the predictive value for either immune or viral outcomes (data not shown).

Immune and viral models were combined to determine the ability to predict outcome based on VFIF, VFIS, or VSIS responses. When the best clinical variables (CD4% plus disease stage) were used, therapy response was predicted correctly for 58%
of the patients (15/26) (data not shown). In contrast, when PR genotype was combined with the two best clinical variables for immune or viral outcome, response was predicted correctly for 81% (21/26) of patients (Table 2-2). Among five patients who were misclassified, two VFIS patients were predicted to have VSIS responses, while three (two VSIS and one VFIS) patients displayed responses that were better than predicted by the model. Overall, the combined model that included PR genotype plus clinical variables predicted correct therapy response in 100% (5/5) VFIF, 78% (21/27) VSIS, and 75% (9/12) VFIS patients.

Discussion

Genetic characteristics of PR or RT, which are targets in the viral genome for drug action, provide prognostic value when combination antiretroviral therapy is switched for patients who fail to maintain viral suppression (92,185,256). In our study HIV-1 PR genotype provided accurate prognostic value for patients who were naïve to PI-combination therapy. The findings indicate that a spectrum of naturally occurring amino acid polymorphisms in PR can be used as biomarkers to predict both immune and viral outcomes to initial combination therapy in HIV-infected patients.

Disease stage, viral load, or CD4+ T cell counts, which are generally applied prognostic measures of treatment outcome (24,142,173) predicted response to combination antiretroviral therapy among 58% of patients in the cohort, which was only slightly better than chance. Limited value of clinical variables was not unexpected in view of inclusion criteria that focused on patients with advanced disease. Pretherapy RT genotype provided little prognostic value for predicting therapy response in the cohort, most likely because patients received extensive RT therapy over a period of years (251) (153). Nonetheless, patients who had RT sensitive genotypes were no more likely than
patients who were resistant to new NRTIs to achieve a successful outcome to combination therapy. In contrast to RT genotype, pretherapy PR genotype, either alone or in combination with clinical variables, led to development of a reliable and sensitive model, which predicted both viral and immune outcome in more than 80% of the cohort of PI-naïve patients.

PR amino acid profiles that included PI resistant polymorphisms were detected at some frequency in virtually all patients in our cohort, similar to pretherapy PR genes in other data sets (9,86,125,209,209,226,228). These polymorphisms are localized predominantly in amino acid positions classified as secondary mutations, persist for years prior to initiation of PI therapy, and are independent of primary mutations that appear in response to drug therapy (9,78). Although a role for secondary mutations alone in diminishing PR sensitivity to inhibitors can be difficult to demonstrate in biochemical or replication assays, secondary mutations in combination with mutations in the active site clearly enhance PR resistance to drugs (39,152). Even if amino acid polymorphisms in secondary sites in PR fail to reduce detectable drug sensitivity, secondary substitutions present prior to the initiation of therapy may provide a replication advantage when the virus is placed under the selective pressure of drug therapy in patients (21). Independent of function, natural polymorphisms in amino acid positions 10, 36, 63, and/or 77 in PR served as accurate biomarkers for therapy outcome.

Several factors were essential for the success of the overall model to predict response. First, stratification of therapy outcome based on both viral and immune parameters was critical. If viral response to treatment had been used to the exclusion of immune response in classifying outcome, the relationship between pretherapy protease
genotype and treatment response would have been less obvious and predictive value of the model would have been diminished. Second, natural polymorphisms in any amino acid position in PR that could be associated with diminished sensitivity to inhibitors were considered. Natural polymorphisms occurred predominantly in any of four positions in PR that were localized outside the active site of the enzyme. If analysis had focused only on polymorphisms in primary sites to the exclusion of secondary sites in PR, pretherapy PR genotype would have limited, if any, predictive value. Finally, defining PR genotype as sensitive or resistant relative to the particular PI received provided a parameter that enhanced both the univariate regression model and the multivariate model.

A striking finding of the study was the value of the multivariate model to predict therapy outcome not only for children who were viral and immune failures or successes, but for children who developed a discordance response as well. Even though the repertoire of viral and host factors that can lead to immune success with viral failure is undefined, use of viral genotype combined with pretherapy clinical parameters resulted in successful prediction of outcome in 75% of VFIS patients (127). Pediatric patients were particularly well-suited for the analysis because of high pretherapy viral burden and greater potential for immune reconstitution in comparison to adults (69,219). Nevertheless, restoration of CD4$^+$ T cell counts despite rebound of viral burden develops in both HIV-infected children and adults treated with combination antiretroviral therapy, and occurs in as many as 50% of patients in some cohorts (49,69,127,219).

While the multivariate model provided a significant level of accuracy in prediction of outcome for a majority of patients, approximately 20% of the patients in the cohort were misclassified, indicating that variables in addition to disease stage, immune
status, and PR genotype impact therapy response. Two patients had PR sensitive
genotypes and clinical parameters that predicted a VSIS response, in contrast to their
actual VFIS therapy response. Genetic determinants outside PR, for example in Gag or in
PR cleavage sites, that modulate PR activity could account for these results (18,62,254).
Alternatively, viral failure in these patients might result from suboptimal drug levels or
immune failure to control viral replication. Three patients experienced viral and/or
immune outcomes that were better than predicted by the model. Host factors, such as
cytotoxic T-cell responses or polymorphisms in chemokines or their receptors, could
impact control of viral replication and contribute to the unexpected success to therapy by
these patients.

Opportunities to control viral replication with combination drug therapies are
limited. Initial intervention prior to the development of cross-resistance and further
immune decline is preferable to salvage therapy, but requires identification of biomarkers
in addition to standard clinical variable that can contribute to prediction of response to
combination antiretroviral therapy. Currently little information supports generalized used
of pretherapy genotyping in clinical practice (92). Our data clearly indicate enhanced
prognostic significance for naturally occurring amino acid polymorphisms in PR
combined with clinical variables among PI-naïve patients with advanced disease and
provide a foundation for design of clinical studies that use pretherapy PR genotype to
optimize combinations of antiretroviral therapeutics.
Figure 2-1. Viral and immune responses during 44 weeks of combination for human immunodeficiency virus (HIV) type 1 infection. Median log10 HIV RNA copies per milliliter (■) are indicated on left axis. Median CD4 T cell counts per microliter (▲) are indicated on the right axis. Vertical bars, 25th-75th quartile range. Top, Viral success and immune success (VSIS; n = 9); center, viral failure and immune failure (VFIF; n = 5); bottom, viral failure and immune success (VFIS; n=12).
### Table 2-1. Clinical characteristics of study population at entry

<table>
<thead>
<tr>
<th>Therapy Response Groups</th>
<th>No. of subjects</th>
<th>Median (range) age, years</th>
<th>Disease Stage A</th>
<th>B</th>
<th>C</th>
<th>CD4 T Cells/μl&lt;sup&gt;a&lt;/sup&gt;</th>
<th>CD4/CD8 Ratio&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Log&lt;sub&gt;10&lt;/sub&gt; Plasma Viral RNA Copies/ml&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSIS</td>
<td>9</td>
<td>8 (6-13)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>256 (78-273)</td>
<td>0.24 (.12-.41)</td>
<td>4.7 (4.3-5.4)</td>
</tr>
<tr>
<td>VFIF</td>
<td>5</td>
<td>10 (9-14)</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>25 (3-45)</td>
<td>0.04 (.03-.12)</td>
<td>5.5 (5.3-5.6)</td>
</tr>
<tr>
<td>VFIS</td>
<td>12</td>
<td>6 (4-9)</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>210 (41-410)</td>
<td>0.28 (.07-.48)</td>
<td>5.2 (4.9-5.7)</td>
</tr>
</tbody>
</table>

P value<sup>b</sup> 0.73 0.02 0.46 0.16 0.07

**NOTE:** VFIF, viral failure and immune failure; VFIS, viral failure and immune success; VSIS, viral success and immune success

<sup>a</sup>Median value (25<sup>th</sup>-75<sup>th</sup> interquartile range).

<sup>b</sup>P values were based on ordinal regression models.
Figure 2-2. Phylogenetic relationships of reverse transcriptase (A) and protease (B) sequences among the cohort of subjects with human immunodeficiency virus (HIV) type 1 infection. Phylogenetic relationships were assessed by using a neighbor-joining algorithm and nucleic acid sequences from pretherapy plasma samples. HIV_{LAI} (*) was used as the outlier sequence. Sequences for each subject are designated by therapy outcome. ▲, Viral success and immune success; ■, viral failure and immune failure; ○, viral failure and immune success. "0.1" Indicates genetic distance.
### Figure 2-3.
Reverse transcriptase (RT) genotypes before initiation of combination therapy for HIV-1 infection.

**A.** Amino acid mutations in RT. Patients are designated by numbers. All patients had received RT inhibitor (RTI) therapy (3TC, lamivudine; D4T, stavudine; DDI, didanosine; EFV, efavirenz; NVP, nevirapine; ZDV, zidovudine). New RTI refers to RTIs included in protease inhibitor combination therapy. Amino acid positions in RT were designated by numbers and were related to resistance to new RTIs (Schinazi, 2000). *White box,* no change from RT genotype of HIV-1_{LAI}; *shaded box,* change nonspecific for new RTI(s); *black box,* change specific for new RTI(s). RT genotype was sensitive (S), resistant (R), or sensitive to one but resistant to a second RTI (S/R). No patient displayed multidrug-resistant (MDR) RT genotypes involving insertions between codons 67 and 7 and A62V, a single MDR mutation, was detected only in viral success and immune success (VSIS) patient 03.

**B.** RT genotypes relative to therapy outcome. *White bars,* VSIS response, *striped bars,* VFIS response; *black bars,* VFIF response.

#### Table 2-3

<table>
<thead>
<tr>
<th>Patient</th>
<th>RTI</th>
<th>Resistant Amino Acid Profiles</th>
<th>RT Genotype</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>D4T/EFV</td>
<td></td>
<td>S</td>
<td>VSIS</td>
</tr>
<tr>
<td>20</td>
<td>3TC/D4T</td>
<td></td>
<td>S</td>
<td>VSIS</td>
</tr>
<tr>
<td>21</td>
<td>D4T/NVP</td>
<td></td>
<td>S</td>
<td>VSIS</td>
</tr>
<tr>
<td>09</td>
<td>D4T</td>
<td></td>
<td>S</td>
<td>VFIF</td>
</tr>
<tr>
<td>13</td>
<td>3TC/D4T</td>
<td></td>
<td>S</td>
<td>VFIS</td>
</tr>
<tr>
<td>15</td>
<td>D4T</td>
<td></td>
<td>S</td>
<td>VFIS</td>
</tr>
<tr>
<td>18</td>
<td>D4T</td>
<td></td>
<td>S</td>
<td>VFIS</td>
</tr>
<tr>
<td>19</td>
<td>D4T</td>
<td></td>
<td>S</td>
<td>VFIS</td>
</tr>
<tr>
<td>22</td>
<td>D4T</td>
<td></td>
<td>S</td>
<td>VFIS</td>
</tr>
<tr>
<td>23</td>
<td>D4T/NVP</td>
<td></td>
<td>S</td>
<td>VFIS</td>
</tr>
<tr>
<td>01</td>
<td>ZDV/3TC</td>
<td>S/R</td>
<td>S/R</td>
<td>VSIS</td>
</tr>
<tr>
<td>04</td>
<td>ZDV/3TC</td>
<td>S/R</td>
<td>S/R</td>
<td>VSIS</td>
</tr>
<tr>
<td>16</td>
<td>ZDV/3TC</td>
<td>S/R</td>
<td>S/R</td>
<td>VSIS</td>
</tr>
<tr>
<td>25</td>
<td>3TC/D4T</td>
<td>S/R</td>
<td>S/R</td>
<td>VSIS</td>
</tr>
<tr>
<td>05</td>
<td>DDI/D4T</td>
<td>S/R</td>
<td>S/R</td>
<td>VFIF</td>
</tr>
<tr>
<td>07</td>
<td>ZDV/3TC</td>
<td>S/R</td>
<td>S/R</td>
<td>VFIS</td>
</tr>
<tr>
<td>08</td>
<td>D4T/3TC</td>
<td>S/R</td>
<td>S/R</td>
<td>VFIS</td>
</tr>
<tr>
<td>14</td>
<td>ZDV/3TC</td>
<td>S/R</td>
<td>S/R</td>
<td>VFIS</td>
</tr>
<tr>
<td>26</td>
<td>3TC/D4T</td>
<td>S/R</td>
<td>S/R</td>
<td>VFIS</td>
</tr>
<tr>
<td>02</td>
<td>ZDV/DDI</td>
<td>R</td>
<td>VSIS</td>
<td></td>
</tr>
<tr>
<td>03</td>
<td>D4T</td>
<td>R</td>
<td>VSIS</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>ZDV/DDI</td>
<td>R</td>
<td>VSIS</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>ZDV/DDI</td>
<td>R</td>
<td>VSIS</td>
<td></td>
</tr>
</tbody>
</table>

---

*Figure 2-3.* Reverse transcriptase (RT) genotypes before initiation of combination therapy for HIV-1 infection. 

A, Amino acid mutations in RT. Patients are designated by numbers. All patients had received RT inhibitor (RTI) therapy (3TC, lamivudine; D4T, stavudine; DDI, didanosine; EFV, efavirenz; NVP, nevirapine; ZDV, zidovudine). New RTI refers to RTIs included in protease inhibitor combination therapy. Amino acid positions in RT were designated by numbers and were related to resistance to new RTIs (Schinazi, 2000). *White box,* no change from RT genotype of HIV-1_{LAI}; *shaded box,* change nonspecific for new RTI(s); *black box,* change specific for new RTI(s). RT genotype was sensitive (S), resistant (R), or sensitive to one but resistant to a second RTI (S/R). No patient displayed multidrug-resistant (MDR) RT genotypes involving insertions between codons 67 and 7 and A62V, a single MDR mutation, was detected only in viral success and immune success (VSIS) patient 03. 

B, RT genotypes relative to therapy outcome. *White bars,* VSIS response, *striped bars,* VFIS response; *black bars,* VFIF response.
### A. Sensitive

<table>
<thead>
<tr>
<th>Patient</th>
<th>PI</th>
<th>10</th>
<th>36</th>
<th>63</th>
<th>71</th>
<th>77</th>
<th>82</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>02</td>
<td>03</td>
<td>10</td>
<td>36</td>
<td>63</td>
<td>71</td>
<td>77</td>
<td>82</td>
</tr>
<tr>
<td>16</td>
<td>07</td>
<td>08</td>
<td>17</td>
<td>20</td>
<td>21</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td>17</td>
<td>20</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>15</td>
<td>16</td>
<td>25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### B. Resistant

<table>
<thead>
<tr>
<th>Patient</th>
<th>PI</th>
<th>10</th>
<th>36</th>
<th>63</th>
<th>71</th>
<th>77</th>
<th>82</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>04</td>
<td>05</td>
<td>06</td>
<td>07</td>
<td>08</td>
<td>09</td>
<td>13</td>
<td>14</td>
<td>VSIS</td>
</tr>
<tr>
<td>05</td>
<td>06</td>
<td>07</td>
<td>08</td>
<td>09</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>VFIS</td>
</tr>
<tr>
<td>05</td>
<td>06</td>
<td>07</td>
<td>08</td>
<td>09</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>VFIS</td>
</tr>
</tbody>
</table>

Figure 2-4. Genotypes of protease (PR) before PR inhibitor (PI) combination therapy for HIV-1 infection. A, PR genotype. Patients are designated by the same numbers as in Figure 2-3. IDV, indinavir; NFV, nelfinavir; RTV, ritonavir. White box, no resistance substitution; shaded box, therapy-nonspecific amino acid substitution; green box, therapy-specific amino acid substitution (Schinazi, 2000). Patients had viral success and immune success (VSIS), viral failure and immune failure (VFIF), or a mixed response (viral failure and immune success (VFIS)). B, Pretherapy PR genotypes relative to therapy outcome. Aqua bars, VSIS response; purplestriped bars, VFIS response; black bars, VFIF response.
Figure 2-4--continued.

B.

<table>
<thead>
<tr>
<th></th>
<th>VSIS</th>
<th>VFIS</th>
<th>VFIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resistant</td>
<td>12</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Sensitive</td>
<td>8</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2-4--continued.
Table 2-2. A combination of clinical and genetic variables predict therapy outcome in subjects with HIV-1 Infection.

<table>
<thead>
<tr>
<th>Actual Response&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Predicted Response&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSIS</td>
<td>7 1 1</td>
</tr>
<tr>
<td>VFIS</td>
<td>2 9 1</td>
</tr>
<tr>
<td>VFIF</td>
<td>0 0 5</td>
</tr>
</tbody>
</table>

NOTE. VFIF, viral failure and immune failure; VFIS, viral failure and immune success; VSIS, viral success and immune success.

<sup>a</sup> Actual response: VSIS, n = 9; VFIS, n = 12; VFIF, n = 5.

<sup>b</sup> The 2 best clinical variables (CD4 percentage and disease stage) were combined with protease genotype as a third variable. Predicted response was calculated using receiver operating characteristic analysis. Boldface numbers represent patients whose predicted response to therapy was identical to their actual response.
CHAPTER 3
GENETIC ANALYSIS OF GAG-PRO FOLLOWING 24 WEEKS OF COMBINATION THERAPY

Introduction

Virus mutants resistant to protease inhibitors used in the clinical setting involve an accumulation of multiple amino acid substitutions (39,41,82,130,152,202,204). Although the accumulation of mutations in protease have been associated with failure to combination therapy, there have been reports of viral rebound in adults who have no protease resistance substitutions (251), suggesting that mutations outside of protease can impact response to therapy. Mutations in the D' and D cleavage sites provide some compensation for resistance associated loss of viral fitness (62) (134) and recent studies have linked these mutations with therapy specific mutations in protease in the adult population (6,43,116).

The paradoxical response of increased CD4⁺ T cells in combination with high viral loads (VFIS) raises a number of questions about the virus and the host immune response. Viruses that rebound during therapy could reflect emergence from a genetic bottleneck by a minor population of viruses or progressive accumulation of resistance among a spectrum of viruses. Analysis by phylogenetic relationships and genetic variability provides quantitative measures of all nucleic acid changes in viral populations and a method by which to access relationships in virus populations before and after 24 weeks of therapy among response groups. A major focus of the Goodenow laboratory is to identify viral parameters that could contribute to the discordant response. Tuttle et al.
demonstrated that changes in envelope do not appear to be responsible for the apparent reduction in pathogenesis of virus rebounding in VFIS patients (232).

At the initiation of this study there was limited knowledge about post therapy gag and pro genotype relative to the therapy response groups identified in Chapter 2 of this dissertation. Genetic characteristics of virus after 24 weeks of therapy were determined by sequencing multiple protease alleles from plasma and cell associated viruses for viral failure patients or from cells alone for viral success patients whose plasma virus levels were below 400 copies per ml. Phlyogenetic analysis of protease from treated patients will provide information about the genetic basis for the emergence of resistance and the pathogenesis of HIV while under the selective pressure of protease inhibitors.

Materials and Methods

Study Patients

The study cohort was comprised of a subset of HIV-infected children and adolescents who were evaluated in Chapter 2 of this dissertation. Of the twenty-six patients analyzed for pretherapy polymorphisms in Chapter 2, nineteen patients fulfilled the inclusion criteria and also demonstrated compliance with combination therapy. Eligible patients were between one and eighteen years of age and received at least 24 weeks of PI therapy. Fourteen patients (74%) received combination therapy with ritonavir (RTV), one (5%) with nelfinavir (NFV), and four (21%) with indinavir (IDV). Selection of PI was based on availability of pediatric formulations, as well as ability of the patient to swallow capsules or tolerate liquid formulations and to adhere to the treatment regimen. Optimal drug dosing was based on pharmacokinetic studies of PIs in children (122, 154, 155).
The cohort had a median age of seven years (range two to sixteen years) and included six adolescents. Nine patients were females and ten were males. Nine children (47.5%) were African Americans, nine (47.5%) were Caucasian, and one (5%) was Asian American. All patients were infected prior to introduction of PI therapy into the adult population. Post therapy median CD4 T cell counts were 524 cells per ul (25th to 75th quartile ranges from 214 to 738 cells/ul), median CD4/CD8 ratio was 0.45, and median HIV plasma RNA levels were $5.1 \log_{10}$ (25th to 75th quartile ranges from 3.7 to 5.5 log10) copies per ml.

**Clinical Monitoring, Sample Collection, and Processing**

Blood samples and a complete clinical evaluation were obtained from each patient at approximately 24 weeks following initiation of antiretroviral therapy. Plasma HIV levels, complete blood counts, and T cell subsets were evaluated. Processing and storage of patient samples and the definition of each clinical response to therapy can be found in the Materials and Methods in Chapter 2 of this dissertation.

**Sequence Analysis**

Plasma and PBMC samples collected following 24 weeks of combination therapy were used to evaluate PR and Gag genotype. Details of viral RNA extraction, amplification, and analysis are included in Chapter 2 of this dissertation. An amino acid alignment of protease alleles analyzed can be found in Appendix A of this dissertation. The Gag alignment may be found in Appendix B of this dissertation. The GenBank accession numbers are pending.

**Amplification of protease from PBMC**

Cell associated viral DNA was amplified from the C terminus of capsid into reverse transcriptase using two rounds of PCR with nested primers as previously
described in Barrie et al. (9). First round amplification was performed using forward primer gag7 (5’-GTTAAAGAGACCATCAAT-3’, nucleotides 935 to 953 in HIV_LAI) and reverse primer pol4 (5’-TCCTACATACAAATCATC-3’, nucleotides 2710-2684) to produce a 1.7 kb fragment. Second round amplification with forward primer g100 (5’-TAGAAGAAATGATGACAG-3’, nucleotides 1363-1380) and reverse primer pol1 (5’-ACTTTTGGGCCATCCATTGCTGC-3’, nucleotides 2170-2193) yielded an 830 bp fragment. Amplification reactions were carried out in a Perkin Elmer 9600 tempcycler. Amplification conditions included an initial denaturation at 95°C for 10 minutes, followed by 35 cycles each consisting of denaturation at 95°C for 1 minute, primer annealing at 55°C for 30 seconds, and extension at 72°C for 3 minutes, and a final elongation at 72°C for 10 minutes. Amplified products were cloned using the pGEM-T Easy® vector system (Promega) and competent DH5α strain of Escherichia coli. Recombinant plasmids were purified by use of a QIAprep Miniprep kit (QIAGEN), and the presence of an insert was verified by digestion with restriction enzymes. Approximately 10 (range, 4–19) plasmids were sequenced per amplification reaction. Sequences were generated with primers p2, p3, p4, or g100 using ABI PRISM BigDye Terminator Cycle Sequencing and ABI 377 automated sequence instrumentation (Perkin-Elmer). Sequences were proofed using SeqEd v1.0.3(ABI, Inc.), then aligned using COMPARE in the DNA sequence Alignment Editor 2.4, (Dr. Alan Goldin, Department of Biology, Cal Tech, 1994) and translated with Gene Runner 3.0 for Windows (Hastings Software, Inc., 1994).
Genetic Analysis

Phylogenetic analysis was performed on both the nucleotide and amino acid sequences of HIV-1 protease obtained prior to and following combination therapy. Protease nucleotide and amino acid sequences were edited and aligned using the Clustal analysis program version at the European Bioinformatics Institute website (http://www2.ebi.ac.uk/clustalw/) (23). Parsimony trees were constructed using PHYLIP (239). Bootstrap analysis for nucleotide sequences was carried out with DNAPARS and CONSENSE and for amino acid trees was performed with SEQBOOT. Bootstrap values were based on generation of 100 replicate trees. Trees were drawn using the phylogram output option and the programs TREEVIEW and PowerPoint were used to insert symbols and reduce the scale of the trees. Distance analysis was calculated using the programs SYNO and SEND. Statistical Analysis was performed with the Sigma Stat software package (Jandel Scientific, San Rafael, CA). The two-way ANOVA test for analysis of variance was used to evaluate and compare means for evolutionary parameters.

Results

Viral and Immune Response to Therapy

Immune success (IS) in response to combination therapy was achieved by 79% (15/19) of the cohort, while viral success (VS) occurred in only 26% (129) of the patients. Five patients (26%) who experience sustained suppression of viral replication and significant improvement in CD4 T cell count, were viral and immune success (VSIS). Four children (21%), who failed to maintain suppression of viral replication or demonstrate improvement in CD4 T cell counts, were viral and immune failures (VFIF). Ten children (53%), who demonstrated only transient suppression of viral replication, but
marked improvement in CD4$^+$ T cell counts, were classified as viral failures and immune successes (VFIS).

**Evaluation of Clinical Variables**

Pretherapy clinical variables of this cohort are similar to the larger cohort described in Chapter 2. Clinical variables at 24 weeks of therapy were statistically different between response groups (Table 3-1). Median post therapy virus levels of $\leq 2.6 \log_{10}$ copies/mL in the VSIS group differed statistically from 5.2 and 5.5 $\log_{10}$ copies/ml among the VFIS ($p<0.001$) and VFIF ($p<0.001$) groups, respectively. Immune success, regardless of viral response was statistically different from the VFIF. CD4$^+$ T cell count of 39 cell/$\mu$L in the VFIF was statistically different from the 524 and 930 cell/$\mu$L among the VSIS ($p=0.008$) and VFIS ($p=0.003$), respectively.

**Amino Acid Polymorphisms in Protease [PR]**

The accumulation of polymorphisms in protease was evaluated following 24 weeks of combination therapy. Analysis focused on amino acid positions in PR that are related to reduced sensitivity to protease inhibitors [PI] (204). Protease alleles are summarized as a consensus sequence for each patient in Figure 3-1. Mutations that are specific to the administered protease inhibitor were scored as solid black boxes if the mutation accumulated during the course of therapy or as striped black boxes if the polymorphism existed at that position prior to the initiation of therapy. Non-specific mutations, which could reduce sensitivity to other PIs, were scored as solid purple boxes if they accumulated during therapy or as hatched purple boxes if the polymorphism existed prior to therapy. Most new mutations were specific for the protease inhibitor
received (black boxes), although occasional mutations that were not directly related to the PI administered accumulated (purple boxes)(Figure 3-1).

The number of therapy specific and non-specific mutations was evaluated independently and cumulatively (i.e. any drug associated polymorphism present in the 24 week allele). Comparison of protease alleles from patients displaying viral success and viral failure demonstrated significant differences with respect to total specific mutations (p=0.003) and cumulative mutations (p=0.01) (Table 3-2). When protease alleles were analyzed with respect to the patient’s pretherapy genotype (protease sensitive or resistant) there was no significant difference. Discordant response patients (VFIS) were divided between protease sensitive and resistant genotypes prior to therapy, yet following therapy neither the quantity or quality of protease mutations were sufficient to distinguish between these patients.

Analysis by response group demonstrated the median number of new, drug related mutations in protease was 3 [range 1 to 7] for VFIS patients and either 0.6 [range 0 to 2] for VSIS or 1.2 [range 0-3] among VFIF patients (Figure 3-2). The VSIS group differed statistically from both the VFIF (p=0.027) and VFIS (p=0.003) with respect to total therapy specific mutations (Table 3-2). No significant differences existed between the VFIS and VFIF groups in terms of protease polymorphisms associated with reduced sensitivity to protease inhibitors.

**Supplemental Polymorphisms in Protease**

A pattern of amino acid substitutions in protease positions that have not been associated with reduced sensitivity to protease inhibitors was identified. The following polymorphisms were detected: I13V, I15V, V32I, E34Q, R42K, E35D, R57K, I64P, K70R, L88S, and I93L. Although a number of these polymorphisms have been recorded
in patients treated with protease inhibitors, there has been no association of these polymorphisms with therapy response and no phenotypic evidence of reduced drug sensitivity (39,41).

The identified polymorphisms referred to as “supplemental” changes appeared with the highest frequency among the viral failure patients (Figure 3-3 A). Supplemental polymorphisms were scored as striped blue boxes if present prior to therapy initiation or as solid blue boxes if they accumulated during the course of therapy. Prior to therapy each response group differed significantly from the other groups with respect to the presence of supplemental changes (blue striped boxes, Figure 3-3 A).

Following therapy, the appearance of new supplemental polymorphisms was associated with clinical response. VSIS patients (0/5) did not demonstrate any supplemental polymorphisms while, 5/5 VFIF and 9/10 VFIS patients displayed at least one supplemental polymorphism (Figure 3-3 A). VFIS patients displayed a median of 2 new supplemental changes compared to the VSIS and VFIF groups, which displayed a median of zero new supplemental polymorphisms. The total number of supplemental polymorphisms present after 24 weeks of therapy distinguished the VSIS group from both the VFIS (p=0.001) and VFIF (p=0.02) groups (Figure 3-3 B).

**Distance Analysis**

The extent of variation within protease was evaluated by calculating genetic distance (D). Median diversity (intra-sample genetic distance) was 2.4% (range 1.6 to 4.3), 1.7% (range 0.7-4.8) and 1.6% (range 1.0-2.2) prior to therapy and 3.1% (range 0.5-4), 2.92% (range 0.6-4.3) and 1.9% (range 0.2-2) following 24 weeks of therapy for VSIS, VFIS, and VFIF, respectively (Figure 3-4). There was no statistically significant
difference in pre- or post therapy diversity between response groups, or between pre and post therapy diversity within a response group.

To examine the extent of variation within protease following 24 weeks of protease inhibitor therapy total distance (or divergence) was evaluated. Overall each response group demonstrated an increase in the genetic distance (Figure 3-5). Median divergence following therapy was 3.1% (range 1.4-5.3%), 3.5% (range 2.1-4.8%) and 1.9% (range 1.8-4.0%) for VSIS, VFIS, and VFIF respectively (Figure 3-9). Total divergence was not statistically different between groups.

Overall distance \([D]\) is comprised of nucleotide changes that cause amino acid substitutions (nonsynonymous \([D_N]\)) and nucleotide changes that do not (synonymous \([D_S]\) substitutions). Synonymous distance increased in all response groups (Figure 3-6). The median increase for the VSIS group was 2% demonstrating that despite successful treatment with antiretrovirals there is continued viral replication. The median synonymous distance increased by approximately 1% in both the VFIS and VFIF groups (Figure 3-9). There is no statistically significant difference in the synonymous distance between response groups. Although a slight increase in nonsynonymous divergence was detected in each response group the increase was more pronounced in VFIS patients (Figure 3-7). The median increase in nonsynonymous divergence for the VSIS and VFIF groups was less than 1% following 24 weeks of therapy (Figure 3-8). In contrast, the VFIS group displayed a median 2% increase in nonsynonymous distance, which differed statistically from the VFIF (p= 0.04) group.

The contribution of \([D_S]\) and \([D_N]\) to the overall divergence was evaluated (Figure 3-7). The median ratio of nonsynonymous distance to synonymous distance \((D_N/D_S)\) was
0.19 (range 0.15 -0.37) for VSIS, 0.72 (range .38-1.33) for VFIS and 0.21 (range 0.04-.33) for VFIF (Figure 3-9). The relative contribution of nonsynonymous mutations to the overall genetic divergence was greater in VFIS patients than the other response groups. The ratio of $D_N/D_S$ for VFIS response was statistically different than both VSIS (p=0.015) and VFIF (p=0.022) response groups.

**Genetic Relationship among Viruses**

Phylogenetic analysis of viral sequences in protease was assessed for patients following 24 weeks of combination therapy. Nucleotide sequences from individual patients formed distinct monophyletic branches, which confirmed the integrity of the data set (Figure 3-10 A, C and E).

The close relationship of pre and post therapy sequences in both the nucleotide and amino acid trees within a patient is consistent with a low level of viral replication and the accumulation of few mutations in the VSIS group (Figure 3-10 A and B). Trees from VFIF patients also demonstrated little evolution from the pre and post therapy alleles. In contrast, at the nucleotide level VFIS patients post therapy alleles are further from the node of the tree in a majority of patients demonstrating that post therapy alleles arise from pretherapy alleles (Figure 3-10 C). Yet at the amino acid level post therapy alleles from several VFIS patients appear closer to the node of the tree than pretherapy alleles (Figure 3-10 C). One possible explanation is convergent evolution of protease between patients while under the selective pressure of therapy. VFIS patients tended to display increases in the nonsynonymous distance, which would relate to the accumulation of mutations associated with decreased sensitivity to protease inhibitors. Since there is a restriction on the location and type of mutations that accumulate in protease in response
to drug, it is plausible that convergent evolution would be seen upon evaluation at the amino acid level.

**Analysis of Gag**

A region of gag from the B cleavage site through p6 was evaluated for polymorphisms in fourteen of the nineteen patients analyzed for protease polymorphisms. This cohort consisted of three VSIS, eight VFIS and two VFIF patients. The amino acid alignment of all clones analyzed is located in Appendix B of this dissertation.

The association of Gag polymorphisms with reduced sensitivity to protease inhibitors has not been fully elucidated therefore in this analysis any amino acid positions that varied from the reference sequence, HIV\textsubscript{LAI}, were scored as polymorphisms. The dominant Gag cleavage site (cleavage site that appeared most frequently) was used in this analysis. The B cleavage site demonstrated low variability prior to and after 24 weeks of protease inhibitor therapy, with only 1/14 patients displaying polymorphisms detected at both timepoints. The D’ site also displayed low heterogeneity with two patients displaying identical polymorphisms at both pre and post therapy timepoints. Following therapy one patient (VFIS12) acquired an aspartic acid in place of glutamic acid at position 428, while two patients (VFIS22 and VFIF09) accumulated a mutation at position 431 (A\(\rightarrow\)V), which has been associated with reduced drug sensitivity (6,43,108,116,134). The C and D cleavage sites demonstrated more variability and were rigorously evaluated to determine the relationship of polymorphisms to therapy response.

The D cleavage site was variable in 10/14 patients. Seven patients had polymorphisms present at the initiation of therapy that persisted. Four patients had polymorphisms in the D cleavage site that reverted to wildtype (LAI like) following therapy. VFIS15 accumulated an arginine in place of lysine at position 452. Mutations of
P453L and L449F, which have been associated with exposure to protease inhibitors (108,116,134), were not detected in this cohort or by Cote et al. (43).

The C cleavage site was variable in all fourteen patients (Table 3-3). Nine patients had polymorphisms prior to therapy that persisted, while the remaining five patients accumulated new mutations within the C site during therapy. The dominant pre and post therapy allele and the frequency are presented in Table 3-3. Polymorphisms are predominantly located in positions P4, P3, P2, and P3’ in the C-sites from all response groups and p5 in VF patients. When patients were stratified by viral success or failure there was no difference with respect to the location or number of positions with polymorphisms in the C site.

The frequency of changes at each position differed between response groups, with fewer amino acid substitutions in the VSIS than VFIF or VFIS patients. All VFIF patients displayed a polymorphism at position p4 of the C cleavage site prior to and following therapy. This same polymorphism was detected in 56% of VFIS patients prior to therapy and 78% following therapy, while the p4 polymorphism was seen in 33% of VSIS patients both prior to and following therapy (Table 3-4). At least half of the VFIS and VFIF patients had a substitution of R3’K while only 33% of VSIS patients had such a polymorphism. No direct relationship between C cleavage site polymorphisms and therapy response was detected.

Gag was analyzed for polymorphisms that could be directly related to therapy response. Any polymorphism that was detected in greater than 50% of each response group was considered to be part of the high genetic background and not considered in this analysis. A number of amino acid positions were found to more polymorphic in the
failure and discordant groups than the success group. Greater than 50% of the VFIS patients displayed polymorphisms in 3 positions of the C site, p4, p2, and p3’, as well as amino acid positions 389 and 403 in p7. The majority of VFIF patients also displayed polymorphisms at position 389 in p7 in addition to 3 polymorphisms in p6 at positions, 485, 488, and 489. Mutations reported to accumulate in Gag during therapy at positions 431, 432, 436, 449, and 453 (6,108,116) were not detected in this cohort.

**Discussion**

The accumulation of multiple amino acid substitutions in protease which cause resistance to protease inhibitors have been associated with failure to combination therapy (39,41,82,130,152,202,204). In our cohort patients with viral failure differed in the number of therapy specific mutations in protease compared to viral successes. Although classic drug resistant mutations are necessary they were insufficient to explain the difference between viral failure with and without immune reconstitution. Supplemental amino acid changes identified in this study distinguished viral failures from viral successes. Discordant patients accumulated more supplemental mutations, than either the VSIS or VFIF. Perhaps supplemental mutations contribute to viral fitness, which may in turn contribute to the persistently high viral load detected in discordant response patients. At the very least supplemental polymorphisms should be considered in the model for predicting therapy response.

The use of genetic analysis as a marker for therapy response was evaluated. Intrasample diversity cannot be used for prediction or identification of response to therapy. The trends for nonsynonymous and synonymous divergence coincided with therapy response. Despite selective pressure synonymous distance was the main contributor to genetic divergence in VSIS and VFIF response groups. The VFIF group
displayed high levels of viral replication but limited accumulation of new mutations. The lack of new protease mutations in VFIF patients may represent the existence of virus that is exquisitely able to replicate in the face of drug, perhaps due to pre-existing supplemental mutations in protease or secondary polymorphisms in Gag. For VSIS patients, viral suppression correlated to reduced nonsynonymous divergence. However, it is important to note that despite the effectiveness of therapy the median level of divergence at the synonymous level is approximately equal for each group, suggesting that replication continues within the VSIS group. Nonsynonymous divergence increased in patients that developed multiple mutations associated with resistance to protease inhibitors, and in this study predominantly reflected the VFIS response group. Evaluation of total, synonymous, and nonsynonymous divergence in patients during a period of up to 6 years prior to therapy demonstrated lower divergence than detected in virus from VFIS patients after only 24 weeks of therapy (data not shown), indicating the increase in divergence is directly related to combination therapy. Genetic distance (divergence) may be used to follow a discordant response.

Despite other reports of compensatory mutations within the Gag D’ and D cleavage sites (62,254), only 2 patients out of 14 patients demonstrated the D’ cleavage site mutation at position 431 while zero demonstrated the L453P mutation in the D cleavage site. Although the C cleavage site displayed the most natural heterogeneity, there appears to be a relationship between mutations at position p5, p4 and p3’viral failure. The direct impact of mutations at these sites should be considered in future experiments. One confounding problem in identifying mutations in Gag is the lack of knowledge concerning the contribution of pre-existing polymorphisms. Biochemical
analysis of the cleavage sites and mutations in Gag will provide insight to the types of mutations that warrant more consideration in a multivariate analysis. Evaluation of pre and post therapy alleles from a larger cohort may allow for the statistical evaluation of mutations relative to therapy response. As demonstrated with protease, mutations in Gag may be related to therapy response and thus should be tested in the predictive model.
Table 3-1. Clinical characteristics of study population at entry and 24 weeks posttherapy

<table>
<thead>
<tr>
<th>Therapy Response Groups</th>
<th>No. of subjects</th>
<th>Median (range) age, years</th>
<th>Median CD4 T Cells/µL&lt;sup&gt;a&lt;/sup&gt; Baseline</th>
<th>Median CD4 T Cells/µL&lt;sup&gt;a&lt;/sup&gt; 24 weeks PI</th>
<th>Median CD4/CD8 ratio&lt;sup&gt;a&lt;/sup&gt; Baseline</th>
<th>Median CD4/CD8 ratio&lt;sup&gt;a&lt;/sup&gt; 24 weeks PI</th>
<th>Log&lt;sub&gt;10&lt;/sub&gt; Plasma Viral RNA, Copies/mL Baseline</th>
<th>Log&lt;sub&gt;10&lt;/sub&gt; Plasma Viral RNA, Copies/mL 24 weeks PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSIS</td>
<td>5</td>
<td>10 (6-14)</td>
<td>256 (176-265)</td>
<td>524 (412-722)</td>
<td>.24 (.16-.26)</td>
<td>0.60 (.44-.76)</td>
<td>4.7 (4.7-5.4)</td>
<td>2.6 * ±</td>
</tr>
<tr>
<td>VFIS</td>
<td>10</td>
<td>6 (2-16)</td>
<td>83 (41-302)</td>
<td>930 (601-786)</td>
<td>.16 (.08-.5)</td>
<td>0.56 (.45-.66)</td>
<td>5.4 (5.1-5.7)</td>
<td>5.2 (4.9-5.4)</td>
</tr>
<tr>
<td>VFIF</td>
<td>4</td>
<td>12 (7-14)</td>
<td>24 (3-45)±</td>
<td>39 (3-75)±</td>
<td>.08 (.04-.12)</td>
<td>0.07 (.03-.11)</td>
<td>5.5 (5.5-5.6)</td>
<td>5.5 (5.3-5.6)</td>
</tr>
</tbody>
</table>

NOTE: VFIF, viral failure and immune failure; VFIS, viral failure and immune success; VSIS, viral success and immune success

<sup>a</sup> Median value (25<sup>th</sup>-75<sup>th</sup> interquartile range).

<sup>*</sup> Median was < 400 copies/mL

<sup>±</sup> Statistically significant difference
<table>
<thead>
<tr>
<th>Patient</th>
<th>Pre Genotype</th>
<th>Pl</th>
<th>Resistant Amino Acid Profiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSIS01</td>
<td>Sen RTV</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>VSIS02</td>
<td>Sen RTV</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>VSIS03</td>
<td>Sen RTV</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>VSIS21</td>
<td>Sen NFV</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>VSIS04</td>
<td>Res RTV</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>VFIS10</td>
<td>Sen RTV</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>VFIS11</td>
<td>Sen RTV</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>VFIS12</td>
<td>Sen RTV</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>VFIS14</td>
<td>Sen IDV</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>VFIS13</td>
<td>Res RTV</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>VFIS15</td>
<td>Res RTV</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>VFIS18</td>
<td>Res RTV</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>VFIS19</td>
<td>Res NFV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VFIS22</td>
<td>Res RTV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VFIS26</td>
<td>Res IDV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VFIF05</td>
<td>Res NFV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VFIF06</td>
<td>Res IDV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VFIF07</td>
<td>Res IDV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VFIF09</td>
<td>Res RTV</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-1. Genotype of HIV-1 protease following 24 weeks of combination therapy. Patients are classified by therapy response group and pretherapy protease genotype. PI (protease inhibitor) administered; IDV, indinavir; NFV, nelfinavir; RTV, ritonavir. White boxes, no resistance substitution; black boxes, new therapy specific mutations; black striped boxes, pre-existing therapy specific mutations; purple boxes, new therapy-nonspecific amino acid substitutions; purple striped boxes, pre-existing therapy-nonspecific amino acid substitutions.
Table 3-2. Statistical analysis of amino acid substitutions in HIV-1 protease associated with reduced sensitivity to protease inhibitors

<table>
<thead>
<tr>
<th>Median Number of Mutations</th>
<th>VS</th>
<th>VF</th>
<th>VFIF</th>
<th>VSIS</th>
<th>VFIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>New therapy specific mutations</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>p value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.021</td>
</tr>
<tr>
<td>Total therapy specific mutations</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>p value</td>
<td>0.003</td>
<td></td>
<td></td>
<td>0.027</td>
<td>0.003</td>
</tr>
<tr>
<td>New non-specific mutations</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>p value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total non-specific mutations</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>p value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative mutations</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>p value</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td>0.01</td>
</tr>
</tbody>
</table>
Figure 3-2. Number of new amino acid substitutions in protease after 24 weeks of therapy. Median number among VSIS, <1; VFIS ~1; VFIS, >3.
Figure 3-3. Supplemental amino acid mutations in HIV-1 protease.  

A. Consensus alleles for each patient.  White boxes, no polymorphisms relative to HIV\textsubscript{LAI}; solid blue boxes, supplemental polymorphisms that accumulated during 24 weeks of combination therapy; blue striped boxes, polymorphisms present prior to therapy, which also persisted.  

B. Statistical analysis of supplemental mutations between groups.  Pre, polymorphisms present prior to therapy; new, polymorphism accumulates during 24 weeks of therapy; total, all supplemental polymorphisms detected in alleles at 24 weeks.
Figure 3-4. Total intrapopulation genetic distance (diversity) in HIV-1 protease at baseline and after 24 weeks of combination therapy including a protease inhibitor. Patients were classified by therapy response group. A, VSIS; B, VFIS; C, VFIF. Symbols are designated next to the respective panel.
Figure 3-5. Total divergence in HIV-1 protease following 24 weeks of combination therapy including a protease inhibitor. A, VSIS; B, VFIS; and C, VFIF. Symbols for each patient are designated next to the respective panel.
Figure 3-6. Synonymous divergence in HIV-1 protease following 24 weeks of combination therapy including a protease inhibitor. A, VSIS; B, VFIS; C, VFIF. Symbols are designated next to the respective panels.
Figure 3-7. Nonsynonymous divergence in HIV-1 protease following 24 weeks of combination therapy including a protease inhibitor. A, VSIS; B, VFIS; C, VFIF. Symbols are designated next to the respective panel.
Figure 3-8. Contribution of nonsynonymous mutations relative to synonymous mutations in HIV-1 protease. A, VSIS; B, VFIS; and C, VFIF. Symbols are designated next to the respective panel.
Figure 3-9. Median genetic divergence following 24 weeks of combination therapy including a protease inhibitor. A, median total divergence; B, median synonymous distance; C, median nonsynonymous distance; D, median ratio of $D_N$ to $D_S$ for each response group. VSIS, diamonds (♦); VFIS, squares (■); and VFIF, triangles (▲).
Figure 3-10. Parsimony analysis of HIV-1 protease isolates prior to and following combination therapy. A and B, VSIS; C and D, VFIS; E and F, VFIF. Panels A, C and E analysis at the nucleotide level; B, D, and F analysis at the amino acid level. Figure key describes symbols for each response group. Closed symbols are pretherapy sequences, open symbols represent post therapy sequences. Numbers on tree branches refer to bootstrap values from 100 trees.
Figure 3-10--continued. (VFIS response).
Figure 3-10--continued. (VFIF response).
Table 3-3. Dominant C cleavage site sequence detected pre- and posttherapy

<table>
<thead>
<tr>
<th></th>
<th>Pre-therapy</th>
<th>24 Weeks Post therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SATIM*MQRGN</td>
<td>%</td>
</tr>
<tr>
<td></td>
<td>p5 p4 p3 p2 p1</td>
<td>p'1 p'2 p'3 p'4 p'5</td>
</tr>
<tr>
<td>VSIS01</td>
<td>- - A - - - K - -</td>
<td>80</td>
</tr>
<tr>
<td>VSIS02</td>
<td>- - - V - - - - -</td>
<td>100</td>
</tr>
<tr>
<td>VSIS03</td>
<td>- S I M - - - - -</td>
<td>100</td>
</tr>
<tr>
<td>VFIS22</td>
<td>- N - - - I - G - -</td>
<td>50</td>
</tr>
<tr>
<td>VFIS11</td>
<td>- - - V - - - K - -</td>
<td>100</td>
</tr>
<tr>
<td>VFIS12</td>
<td>- - - V - - - - -</td>
<td>57</td>
</tr>
<tr>
<td>VFIS15</td>
<td>- Q I V - - - K - -</td>
<td>83</td>
</tr>
<tr>
<td>VFIS13</td>
<td>- P A - - - - -</td>
<td>100</td>
</tr>
<tr>
<td>VFIS18</td>
<td>T - A M - - - - -</td>
<td>27</td>
</tr>
<tr>
<td>VFIS19</td>
<td>- - - - - - - K S -</td>
<td>100</td>
</tr>
<tr>
<td>VFIS10</td>
<td>- V N - - - - -</td>
<td>50</td>
</tr>
<tr>
<td>VFIS14</td>
<td>- T A - - - - - K - -</td>
<td>80</td>
</tr>
<tr>
<td>VFIF09</td>
<td>P T - - - - - - -</td>
<td>50</td>
</tr>
<tr>
<td>VFIS05</td>
<td>N S N - - - - - -</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 3-4. Percent of patients displaying polymorphism at the respective amino acid positions in the C cleavage site

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>A</th>
<th>T</th>
<th>I</th>
<th>M</th>
<th>*</th>
<th>M</th>
<th>Q</th>
<th>R</th>
<th>G</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P5</td>
<td>P4</td>
<td>P3</td>
<td>P2</td>
<td>P1</td>
<td>P1'</td>
<td>P2'</td>
<td>P3'</td>
<td>P4'</td>
<td>P5'</td>
<td></td>
</tr>
<tr>
<td>VSIS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre</td>
<td>0</td>
<td>33</td>
<td>66</td>
<td>66</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>33</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>0</td>
<td>33</td>
<td>66</td>
<td>66</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>33</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>VFIS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre</td>
<td>11</td>
<td>56</td>
<td>56</td>
<td>44</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>56</td>
<td>11</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>11</td>
<td>78</td>
<td>44</td>
<td>44</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>56</td>
<td>11</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>VFIF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>50</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 4
SEGREGATION OF HIV-1 PROTEASE GENOTYPIC RESISTANCE WITHIN BLOOD COMPARTMENTS IDENTIFIES CD4+ CD45RA T LYMPHOCYTES AS A LONG LIVED VIRAL RESERVOIR.

Introduction

A viral reservoir can be defined as a cell type or anatomical location in which a replication competent form of HIV-1 accumulates and persists with more stable kinetics than the main pool of actively replicating virus (15). As a consequence of reservoir stability, viral species arising at various times in the infection should be present in the reservoir. In short, recent and archival sequences should not only be detected but also capable of being released (ie replication competent) (15). A number of cell types have been implicated as viral reservoirs, including macrophages, resting memory CD4+ T cells, and follicular dendritic cells.

Targets for HIV-1 infection include thymocytes, macrophages and peripheral CD4+ and CD8+ T cells. Both memory and naïve CD4+ T cells can be infected. However, CD4+ memory cells are the predominant cell population infected with HIV-1 in both children and adults (170,217). HIV-1 also preferentially replicates in cells with a memory phenotype (205,250).

Varying molecular weight isoforms of the CD45 membrane protein (31) expressed on CD4+ T cells have been associated with cell differentiation. Cells expressing the high molecular weight (CD45RA) isoform are frequently considered immature or “naïve” while cells expressing the low molecular weight isoform (CD45RO) are identified as memory CD4+ T cells (31,206). A number of studies suggest that
CD45RO and CD45RA expression reflects the activation state of cells rather than differentiation (11,12,85). Nevertheless, most researchers in the field believe that a change in expression from the CD45RA isoform to the CD45RO isoform simply reflects a unidirectional differentiation into a memory phenotype.

Coreceptors used for HIV-1 infection are differentially expressed on CD45RO and CD45RA cells. The dominant coreceptor for macrophage tropic viruses, CCR5, is largely restricted to expression on the CD45RO T cells (1,16,53,58,59,63). CXCR4, the dominant coreceptor for T cell tropic viruses, (42,72,215) is found in greater levels on naïve cells, although memory cells also express CXCR4 (17,148).

Despite the apparent cellular niches established on the basis of coreceptor expression, NSI [CCR5 utilizing] virus is found in CD45RA T cells prior to conversion to an SI, CXCR4-using phenotype (236). SI virus can be found in both the CD45RA and CD45RO cells (14). Currently there is no explanation for the mechanism of entry for these viruses. Evidence suggests that SI and NSI viruses evolve independently. Yet, there is no phenotypic or genotypic compartmentalization of viruses between CD4⁺ T cell subsets, suggesting a dynamic exchange of viruses in vivo within the CD45RA and CD45RO T cell subsets (170,236).

The disadvantages of utilizing envelope genotype or phenotype for evaluating viral trafficking in CD4⁺ T cell subsets include the lack of sequential polymorphisms and the difficulty in detecting intermediates. In contrast, when protease is under the selective pressure of antiretroviral therapy there is an accumulation of multiple mutations that persist with a maintained therapy regimen. Therefore protease genotypic resistance provides a better marker for evaluating viral trafficking in CD4⁺ T cells.
To explore the trafficking of virus within the plasma, CD4+ CD45RA and CD45RO T cell subsets, we evaluated pro genotype in a panel of pediatric patients following treatment with combination therapy, which included a protease inhibitor. Envelope genotype was also evaluated in representative pediatric discordant response patients to establish that env did not segregate in this cohort according to cell phenotype.

**Materials and Methods**

**Study Patients and Samples**

Peripheral blood samples were collected from HIV-1 infected pediatric patients and were processed as described in Chapter 2 of this dissertation.

**Separation of CD4+ T Cell Subsets**

CD4+ T cell subsets, CD45RA+ and CD45RO+, were separated from either freshly processed or cyropreserved samples using microbeads with the appropriate monoclonal antibody (118,217). Briefly, CD4+ T cells were positively selected, followed by positive selection of CD45RA cells. Since the majority of CD45RA cells in pediatric patients co-express CD62L (191) and a limited number of cells were available, CD62L was not used for cell separations. Therefore the CD45RA cell population consists of cells that co-express CD45RA and CD45RO. Nucleic acids were extracted by incubation of cells in K buffer [50mM KCl, 10mM Tris/Cl, pH 8.3, 2.5mM MgCl2, 0.1mg/ml gelatin, 0.45% NP40, 0.45% Tween 20, plus 100mg/ml fresh proteinase K] for two hours at 55°C. Proteinase K was heat inactivated by incubation at 95°C for 5 minutes.

**Generation of Protease Sequence**

Allele specific protease sequences were generated from the plasma, unfractionated PBMC (peripheral blood mononuclear cells), CD4+ CD45RA and CD4+
CD45RO T cells. Methods for amplification cloning and sequencing are described in Chapters 2 and 3 of this dissertation.

**Generation of Envelope Sequence**

Allele specific envelope sequences were generated from the CD4\(^+\) CD45RA and CD45RO cells. A 1.2 kb region of the HIV-1 envelope including hypervariable regions V1-V5 were amplified using primers LV15 (5’-GCCACACATGCCTGTGTACCCACA-3’, nucleotides 6432-6456 in HIV\(_{HXB2}\)) and 194G (5’-GGTAGAACAGATGCATGAGGAT-3’, nucleotides 7666-7645 in HIV\(_{HXB2}\)) for two rounds of PCR. PCR amplifications were carried out in a similar fashion as the *gag-pol* amplifications with the exceptions of the magnesium concentration (1.75 mM) and the thermal cycler conditions, (annealing temperature 55°C and extension time to 2 minutes).

**Genetic Analysis**

Phylogenetic analysis was performed using neighbor-joining and maximum parsimony in the PHYLIP software package (71). Bootstrap values were based on generation of 100 replicate trees. The integrity of genetic data was verified by using amino acid alignments and construction of phylogenetic trees, to compare each sequence with all *pol* or *env* sequences generated in our laboratory. Any detectable amino acid polymorphisms at positions in PR that are associated with reduced sensitivity to drugs were scored as resistant in the same manner as previous chapters.

Envelope net V3 charge was determined by subtracting the negatively charged amino acid residues aspartic acid (D) and glutamic acid (E) from the sum of the positively charged amino acid residues lysine (K) and arginine (R). \([(K+R)-((D+E))].\)
Culture of CD4+ T Cell Subsets

CD45RA and CD45RO cell subsets (either 2 x 10^5 or 5 x 10^5 cells, depending on the experiment) were cultured with an equal number of uninfected target PBMC that were stimulated with phytohemagglutinin (PHA) and cultured in RPMI 1640 containing 15% FCS and interleukin-2 (30 U per ml). Supernatants were collected every 3-4 days and replaced with media alone or media with new PBMC (2 x 10^5 cells) once a week. Supernatants were stored at -80°C, and analyzed for p24 antigen by ELISA (Coulter, Miami, FL). After 30 days DNA was extracted, amplified and sequenced as described in Chapter 2 of this dissertation.

Results

Identification of Patients for Separation of CD4+ T Cells Subsets

The cohort for this study was a panel of HIV-1 infected pediatric patients demonstrating sufficient reconstitution in CD4+ T cells for separation in combination with the accumulation of genotypic resistance mutations in PR. Although some viral success and immune success patients meet this criteria, the high rate of viral replication and the accumulation of PR mutations in discordant response (viral failure/immune success) patients (184) make them an ideal population for the study of viral trafficking by protease genotypic resistance. The cohort displayed a reduction in viral burden, which was transient and rebounded to baseline levels by 16 weeks of therapy (Figure 4-1 A). Despite poor viral response to therapy, patients displayed sustained increases in CD4+ T cell counts (Figure 4-1A). Patient VFIS15, who had the lowest CD4+ T cell count prior to therapy demonstrated an initial rebound of predominantly CD45RO cells followed by an increase in the CD45RA cells (Figure 4-1 B panel 1). VFIS12 and VSIS03 also demonstrated increases in both cell subsets (Figure 4-1 B panels 2 and 3). The increase
in total CD4\(^+\) T cell count and reconstitution of both cell subsets made separation of the CD4\(^+\) CD45RA and CD45RO cells possible.

**Envelope Genotype Demonstrates Independent Viral Evolution Based on Coreceptor Use but Not CD4\(^+\) CD45 T Cell Subsets**

Viral V3 envelope alleles isolated from the CD4\(^+\) CD45RA and CD45RO cells were evaluated using parsimony analysis. Sequences with a low V3 net charge (3 or 4) appear on an independent branch from those with high V3 charge (≥5) in VFIS15 (Figure 4-2). V3 envelopes with low charge are predominantly identified in the CD45RO cell subset, while alleles with high charge are found in both the CD45RA and CD45RO cells. However, a low charge envelope allele detected within the CD45RA subset after 52 weeks of therapy was found to segregate with CD45RO alleles (arrow on bottom branch). Low charge alleles in the CD45RA cells that are phylogenetically related to alleles identified in CD45RO cells have been detected in additional patients with a frequency between 6-10% (data not shown). Based on viral tropism it is unlikely that de novo infection would explain the appearance of a low charge envelope allele in the CD45RA subset. One explanation is the reversion of cells expressing CD45RO to a CD45RA phenotype. Virus detected in the CD45RO cells (top branch) gives rise to envelope alleles with a high (≥ 5) charge (Figure 4-2 A). Alleles from both the CD45RA and CD45RO cells with high charge were detected on the top branch. In concordance with other studies (236), high and low charge virus evolve independently although independent evolution is not detected within the CD4\(^+\) T cell subsets.

The predominant net V3 envelope charge in patient VFIS12 was five. A charge of 5 would allow for dual tropism, or the ability to infect both CD4\(^+\) T cell subsets. Virus
that can equally infect the two cell populations is unlikely to segregate by cell type. This is the case for the envelope tree for VFIS12.

**Pretherapy Protease Genotype Is Indistinguishable Between Compartments**

Pretherapy HIV-1 protease genotype was evaluated in 11 patients that were immune suppressed and naïve to protease inhibitor therapy. Analysis focused on amino acid positions in PR that are related to reduced sensitivity to PIs (204). Pretherapy polymorphisms detected at amino acid positions 10, 36, 63, 77, and 82 were mainly therapy non-specific (Figure 4-3). Protease inhibitor sensitive alleles were detected in both the plasma and PBMC in four patients. Protease alleles in two patients, VFIS15, and VSIS21, displayed therapy specific mutations in both PBMC and plasma. The only instance of discordance between compartments prior to therapy was in patient VFIS14 within whom an allele containing a therapy specific polymorphism was detected in the plasma but not the PBMC. Protease genotype within the CD4⁺ CD45RA and CD45RO cell subsets was found to be identical in four patients. Prior to the initiation of PI therapy polymorphisms in protease cannot be used to distinguish between compartments.

**Diversity of Protease Alleles Within Compartments Following 24 Weeks of Therapy**

Following 24 weeks of therapy, protease genotype in the plasma and PBMC of six patients was evaluated for therapy specific mutations. Alleles between the compartments were identical in VFIS14 and VFIF17 (Figure 4-4). Plasma alleles from the other four patients had numerous therapy specific mutations (range 3 to 7). In contrast, alleles isolated from the PBMC were more diverse in these patients, with therapy specific mutations ranging from 0 to 7. Pretherapy protease alleles persisted within the cellular compartments of VFIS12 and VFIS15 despite the appearance of alleles with multiple therapy specific polymorphisms (data not shown). To determine if the presence of two
distinct genotypes within PBMC represents infection of cellular niches we evaluated protease genotype in the CD4$^+$ CD45RA and CD45RO cells.

**PR Genotype Differs in CD4$^+$ T Cell Subsets Following 24 Weeks of Therapy**

Protease alleles from the CD4$^+$ CD45RA and CD45RO cells were evaluated in 6 patients (2 VSIS, and 4 VFIS). PR alleles in the CD45RA cells of VSIS02 after 20 weeks of therapy were identical to pretherapy alleles (Figure 4-5 A), while alleles in the CD45RO cells accumulated therapy specific mutations at amino acid positions 54, 82, and 84. Following 24 weeks of therapy, plasma alleles from VFIS12 were more polymorphic than those found in either cell subset (Figure 4-5 B). Therapy specific mutations detected in the CD45RO T cell subset reflected a portion of those seen in the plasma. Virus harboring resistance associated mutations were detected in the plasma and CD45RO cells following 24 weeks of therapy while the appearance of resistance associated mutations in the CD45RA was minimal (VFIS15) or non-existent (VSIS02 and VFIS12) (Figure 4-5). Possible explanations include independent protease evolution within the CD4RA and CD45RO cells, or viral trafficking within subsets.

To establish if protease genotypic segregation in the CD4$^+$ T cell subsets represents a transient event protease alleles from VFIS15 were evaluated after 52 weeks of therapy (Figure 4-5 C). Alleles with a high number of therapy specific mutations, like those found in the CD45RO cells and plasma at 24 weeks were now detected in the CD45RA cells. Despite the appearance of new alleles in the CD45RA cells there is persistence of alleles identified 6 months earlier, suggesting that these cells serve as a reservoir for HIV-1.

Protease genotypic resistance serves as a good marker for segregation within compartments before a steady state is re-established. The kinetics of protease allele
steady state re-establishment is variable. Re-establishment occurred quickly, in less than 28 weeks, between the cells and plasma for VFIS11 (data not shown), while a new steady state did not occur within the CD4+ T cell compartments until after 52 weeks of therapy for VFIS15.

**CD4+ CD45RA T Cell Subset: Reservoir for HIV-1**

The significance of viral genotype within the CD4+ CD45RA cells has frequently been questioned. One possibility is that virus harbored within CD45RA cells is not replication competent, or in other words a dead end virus. To address this concern CD45RA and CD45RO cells from 3 patients (in four experiments) were cultured in the presence of uninfected target cells. Supernatants collected from both cell subsets were positive for p24 production in each experiment (data not shown), indicating that virus within each compartment capable is replication competent.

*Gag-pro* genotype was evaluated following culture termination in each experiment. After 83 weeks of therapy, the CD45RA cells from patient VFIS15 harbored a single allele with many therapy specific mutations (Figure 4-6). The CD45RO cells displayed numerous alleles, harboring mutations similar to those in the CD45RA cells, in addition to polymorphisms at amino acid positions 46, 60, and 90. This new segregation of alleles within the cell subsets resulted from of a switch in the patient’s drug regimen from RTV/D4T to IDV/3TC/ZDV 15 weeks prior to this sample collection. Weeks of therapy represent a summation of all weeks of protease inhibitor exposure, therefore 83 and 92 weeks of therapy represent 15 and 24 weeks respectively on the new protocol. Despite the appearance of protease allele steady state within this patient’s CD45RO and CD45RA prior to protocol switch there is once again the accumulation of new polymorphisms in the CD45RO cells, followed by the accumulation of similar alleles in
the CD45RA cells after an additional 9 weeks of therapy. The M36I mutation, which confers decreased sensitivity to the first PI but not the new PI, was no longer detected in the CD45RO cells, however this polymorphism was detectable within the CD45RA population following 24 weeks of therapy with the second PI (Figure 4-6). CD4+ CD45RA cells produce virus that is genotypically different from virus in CD45RO cell subset, identifying it as a unique reservoir that can harbor replication competent virus for at least 6 months if not longer, making it an important source for virus exhibiting resistant alleles from past therapies.

**Protease Genotype Indicates Viral Trafficking Thru the Blood Compartments**

To establish if PR alleles isolated from the CD45RA subset result from a long lived cellular reservoir or independent evolution of virus within the cell subset a parsimony analysis of protease alleles from patient VFIS15 was used (Figure 4-7). Pretherapy alleles (stars) appear near the root of the tree, with one allele giving rise to all plasma and CD45RO cell alleles detected at 24 weeks. In contrast, alleles identified in the CD45RA cells at 24 weeks were closely related to pretherapy alleles and are detected closer to the root of the tree. Following 52 weeks of therapy, alleles identified in the CD45RA cells were related to pretherapy alleles, as well as alleles detected in the CD45RO cells at 24 and 52 weeks of therapy. Evolution of protease from both subsets on a common branch of the tree, suggests that the virus is trafficking between the compartments. The persistence of alleles closely related to baseline in the CD45RA cells for 6 months to a year identifies this cell subset as a long-lived reservoir for virus.

Following a switch in inhibitor regimen virus in the CD45RO cells evolved rapidly, while virus in the CD45RA subset continued to be phylogenetically closer to protease alleles from earlier samples than to alleles from CD45RO cells at the same time.
sample collection. After additional weeks on therapy a new steady state is achieved between compartments with the alleles evolving accordingly.

**Discussion**

HIV-1 genotype may be used as a probe in human cells. Mutations to the viral genome may provide information about normal processes in a manner similar to the use of human DNA mutations in cancer patients. The use of protease genotypic resistance requires persistent replication and the accumulation of polymorphisms while under the continued selective pressure of protease inhibitors. These conditions are met in patients who demonstrate a discordant response to therapy and to some extent those experiencing both viral and immune success. The observation that HIV-1 protease genotypic resistance can be detected first in the plasma, followed by the CD4\(^+\) CD45RO and finally in the CD45RA cells, suggests trafficking of virus within blood compartments. Two mechanisms that may explain trafficking of virus include the evolution of envelope or the reversion of CD45 isoform on CD4\(^+\) T cells.

In concordance with other studies, we found that high and low charge viruses evolve independently, although independent evolution was not detected within the CD4\(^+\) T cell subsets (236). Genetic modifications in *env* do not alter tropism in discordant patients following therapy (232), suggesting that viral trafficking of protease resistant virus is independent of envelope.

Virus entry into CD45RA cells may occur in one of three ways. The first is *de novo* infection with a virus using the CXCR4 co-receptor. However, this route of infection does not explain how a low charge virus, which would use the CCR5 coreceptor, gains entry into CD45RA cells. The second possibility is that low charge virus may gain access to CD45RA cells during thymopoiesis. Brooks et al. demonstrated
in a SCID-hu model that both CCR5 and CXCR4 virus could infect thymocytes that were exported to the periphery (20). While thymocyte infection could indeed explain viral entry into CD4$^{+}$ CD45RA cells, it does not account for the temporal segregation of PR genotype within the compartments. If indeed the CD45RA cells were infected as thymocytes, then virus in that cell subset would be expected to be genotypically identical to virus detected in the CD45RO cells. The third possibility is that CD45RO expression back reverts to CD45RA, which is supported by the appearance a low charge envelope allele in CD45RA cells (Figure 1B branch 2). Back reversion of CD45 cell phenotype would also account for the appearance of CD45RA protease alleles at 52 weeks giving rise to alleles detected at 24 weeks in the CD45RO cells. Isoform reversion would also explain the persistence of PR sensitive alleles in the CD45RA cells, in addition to alleles with an intermediate number of polymorphisms at 24 weeks of therapy. Phylogenetically, protease alleles from CD45RA cells are more closely related to pretherapy alleles, suggesting the involvement of a long lived cell rather than independent viral evolution.

The controversial concept that a move in the expression from CD45RA to the CD45RO isoform expression reflects a unidirectional differentiation to the memory phenotype was addressed by the work in this chapter. Our study provides direct evidence to suggest reversion occurs in humans. CD45RA cells do not express sufficient quantities of the CCR5 coreceptor to allow infection by virus expressing a low V3 charge, yet we and others have identified low charge virus within the CD45RA cells. This study identified alleles in the CD45RA cells that could not have gained entry via de novo infection and have segregated phylogenetically with alleles detected in the CD45RO
population from an earlier timepoint. The half-life of the CD45RA and CD45RO cells in humans has been shown to vary dramatically. The persistence of dicentric chromosomes from x-ray irradiated patients in the CD45RO lymphocytes were lost within a year, while persistence occurred for up to 10 years in the CD45RA cells (146). The extended life of CD45RA expressing cells suggests that, over a period of up to ten years, any genotypic mutations detected within a patient may persist within this cell population.

Due to evidence that CD45RA and CD45RO expression is not limited to naïve and memory cells, respectively, it has been proposed that CD62L should be used in combination with CD45RA to identify truly naïve CD4\(^+\) T cells. The use of CD62L, however, is limited as CD62L is a labile marker and its expression may change depending on the handling of cells, including cryo-preservation and room temperature storage overnight prior to processing (89). Additionally, CD62L expression on cells is a continuum, ranging from bright to dim. Flow cytometry can only denote CD62L as being present or absent, which could lead to inappropriate cell sorting. In contrast to low expression in adults, greater than 80% of CD4\(^+\) CD45RA cells coexpress CD62L in children (191). The fact that most CD45RA cells co-express CD62L in children, in combination with the limited quantity of blood that can be obtained at a given time, we limited our cell separations to CD45 isoform expression alone.

The unique genotype trafficking and evolution observed could result from contamination of the CD45RA subset with CD45RO cells. To address this possibility, we considered the number of infected cells and the level of CD45RO cell contamination that would be necessary for detection. Discordant and success patients averaged 1 infected cell per 200 CD4\(^+\) T cells between 24 and 52 weeks of therapy (169). During
this timeframe, the CD4 T cell population displayed approximately equal distributions of CD45RO and CD45RA cells. Given that 200,000 cells were cultured, approximately only 1,000 of these cells would have been infected. Detection by our cloning and sequencing procedures would have required such extensive contamination by CD45RO cells the separation procedure would have been considered a failure. The success of our separations were demonstrated by flow cytometry as well as through experiments that demonstrated CD45RA and CD45RO cell subsets have distinct T cell receptor diversity (118). Multiple criteria argue that this data is not the result of subset contamination.

CD4⁺ CD45RA cells fit the criteria for a viral reservoir. These cells are long-lived and can harbor replication competent HIV-1 displaying both archival and current protease genotypes. Persistence of archival sequences is based on the cell survival rate. With a half-life of 10 years, CD45RA cells provide a location where archival virus could exist. The significance of CD4⁺ RA T cells subsets serving as viral reservoirs has not yet been explored. CD4⁺ CD45RA cells should be considered when targeting the resting pool of infected cells.
Figure 4-1. Viral and immune response to combination therapy for HIV-1 infection. A, Median log₁₀ HIV-1 RNA copies per milliliter (■) are indicated on left axis. Median CD4⁺ T cell counts per microliter (▲) are indicated on the right axis. VFIS patients, n=11. B, Absolute CD4⁺ T cell count on the y axis. Percent CD45RA cells (■), and percent CD45RO cells (▲). Panel 1 VFIS12; Panel 2, VSIS02; Panel 3, VFIS15.
Figure 4-2. Phylogenetic analysis of envelope V3. Numbers following a symbol represented the calculated net V3 charge.
Figure 4-3. HIV-1 protease genotype prior to combination therapy including a protease inhibitor. Patients are designated with the same numbers as previous chapters. Amino acid positions in protease are designated by numbers and were related to resistance (204). White box, no resistance substitution, gray box (■) therapy nonspecific amino acid substitution, black box (●), therapy specific amino acid substitution. Source of DNA for genetic analysis; p, PBMC; r, plasma RNA; o, CD4+ CD45RA T cells; and o, CD4+ CD45RO T cells.
Figure 4-4. Presence of alleles with therapy specific mutations following 24 weeks of combination therapy. Independent symbols were used for each patient. Open symbols, alleles obtained from peripheral blood mononuclear cells. Closed symbols, alleles obtained from plasma RNA.
Figure 4-5. Protease genotypic resistance segregates within CD4⁺ CD45 T cell subsets. The number of pre- and post therapy alleles are indicated. Amino acid mutations in protease are designated across the top. White box, no resistance substitution; black box, therapy specific mutations; and gray shaded box, therapy nonspecific amino acid substitutions.
Figure 4-6. Distinctive PR alleles were identified in cultured cell subsets. The total weeks on therapy, in addition to the time on each protease inhibitor are presented in the first column. White box, no resistance substitution; black box, therapy specific mutations; gray shaded boxes, therapy nonspecific amino acid substitutions; striped boxes are supplemental polymorphims.
Figure 4-7. Phylogenetic tree of protease alleles obtained from patient VFIS15. Open symbols, alleles obtained from CD4⁺ CD45RA cells; closed symbols, CD4⁺ CD45RO; and hatched symbols, plasma. Pretherapy alleles isolated from unfractionated PBMC.
CHAPTER 5
REPLICATION OF HIV-1 IS INFLUENCED BY \textit{GAG-PRO} AND TARGET CELL TYPE

\textbf{Introduction}

Upon considering the therapy response groups, it is intuitive that patients with viral success, or reduced levels of viral replication, would also demonstrate a reduction in the death of CD4$^+$ T cells. Conversely, patients demonstrating a viral failure, or continued high levels of viral replication, would display a continued decline in the CD4$^+$ T cells, or immune failure. However, the viral failure immune success (VFIS) response, which is characterized by restoration of CD4$^+$ T cells despite high viral replication, is paradoxical. A number of factors could contribute to this phenotype. In addition to the patient’s immune response and changes in the virus itself, a reduction in the pathogenic properties of rebounding virus may also occur. Unpublished studies by Tuttle et al. have demonstrated that viruses in VFIS patients following therapy do not exhibit changes in envelope that would alter or extend tropism (232). VFIS patients might develop different patterns of Gag-PR mutations in response to therapy that would modulate fitness.

Fitness can be defined in strict Darwinian terms as the selective advantage for replication among viral variants competing in a particular environment. The definition of fitness is less strict in the current literature surrounding HIV-1, and is used to describe any competitive advantage or disadvantage of a drug resistant variant relative to wild-type or other mutants in the presence or absence of drug (135). The latter definition of fitness will be used in this dissertation. Mutations in protease associated with reduced
sensitivity to protease inhibitors translate to reduced replicative ability or fitness (19,44,94,137,166,193,252). Polymorphisms in \textit{gag} can restore fitness to viruses that harbor protease mutations (62,134,254). A recent study by Stoddart \textit{et al.} demonstrated that viral fitness can be related to determinants in protease, and that although viruses were moderately impaired for replicative ability in PBMC, a more pronounced effect in thymic organ cultures was observed (224). Viruses that rebound following treatment interruption were more fit than viruses isolated during protease inhibitor treatment (224).

VFIS patients have reduced levels of infected CD4\(^+\) T cells during therapy (169). If drug resistant viruses have diminished replicative ability in thymocytes, and a lower number of CD4\(^+\) T cells are infected, then drug resistant virus production must originate from other compartments or cells, such as macrophages, to account for the persistence of high levels of plasma virus in the VFIS response. Low levels of HIV-1 infection in peripheral CD4\(^+\) T cells, despite high levels of plasma virus, implies a restriction in host cell range for virus within VFIS patients that extends beyond CD4\(^+\) thymocytes to include peripheral CD4\(^+\) T cells. Chapter 3 of this dissertation demonstrated that the VFIS patients accumulated more therapy specific and supplemental mutations in protease than the other response groups. Drug resistance and fitness are critical components of the VFIS response.

These studies were designed to evaluate the impact on Gag-PR on viral fitness. Three main parameters were evaluated: one, the role of natural polymorphisms, which impact proteolytic processing, on the modulation of viral replication; two, the impact of post therapy Gag-PR regions on fitness relative to the patient’s pre Gag-PR; and three, the replicative potential of multi-drug resistant Gag-PR recombinant viruses within
peripheral blood mononuclear cells and macrophages. Our approach has several novel aspects: one, recombinant viruses were constructed with $gag$-$pro$ regions from patient isolates while maintaining reverse transcriptase from the molecular clone; two, wild type protease alleles were obtained pretherapy from protease inhibitor naïve patients, rather than from patients whose virus rebounded following therapy interruption; and three, $gag$-$pro$ recombinant viruses were constructed with select envelopes having specificity for CCR5- or CXCR4-expressing target cells, including subsets of lymphocytes and monocyte-derived macrophages. A portion of this work has been published in Virology, volume 292 in 2002 (77).

**Materials and Methods**

**Recombinant Viruses**

HIV-1 $gag$-$pro$ patient isolates were obtained by cloning into pGEM-T vector system™ (Promega Madison, WI) a 1.7kb region that was amplified from peripheral blood mononuclear cells (PBMC), as described previously (9). The HIV-1 molecular clone pLAI.2 (provided generously by Keith Peden) was used for construction of recombinant viruses (179). This 11.8 kb plasmid consists of 2 kb of vector with HIV-1$_{LAI}$ and minimal host cell DNA (179). Unique restriction sites were sought for direct insertion of patient alleles into the pLAI.2 background. The BSTZ1107I site at nucleotide position 3031 and the SpeI at nucleotide position 1553 were unique in the viral sequence however, an SpeI site was present in cellular material included in the plasmid. The SpeI site upstream of the 5’ LTR at nucleotide position 7 was removed by oligo-linker insertional mutagenesis to produce pLAI.4 (Figure 5-1).
**Generation of pLAI.4**

A 4.6kb fragment of pLAI.2 (nucleotides positions 8568 to 1493), which included the targeted SpeI restriction site, was digested using BamHI (New England Biolabs, Inc. Beverly, MA, NEB) and SphI (NEB) and cloned into pGEM-3Z (Promega) for mutagenesis. Conveniently the polylinker region of pGEM-3Z contained both BamHI (NEB) and SphI (NEB) restriction sites. After digestion, the 2.7Kb fragment of pGEM-3Z was treated with calf alkaline phosphotase (CIAP) (Promega) to prevent plasmid religation. The 4.6kb fragment of pLAI.2 was ligated into the 2.7kb fragment of pGEM-3Z using T4 DNA ligase (Promega). The resulting 7.3 kb plasmid, referred to as pGEM-3Z-4.6kb LAI (Figure 5-1) was transformed into DH5α cells (Gibco-BRL, Rockville, MD), plated on Luria broth with ampicillin (100mg/ml) and incubated overnight at 37°C. Ten clones were selected and grown in Luria broth with ampicillin (100 µg/ml) at 37°C overnight. Plasmids were extracted using QIAprep™ Miniprep Kit (QIAGEN, Valencia, CA) following the manufacturer’s protocol. Insertation of the 4.6 Kb fragment was confirmed by restriction digest using BamHI and SphI.

To prepare for mutagenesis pGEM-3Z-4.6kb LAI was digested with SpeI (NEB), CIAP (Promega) treated, and gel purified. Oligos were designed to incorporate the digested SpeI site with the insertion of an ApaI site (Figure 5-2). Primers ΔSpeI + ApaI forward (Gibco #213, CTAGGGGCCC) and ΔSpeI + ApaI reverse (Gibco #214, CCCGGGGATC) were annealed at 85°C for 5 minutes and slow cooled to room temperature. Annealed oligos were ligated into the SpeI digested pGEM-3Z-4.6kb LAI with T4 DNA ligase overnight at 4°C. Transformation into DH5α cells, overnight cultures and plasmid preps were described above. Clones were digested with SpeI or
ApaI to confirm the inability of SpeI to digest the plasmid and the addition of the ApaI site. The mutagenized 4.6kb fragment of pLAI.2 was digested out of pGEM-3Z with SphI and BamHI, CIAP treated and re-ligated into the 7.1kb pLAI.2 fragment created earlier. Mutagenesis of the SpeI site and insertion of the ApaI was confirmed in clones by restriction digest and sequence analysis. The resulting plasmid was named pLAI.4.

**Generation of Patient Gag-pro Replication Competent Recombinant Viruses**

**Strategy 1**

Patient *gag-pol* regions (1.7kb fragment) were excised from pGAGPOL (obtained from Dr. Ben Dunn) by digestion with SpeI and BstZ1107I, and ligated into pLAI.4 (Figure 5-3A). The plasmids were transformed, and prepped as described above. The proper insert was confirmed with restriction digest and sequence analysis.

**Strategy 2**

A second strategy for recombinant viruses was necessary because many *gag-pol* alleles of interest were initially amplified using alternate primers that generated an 830bp fragment. Unique restriction sites were not available so a shuttle vector was constructed by inserting the region of pLAI.4 from SphI (nucleotide 1493) to EcoRV (nucleotide 3061) into pGEM5Zf+ (Promega)(Figure 5-4). Site directed mutagenesis was used to introduce an XbaI site at nucleotide position 1861 (pLAI.2) and a KpnI site at nucleotide position 2652. The introduction of these restriction sites did not generate a change at the amino acid level. Primers designed in Dr. Dunn’s laboratory for mutagenesis are as follows: HIVXbaU2, 5’-CCA GCG GCT ACT CTA GAA GAA ATG-3’; HIVXbaCU, 5’-CAT TTC TTC TAG AGT AGC CGC TGG-3’; HIVKpnUM, 5’-GCC CTA TTG AGA CGG TAC CAG TAA A -3’; and HIVKpnLM, 5’-TTT ACT GGT ACC GTC TCA ATA GGG C-3’. The protocol accompanying Strategene’s Site directed mutagenesis kit
was used to first insert the XbaI site and then the KpnI. The double mutation was confirmed by restriction digest and sequence analysis.

For insertion into the shuttle vector patient *gag-pol* alleles were amplified with primers designed in the Dunn lab to generate a fragment with an XbaI site at the 5’ end (forward primer HIVXbaU2, 5’-TAC TCT AGA AGA AAT GAT GAC-3’) and a KpnI site at the 3’ end (reverse primer, KPNL2 5’-GGT ACC GTC TCA ATA GGG CTA ATG GG-3’). Ligations, transformations and screening of clones were performed as described previously. A 1.7kb fragment from SpeI to BstZ1107I, which encompasses the 830 bp patient *gag-pol* fragment was digested and inserted into pLAI.4 (Figure 5-3B). A total of eleven *gag-pol* recombinant viruses were constructed with the HIV_{LAI} envelope (Table 5-1).

**Generation of Envelope Recombinant Viruses**

Recombinant viruses with the V1-V5 of HIV_{JRFL} envelope were generated in the pLAI.4 background. Unique restriction sites were not available so a shuttle vector was constructed by inserting a 2.7 kb fragment of pLAI.4 from restriction sites SalI (HIV_{LAI.2} nucleotide 5867) to BamHI (nucleotide 8568) into pGEM-11Zf(-) (Promega). Within the 2.7 kb fragment restriction sites KpnI (HIV_{LAI.2} nucleotide 6429) and MfeI (nucleotide 7748) flank the desired envelope region. Amplification of the HIV_{JRFL} envelope with primers D1 (5’-CACAGTCTATTATGGGGTACCTGTGGAA-3’, nucleotides 6332-6362 in HIV_{HXB2}) and 194G (5’-GGTAGAACAGATGCATGAGGAT-3’, nucleotides 7666-7645 in HIV_{HXB2}) introduced the KpnI and MfeI restriction sites. The 1.3 kb HIV_{JRFL} envelope fragment was ligated into the shuttle vector. The 2.7 kb fragment was digested from the shuttle vector and ligated into the pLAI.4 *gag-pol* recombinant backbone. The replacement of HIV_{LAI.2} envelope with HIV_{JRFL} envelope was confirmed
by sequence analysis. A total of twelve gag-pol recombinant viruses were made with the HIVJRFL envelope (Table 5-1).

**Production of Recombinant Virus Stocks**

Virus stocks were generated by transfection of HEK293 cells with the protease recombinant construct. The day prior to transfection, 3.9 x 10⁶ 293 cells were seeded in Dulbecco’s modified Eagle’s media (DMEM) (Gibco-BRL) supplemented with 10% fetal calf serum (FCS) (Gibco-BRL) (Rice and Mathews, 1988), penicillin (100U/ml), and streptomycin (100µg/ml)(Gibco-BRL), into T75 flasks. The cells were incubated at 37°C, 5% CO₂ for 24 hours to yield approximately 50-60% confluency. On the day of transfection 15µg protease recombinant construct was diluted into a total of 700µl unsupplemented DMEM with 15µg of Superfect transfection reagent (QIAGEN). This transfection mixture was vortexed for 10 seconds and incubated at room temperature for 10 minutes. During this incubation, media was removed from the flasks containing the 293 cells and washed 1 time with 15ml of 1X phosphate buffered saline (PBS). After washing, 6 ml of supplemented DMEM was added to each flask. The transfection mixture was added directly to the cells in the T75 flask and incubated at 37°C, 5% CO₂ for 3 hours. Following the incubation, media was removed and the cells were washed twice with 15 ml of 1X PBS. Six ml of supplemented DMEM was added to each T75, and incubated at 37°C, 5% CO₂ Twenty-four hours post-transfection the supernatant was removed and 6 ml fresh supplemented DMEM was added. Supernatants were collected after 48 hrs, syringe filtered (.45µm Acrodisc), stored at -80°C, and analyzed for p24 antigen by ELISA using the manufacturer’s instructions (Coulter, Miami, FL).
Virus stocks were titered in peripheral blood mononuclear cells (PBMC) that were stimulated with phytohemagglutinin (PHA) and cultured in RPMI 1640 (Gibco-BRL) containing 15% FCS, Interleukin-2 (Roche, Boheinger-Mannhiem) (30 U per ml), Penicillian (100U/ml), and Streptomycin (100 µg/ml). TCID$_{50}$ values were calculated by the Spearman-Karber method (57).

**Isolation of Target PBMC and Macrophages for Infections**

Leukopak blood from normal (uninfected) donors was processed for both peripheral blood mononuclear cells (PBMC) and monocyte derived macrophages (MDM). PBMCs were isolated from 15 ml of blood by centrifugation over a Ficoll gradient (described in chapter 3). Cells were suspended in PBMC media (RPMI 1640 + L-glutamine, 15% FBS, Penicillin (100U/ml), Streptomycin (100 µg/ml), and IL-2 (30 U/ml). PBMC were stimulated with 5 µg/ml phytohemagglutinin (PHA) (Sigma) for 15 minutes (incubation 37°C, 5% CO$_2$). Following the incubation, PBMC media was added to dilute the PHA to a final concentration of 2µg/ml and the cells were incubated for ~18 hours at 37°C, 5% CO$_2$. On day 1 the cells were washed to remove the PHA by transferring them into a 50 ml conical tube and centrifugation at 1000 RPM for 10 minutes. PBMCs were resuspended in PBMC media and split between the existing flask and a new flask. In a typical experiment 60 x 10$^6$ cells were stimulated, therefore each flask would contain 30 x 10$^6$ in 30 ml (1 x 10$^6$ cells/ml). Cells were incubated at 37°C, 5% CO$_2$ for an additional 2-3 days before infection.

Macrophage derived monocytes (MDM) were isolated from the remaining blood from a leukopak, using the RosetteSep Antbody Cocktail (Stemcell Technologies Inc, Vancouver, British Columbia) following the laboratory protocol. MDMs were plated at 1
x $10^6$ cell/ml in a 48 well plate with macrophage media (RPMI 1640 + L- Glutamine, 15% human AB serum (Sigma), penicillin (100U/ml), streptomycin (100 µg/ml), HEPES (2.5mM)) supplemented with GM-CSF (1 ng/ml). MDMs were incubated at 37°C, 5% CO₂ for ~5 days prior to infection.

**Virus Replication Assays**

Kinetics of replication were evaluated in Jurkat T cells (Weiss et al., 1984) and in PBMC. Jurkat cells were infected with 1000 TCID₅₀ of DNAse treated virus stock per 1.2 x $10^6$ cells. Infections were carried out in 50ml conical tube at 37°C, 5% CO₂, with gentle agitation once every 30 minutes for 2 hours. Cells were washed by adding 10 ml of PBS and pelleted by centrifugation at 1000 RPM for 10 minutes. Cells resuspended in RPMI 1640 with 10% FBS (Gibco-BRL). Supernatants from 3 wells were pooled at 2 day increments and stored at −80°C. Supernatant p24 antigen levels were determined by ELISA (Coulter) following the manufacturer’s protocol. PHA stimulated peripheral blood mononuclear cells were infected as described above for Jurkat T cells. PBMC cultures were maintained in PBMC media.

**Infection of MDM**

On the day of infection, MDMs were checked for macrophage morphology and confluency. Cells were washed twice with 1X DPBS (Gibco) prior to the addition of 1500 TCID₅₀ of virus per well. The final volume during infection was brought up to 100µl with fresh macrophage media and incubated for 2 hours at 37°C, 5% CO₂. Wells were washed 2 times with 1 ml DPBS and 1 ml of macrophage media was added. Cultures were returned to the incubator and supernatants from 3 replicates were pooled at 2 day intervals.
Gag-PR Processing in Jurkat Cells

Jurkat E6-1 (255) cells (30 x 10^6) were transfected with 20 µg of recombinant gag-pol virus plasmid DNA using SuperFect™(QIAGEN, Valencia, CA), following the manufacture’s protocol. The cells were washed 2 hours post infection by addition of 1X PBS (phosphate buffered saline) followed by centrifugation. Cells were resuspended in 5 ml fresh media (10% FCS, penicillin (100U/ml), streptomycin (100 µg/ml), in RPMI 1640 containing L-glutamine) for a final concentration of 2.5 x 10^5 cells per ml and incubated in T-25 flasks. One flask was collected at 0, 48, 60, 72, 80, and 96 hours for lysate preparation. Cells were pelleted, washed twice with PBS and lysed with M-Per™Mammalian Protein Extraction Reagent (PIERCÉ, Rockford, IL) in the presence of a protease inhibitor cocktail (Sigma, St. Louis, Missouri) for 10 minutes at room temperature on a shaker. Samples were centrifuged at 27,000 x g at 10°C for 15 minutes to remove cellular debris. The supernatant was stored at -80°C until further use.

Total protein was determined for each sample using BCA-200 protein assay kit using the manufacture’s protocol (PIERCÉ). Samples were boiled with LSB (Laemmli sample buffer, 100 mM Tris-Cl [pH 6.8], 200 mM dithiothreitol, 4% SDS, 0.2% bromophenol blue, and 20% glycerol) for 10 minutes and 25µg protein was loaded per sample on a 10% Tris-tricine SDS-polyacrylamide gel electrophoresis (171) for approximately 1 hour at constant voltage.

Processing of Gag-PR polyprotein over time was visualized by immunoblotting. Gels were blotted onto Immobilon™-P membranes (Millipore Corporation, Bedford, CT) in 10 mM MES (pH 8.3) buffer for 10-12 hours at 25 milliamps at 10°C. Membranes were blocked for 1 hour in the presence of 5% non-fat dried milk in 0.1% Tween 20 in
Tris-buffered saline (TTBS) containing 25 mM Tris (pH 7.4). A primary mouse monoclonal antibody to p24CA was obtained from the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH (Steimer et al., 1986). Incubation of membranes with a 1:10,000 dilution of the primary p24 antibody was performed in TTBS buffer containing 5% non-fat dried milk. Binding of the primary antibody was detected using a secondary goat anti-mouse antibody conjugated to horseradish peroxidase (Biorad) and visualized by enhanced chemiluminescence (Amersham Life Science, Buckinghamshire, England) using the standard protocol supplied by the manufacturer.

Results

Processing Phenotype is Reflected by Virus Replication

To assess for a functional relationship between Gag-PR processing and viral replication, three recombinant viruses were constructed and evaluated for initial replication compared with wild type HIVLAI in Jurkat T-cells or PBMC (Figure 5-5). Initial constructs were prepared using pretherapy protease alleles from three highly related viruses (a mother MD.03, and her two children D1.10 and D2.21).

Processing phenotypes determined in the Dunn lab were discussed in Chapter 1 of this dissertation and have been published (77). Recombinant viruses expressing a Type I processing Gag-PR region from D1.10 produced three-fold greater levels of p24 antigen than HIVLAI, a Type II processing phenotype, at four days of culture (Figure 5-5 A). Recombinant viruses with D2.21 gag-pro region, which had a Type II processing phenotype, replicated over the course of five days to levels that were similar to HIVLAI (Figure 5-5 B). Viruses with the gag-pro region from patient MD.03, which displayed slow, incomplete processing (Type III phenotype) were replication competent, but produced supernatant levels of p24 antigen that were 2- to 2.5-fold reduced from the
other viruses tested (Figure 5-5 B). Replication by recombinant viruses in PBMC produced levels of p24 antigen that were about ten-fold greater than in Jurkat cells (Figure 5-5 C). Yet, the relationship among the viruses with variant gag-pro regions for replication levels in PBMC were similar to results from Jurkat cells. By four to six days of infection in PBMC, D1.10 recombinant virus produced levels of p24 antigen that were two- to three-fold higher than levels produced by HIV\textsubscript{LAI} or D2.21, and five-to six-fold greater than levels produced by MD.03 (Figure 5-5 C). These results indicate that processing kinetics are a factor in virus replication; the more rapidly a virion matures based on Gag-PR processing, the more fit the virus and the more rapidly the virus can spread.

**Gag Processing in Jurkat Cells**

A pilot study identified that transfection, rather than infection, of Jurkat cells would provide a faster and more reliable method to acquire cell lysates for analysis of processing intermediates. Gag processing in Jurkat cells by recombinant viruses with protease variants was evaluated over time. Initial protease processing in Jurkat cells was undetectable during the first 24 hours with the p24 monoclonal antibody (data not shown). The Gag products detected by 48 hours included Pr55, p41, and p24\textsuperscript{CA} (Figure 5-6). Each virus produced similar steady state levels of the three Gag proteins for as long as four days (Figure 5-6).

**Pre and Post Therapy Alleles from a Discordant Patient**

An allele isolated from discordant response patient D1 prior to the initiation of combination therapy including a protease inhibitor (pre sample) had limited polymorphisms in Gag and Protease (Figure 5-7). Prior to therapy only L63P, which has been associated with reduced sensitivity to protease inhibitors, was detected. Gag-PR
regions isolated from patient D1 following therapy (samples post1 and post2) displayed a
number of therapy specific mutations in protease as well as polymorphisms in gag. The
pretherapy PR allele differed from both post therapy alleles, which accumulated a
number of mutations associated with decreased protease inhibitor sensitivity. The
following therapy associated mutations were detected in post1: L10I, M36I, I54V, L63P,
and V82A; and in post 2: L10I, I54V, L63P, V82A and L90M (Figure 5-7). Supplemental mutations I15V, E34Q, and Q58E accumulated in both post therapy isolates. An additional supplemental mutation was detected in Post2 (Figure 5-7).

Gag sequences also differed at a limited number of positions. The B and D
cleavage sites were identical in all three sequences. Pre Gag-PR differed from both post
sequences in the D’ site at position P2’ (residue 450) while in p7 only post2 displayed a
mutation at position 398. Most polymorphisms were concentrated in the C cleavage site
at positions 374, 375, 376, and 380. None of these mutations have been associated with
increased viral fitness. To evaluate how these subtle polymorphisms in gag and protease
impact viral fitness replication competent recombinant viruses were constructed using
each of these alleles.

**Replication of Recombinant Virus in PBMC**

Recombinant viruses with the HIV$_{LAI}$ X4 envelope were used to infect PBMC. This infection predominantly targets CD4$^+$ CD45RA cells since they express the CXCR4
coreceptor. Between 2 and 5 independent infections were performed using each
recombinant virus. The range of p24 produced varied dramatically depending on donor
of target cells used for infection. One way to standardize the results was to evaluate p24
production of each virus relative to HIV$_{LAI}$, the wildtype control virus used in all
infections targeting X4 PBMC (Figure 5-8 A). The p24 produced relative to LAI varied
dramatically from 0 to 9 fold for the pre recombinant virus, and to a lower extent in the post viruses. Regardless of the relative amount of p24 the rank order based on peak p24 production was consistently pre, LAI, post 2 and post 1 viruses.

The growth curves of X4 PBMC infections were evaluated and displayed as the median log_{10} pg/ml in Figure 5-9. After six days of infection post1 recombinant demonstrated a dramatic 74-fold reduction in supernatant p24 relative to the pre recombinant virus. Post2 virus demonstrated a modest 3-fold reduction in p24 at day 6. This data suggests that polymorphisms in post2 compensate for reduced fitness of post1 in X4 PBMC. Results from a series of real time PCR experiments demonstrated that differences in replication are not the result of differences in the inoculums.

**Replication of R5 Virus in PBMC**

Recombinant viruses with the HIV_{JRFL} R5 envelope were used to infect the CD4^+ R5, memory component, of PBMC in 4 to 6 independent donors. Although a range of p24 relative to LAI was detected, it was less dramatic than what was seen in X4 PBMC (Figure 5-8B). Growth curves from infection by R5 viruses were distinctly different from infections of X4 PBMC. The magnitude of p24 produced was lower, most likely resulting from the reduced number of R5 target cells in stimulated PBMC. Both post1 and post2 viruses displayed modest reductions in replicative ability relative to the pre virus (Figure 5-10).

**Replication of Recombinant Virus in Macrophages**

Recombinant gag-pro viruses with the HIV_{JRFL} R5 envelope were used to infect MDM (monocyte derived macrophages) from 4 independent donors. Supernatant p24 levels from the pre virus ranged from 1 to 4 fold relative to LAI, reflecting a donor cell impact (Figure 5-8C). Median p24 production of post1 was virtually identical to the pre
virus 8 days post infection (Figure 5-11). In contrast, post2 virus, which replicated better in X4 PBMC, was diminished in MDM. Mutations that restore viral fitness in PBMC appear to have a detrimental effect in the context of macrophages.

**Discussion**

A small number of natural amino acid polymorphisms in HIV-1 Gag can have a dominant impact on PR activity (77). This study determined that natural polymorphisms in gag-pro impacted growth by recombinant viruses and that rapid Gag-PR processing in bacteria is concordant with accelerated production of recombinant viruses in both T cell lines and in PBMC. This panel of experiments demonstrated the natural variability that develops in HIV-1 Gag in patients without drug selection points to selective pressures that improve PR processing function. Cleavage site polymorphisms impacted growth by recombinant viruses.

Pretherapy Gag-PR from a discordant response patient (D1) processed rapidly to p24Ca in a bacterial expression system (77), while processing by two post therapy Gag-Pr alleles displayed reduced ability to process Gag completely to p24 (163). Gag-PR sequences tested in the bacterial expression system were introduced into recombinant viruses for evaluation of replication kinetics. Replication of post1 gag-pro recombinant virus in X4 PBMC demonstrated up to a 74-fold reduction in supernatant p24 relative to the pre recombinant virus at 6 days post infection (Table 5-2). Post2 gag-pro virus has a modest 3-fold reduction in replicative ability relative to the pre virus, suggesting that polymorphisms between post1 and post2 alleles were responsible for the apparent restoration of fitness in X4 PBMC. The post alleles differed at amino acid positions 373, 374, 375, 376, and 380 in Gag. Residues 36, 60 and 90 varied in PR.
Replication by post gag-pro recombinant viruses in R5 PBMC was virtually identical to the pre recombinant virus. Post1 and post2 viruses displayed a 1.8- and 2.4-fold reduction, respectively in supernatant p24 relative to the pretherapy recombinant virus. This modest reduction in replicative ability should be considered with caution, since this fold difference is within the margin of error for viral titering.

An unexpected result was that protease resistance in post1 virus had no impact on replication when targeted to macrophages. The ability of protease resistant virus to replicate in macrophages to levels equal to protease sensitive viruses suggests that there are cell type specific influences on the virus, either at the level of viral RNA production and viral proteins, or at the level of protease processing. The fact that a therapy resistant virus replicates to levels equal to the pre therapy allele suggests that this cell could be a viral reservoir.

Host genetic factors, in combination with viral properties, determine the susceptibility of an appropriate target cell for HIV-1 infection as well as the replication potential of the virus in the cell. The in vivo consequences of host genetic factors are obvious, yet a discernable impact is imposed by human donor target cells, during in vitro experiments. One difficulty with these experiments was target cell variability. While the magnitude of difference between viruses within various target cells varied, the rank order of viruses with respect to peak p24 production was constant. A large number of infections (5-7 donors) were required to discern the true impact of gag-pro versus host cell effects. In vitro studies with T cell adapted HIV isolates demonstrated that PBMCs from different donors have varied susceptibility with differences up to 1000 fold (247), while other groups found a maximum of only 40 fold variation in HIV-1 production in
PBMC with low passage primary isolates (222). HIV-1 replication kinetics in monocytes/macrophages derived from identical twins compared to age matched donors were discordant, suggesting a host cell genetic effect on productive viral replication (27,158). Host cell restrictions have been associated with RT processes (66). In our system RT is constant in all constructs, indicating that the host cell restrictions act at other steps in the life cycle. Interpretation of future experiments would be simplified if a single source (donor) of cells could be identified.

A standard fitness assay has not been adopted across laboratories, making comparisons between studies difficult. One current strategy uses a single cycle recombinant virus containing patient derived reverse transcriptase and protease (187), however, this strategy fails to address the role of gag mutations in fitness. The advantages of our recombinant replication competent virus system are the inclusion of gag and the ability to measure both the kinetics of supernatant p24 production and the spread of cell-associated virus during several rounds of infection in different cell types. Our strategy focused specifically on Gag-PR from the C terminus of p24CA thru PR. Standard strategies of measuring fitness will aid in the identification and relative impact of mutations.

The mechanism(s) to explain the different replicative ability in macrophages and T cells is elusive. Some possibilities at the virion level include modifications of protein folding, proteolytic processing, and RNA dimer stability all of which could lead to reduced production of new infectious virions. At the cellular level, there are a number of unexplored possibilities. A study in the Goodenow lab demonstrated viral induced changes in cellular genes during macrophage infection (33). One possibility is that viral
regulation within cell types may vary and could play a role in replication. Protease may interact with cellular proteins altering replicative ability. *In vivo* there could be an impact at the cellular level by protease inhibitors on the cell cycle.

Future studies need to map the specific determinants responsible for fitness in macrophages and PBMC. One approach would be to introduce the *gag-pro* polymorphisms identified in the post2 virus individually and in combination into the post1 virus and assess the impact on fitness relative to the pre *gag-pro* recombinant virus. Competition experiments in the presence and absence of drug would be the most accurate assessment of viral fitness. While the identified Gag-PR polymorphisms may have a direct impact on viral fitness, it is likely that mutations at a number of other positions may yield similar results.
Figure 5-1. Strategy 1 for the production of gag-pro recombinant viruses. Top left, restriction sites in shuttle vector pGEM-3z. Top right, pLAI.2, restriction map identifying the SpeI site targeted for ablation as well as the BamHI and SphI sites used for generation of pGEM-3z-4.6LAI.
Figure 5-2. Linker Insertion Mutagenesis. Annealed oligos were ligated into the SpeI digested pGEM-3Z-4.6LAI, deleting the SpeI site at position 7 and inserting an ApaI restriction site.
Figure 5-3. Map of recombinant viruses. A, location of the 1.5kb gag-pro region inserted using strategy one (purple box); B, strategy two, insertion of an 830bp gag-pol fragment (aqua line); C, the 1.3kb v1-v5 fragment HIV_{JRFL} inserted into pLAI.4 (yellow box).
Site directed mutagenesis adding XbaI & KpnI sites
Removal of pLAI.4 830bp fragment and insertion of 830 bp patient fragments

Digestion with SpeI and BstZ17I yields a 1.5 kb fragment for ligation into pLAI.4.

Figure 5-4. Strategy 2 for the production of gag-pro recombinant virus. A 1.56kb region of pLAI.2 (SphI to EcoRV) was inserted into pGEM5zf. Site directed mutagenesis was used to insert XbaI and KpnI restriction sites at positions 1861 and 2652 (pLAI.2), respectively. The XbaI and KpnI sites are used to insert the 830bp fragment into the shuttle vector, from which the 1.5kb fragment can be isolated and ligated in pLAI.4.
Table 5-1 Protease recombinant viruses

<table>
<thead>
<tr>
<th>gag-pol region</th>
<th>Envelope</th>
<th>therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAI.4</td>
<td>X4</td>
<td>wildtype</td>
</tr>
<tr>
<td>LAI.4</td>
<td>R5</td>
<td>wildtype</td>
</tr>
<tr>
<td>D1.10</td>
<td>X4</td>
<td>pre</td>
</tr>
<tr>
<td>D1.10</td>
<td>R5</td>
<td>pre</td>
</tr>
<tr>
<td>D1 (3188A01)</td>
<td>X4</td>
<td>post</td>
</tr>
<tr>
<td>D1 (3188A01)</td>
<td>R5</td>
<td>post</td>
</tr>
<tr>
<td>D1 (3188O05)</td>
<td>X4</td>
<td>post</td>
</tr>
<tr>
<td>D1 (3188O05)</td>
<td>R5</td>
<td>post</td>
</tr>
<tr>
<td>D1.21</td>
<td>X4</td>
<td>pre</td>
</tr>
<tr>
<td>D1.21</td>
<td>R5</td>
<td>pre</td>
</tr>
<tr>
<td>D2.21</td>
<td>X4</td>
<td>pre</td>
</tr>
<tr>
<td>D2.21</td>
<td>R5</td>
<td>pre</td>
</tr>
<tr>
<td>MD.03</td>
<td>X4</td>
<td>pre</td>
</tr>
<tr>
<td>MD.03</td>
<td>R5</td>
<td>pre</td>
</tr>
<tr>
<td>LAI.4 1628U02</td>
<td>X4</td>
<td>post</td>
</tr>
<tr>
<td>LAI.4 1628U02</td>
<td>R5</td>
<td>post</td>
</tr>
<tr>
<td>2590U01</td>
<td>X4</td>
<td>post</td>
</tr>
<tr>
<td>2590U01</td>
<td>R5</td>
<td>post</td>
</tr>
<tr>
<td>1877U04</td>
<td>X4</td>
<td>pre</td>
</tr>
<tr>
<td>1877U04</td>
<td>R5</td>
<td>pre</td>
</tr>
<tr>
<td>2308U05</td>
<td>X4</td>
<td>post</td>
</tr>
<tr>
<td>2308U05</td>
<td>R5</td>
<td>post</td>
</tr>
<tr>
<td>2309A02</td>
<td>X4</td>
<td>post</td>
</tr>
<tr>
<td>2309A02</td>
<td>R5</td>
<td>post</td>
</tr>
</tbody>
</table>
Figure 5-5. Replication kinetics by recombinant viruses containing natural gag-pro variants. Equal TCID₅₀ of each recombinant virus stock was used to infect cells. A and B, replication in Jurkat T cell line; C, primary PBMC. Accumulation of p24 antigen levels in supernatants (ordinate) is indicated for days post-infection (abscissa). Data are representative experiments that were repeated three times. Gag-pro recombinant viruses are designated by symbols: (■), D1.10; (●) D2.21; (Δ) MD.03; (X) HIV₅₅₆₅. 
Figure 5-6. Immunoblot analysis of Gag-Pol processing in Jurkat cells by recombinant viruses. (77)
**Gag**

<table>
<thead>
<tr>
<th>370</th>
<th>380</th>
<th>400</th>
<th>420</th>
<th>431</th>
<th>453</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>gagLAI</td>
<td>KARVLAEMSQVTNSATIMQRGFRNQRKIVKCFNCGKEGHIARNCRAPRKGCWKCNGKEGHQMKDTERQANFLGKIWFSYKGRPNFQLSRP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre</td>
<td>-----------</td>
<td>A---QI---K</td>
<td>----------------------</td>
<td>K</td>
<td>--------------</td>
</tr>
<tr>
<td>post 1</td>
<td>A---QIV---K</td>
<td>----------------</td>
<td>V</td>
<td>---K</td>
<td>----------------</td>
</tr>
<tr>
<td>post 2</td>
<td>A---APIL</td>
<td>----------------</td>
<td>K</td>
<td>--------------</td>
<td>-----------------------------</td>
</tr>
</tbody>
</table>

**Protease**

<table>
<thead>
<tr>
<th>133</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
</tr>
<tr>
<td>LAI</td>
</tr>
<tr>
<td>pre</td>
</tr>
<tr>
<td>post 1</td>
</tr>
<tr>
<td>post 2</td>
</tr>
</tbody>
</table>

Figure 5-7. Amino acid sequence of Gag and PR alleles isolated from patient D1 prior to and following therapy that were used to construct recombinant viruses. Highlighted positions in protease are mutations associated with reduced sensitivity to protease inhibitors.
Figure 5-8  Supernatant p24 relative to LAI following infection by recombinant viruses.  
A, X4 PBMC 6 days; B, R5 PBMC 6 days; and C, MDM infection 8 days post infection. 
by pre- and post therapy alleles from a discordant response patient. Bars represent the 
median p24 relative to HIV\textsubscript{LAI}.
Figure 5-9. Median supernatant p24 following infection of X4 PBMC by *gag-pro* recombinant viruses. LAI (♦); pre (■); post1 (▲); post2 (X).
Figure 5-10. Median supernatant p24 following infection of R5 PBMC by *gag-pro* recombinant viruses. LAI (♦); Pre (■); post1 (▲); post2 (X).
Figure 5-11. Median supernatant p24 following infection of macrophages by R5 *gag-pro* recombinant viruses. LA1 (♦); pre (■); post1 (▲); post2 (X).
Table 5-2. Fold Difference in Supernatant p24 Production Relative to Pretherapy

<table>
<thead>
<tr>
<th></th>
<th>Post 1</th>
<th>Post 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>X4 PBMC</td>
<td>↓ 74X</td>
<td>↓ 3X</td>
</tr>
<tr>
<td>R5 PBMC</td>
<td>↓ 1.8X</td>
<td>↓ 2.4X</td>
</tr>
<tr>
<td>R5 MDM</td>
<td>↑ 1.3X</td>
<td>↓ 2.5X</td>
</tr>
</tbody>
</table>
CHAPTER 6
DISCUSSION

Protease Genotypic Resistance as a Novel Marker

Studies in retroviruses have yielded findings that have had a broad impact on biology and medicine. For example, the role of oncogenes in cancer, as well as many molecular mechanisms of cell growth and differentiation were identified thru the study of retroviruses (238). Subsequently mutations in HIV-1 resulting from therapy may provide further insight to normal cellular processes.

Results presented in this dissertation established that genotypic resistance to antiretroviral therapy can be a valuable probe for human cells. Specifically, genotypic resistance in protease offers the best genotypic marker when compared to reverse transcriptase [RT] or envelope. In general, genotypic resistance to reverse transcriptase inhibitors can occur rapidly after therapy induction requiring only few mutations. As a result, RT genotypic resistance is a poor marker for this type of analysis due to the lack of intermediates. Envelope genotype also fails to serve as a marker because of the limited ability to detect intermediates. In contrast, protease resistance is associated with the accumulation of many different mutations that frequently occur in a stepwise pattern, and produce detectable intermediates, which persist while under the selective pressure imposed by therapy. While this study used protease genotypic resistance as a marker for viral trafficking it may be useful as a lineage marker, to identify cell subsets for future studies.
Response to Protease Inhibitor Therapy, Implications of Viral Genetics

At the initiation of this study a discordant response to combination therapy including a protease inhibitor was a novel, unexpected result that was only beginning to gain recognition in the field. A discordant response (VFIS), which occurs in up to 40% of treated patients, is characterized by high viral load, like the VFIF group yet with an increase in CD4+ T cells resembling the VSIS group. Despite high viral load, discordant patients do not show evidence of disease progression. A number of studies have demonstrated that a discordant response may persist for 3 or more years (48,101). There is evidence suggesting that an increase in CD4+ T cells is accompanied by functional reconstitution in VSIS and, to a lesser extent, in VFIS patients (118,176). The level of infected CD4+ T cells following 56 weeks of therapy remained high in VFIF patients (2.4%), while only 0.3% and 0.02% of VSIS and VFIS CD4+ T cells were infected, respectively (169). The small fraction of infected CD4+ T cells, despite of high viral load, suggests that an altered tropism or attenuation of the virus has occured. Following therapy, viruses that rebound in VFIS patients do not exhibit changes in envelope that would alter or extend tropism (232).

The importance of evaluating viral parameters associated with discordant response became the focus of my research. The work in this dissertation contributes knowledge about the polymorphisms in gag-pro relative to therapy outcome and evaluation of the role of gag-pro in the reduced pathogenesis of rebounding virus in discordant response patients. Pre- and post therapy gag-pro genotypes were evaluated from each response group in a cohort of pediatric patients, to identify genotypic markers. A spectrum of naturally occurring amino acid polymorphisms in PR can be used as biomarkers to predict both immune and viral outcomes to initial combination therapy in
HIV-1 infected patients. Following 24 weeks of treatment, the number of mutations associated with reduced inhibitor sensitivity were indistinguishable between viral failures, regardless of immune response. In contrast, patients exhibiting viral failure had significantly more supplemental mutations in PR and C cleavage site polymorphisms. These mutations may contribute to viral fitness in drug resistant variants. Mechanisms by which Gag mutations may restore viral fitness are still under investigation. Mutations in the C cleavage site may increase proteolytic processing (163), while mutations in Gag may impact other essential steps in virion assembly such as RNA dimerization (211) or cyclophilin A incorporation (76). Identifying positions in Gag and PR associated with the discordant response to therapy was the first step in evaluating whether the apparent pathogeneic reduction in virus from VFIS patients mapped to these genes.

Pre- and post therapy Gag-PR regions were tested for their impact on the replication abilities using a replication competent recombinant virus system. A dramatic reduction in viral fitness was observed when X4 PBMC were infected by resistant post therapy recombinant viruses. However, pre- and post therapy primary isolates were found to have identical growth kinetics when used to infect PBMC. While mutations in Gag and PR contribute to viral fitness they alone appear to be insufficient to explain the discordant response. Thus, a complex interplay between several regions in the genome must be responsible for the altered pathogenesis in therapy altered viruses.

The low level of infected CD4\(^+\) T cells (224), in combination with the reduced replicative ability by protease resistant virus, suggests that rebounding virus in discordant patients is not replicating in CD4\(^+\) X4 PBMC. Although typically only a small percentage of tissue macrophages are infected (140,214) the level may rise in patients
during end stage disease with exposure to opportunistic infections (167). Patients included in our cohort displayed disease progression, increasing the possibility of macrophage infection. Viruses associated with monocytes, the precursors of macrophages, were identical to the actively replicating pool of viruses found in the plasma (255). These monocytes, once differentiated, would be infected with recently evolved drug resistant variants, which could produce the next round of plasma virus. A limited number of infected cells would be required to produce a high level of plasma virus. The ability of protease resistant virus to replicate in macrophages to levels equal to protease sensitive viruses suggests that there are cell type specific influences on the virus, either at the level of viral RNA production and viral proteins, or at the level of protease processing. Unaltered replication of therapy resistant viruses in macrophages suggests that this cell lineage could be a viral reservoir. These results have significant implications for the discordant response. The reduced replication in thymocytes (224) and lymphocytes can account for reconstitution of CD4+ T cells, while persistent replication in macrophages can contribute to the high levels of plasma virus, detected in patients with a discordant response.

If virus replication in macrophages persists with therapy, while replication in lymphocytes is impaired, then differential impact by drugs on cell lineages might provide one explanation for a VFIS response. One possibility for reduced pathogenesis is the viral protease interacts with other viral proteins, such as regulatory proteins which in macrophages may not be as critical as they are in lymphocytes. A significant question remaining is how virus replication can be restricted in selected host cells, while continuing to produce high viral levels.
Virus that rebounds in VFIS patients is not as infectious. The mechanism(s) responsible for reduced pathogenesis remain undefined, although host pharmacogenetics and viral genetics may be contributors. Future studies to identify target cell gene regulation in the presence or absence of drug will address the possibility that drugs may alter the host cells. Based on knowledge of envelope genotype, viruses retain the ability to enter cells in VFIS patients, however, the proviral load following therapy indicates that virus is not integrating into CD4+ T cells, suggesting a post entry block. Additionally the relative infectiousness of viral particles has yet to be established, in other words the number of mature infectious virions relative to all that are produced. Post therapy Gag-Pol variants showed diminished processing abilities in a bacterial expression system (163). Virions in the plasma have an extremely short half-life estimated to range from minutes to hours (182,253). Since mutations associated with drug resistance decrease processing it is possible that some virions are only partially processed prior to the initiation of degradation, making these virions uninfected, which may result in high viral burden associated with low levels of infection. Virions associated with follicular dendritic cells remain infectious for at least 25 days (220) in human cells. The extended half-life of virus on the surface of dendritic cells may allow for complete processing to occur and the virion to mature prior to contact with target cells.

The reduced pathogenesis in HIV-1 variants following therapy may parallel the apparent reduced pathogenesis or attenuation of SIV (simian immunodeficiency virus) in African green monkeys and sooty mangabey monkeys. These animals can be infected for life, yet, fail to show clinical symptoms or develop disease. However, if these viruses are placed in another host pathogenesis is observed(70). This model suggests that host
factors are intricately involved. The fact that SIV maintains its reduced pathogenesis in the same species but not in another suggests a different mechanism from HIV-1; the VFIS response occurs in the same host suggesting the virus must be changing.

Most successful vaccine campaigns have used live attenuated virus. This option for HIV-1 can only be considered after the factors responsible for reduced pathogenesis are further defined. In Africa where the risk of HIV-1 infection is extremely high, such a vaccine may be reasonable. After all, if the odds are great for infection anyway, a variant with low pathogenic potential might be more favorable. Of course, one serious risk remains the possibility that the virus could revert back to its original pathogenic level.

Although multiple factors, including clinical variables and PR genotype (184), influence response to PI therapy, determinants in Gag that regulate processing and virus replication could be additional factors in success of antiretroviral protocols. The work in this dissertation contributes knowledge about the polymorphisms in gag and pro relative to therapy outcome and the effect of these mutations on the virus. Understanding the contribution of viral genetics to therapy response may allow for more informed decisions about therapy and lead to better designs of specific inhibitors to HIV-1 used for the battle against infection by HIV-1.
VSIS03

Pretherapy

1794u03  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u04  ------------------------------------N-T-K------------------E--P------------------I-----------------
1794u05  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u06  ------------------------------------N-T-K------------------E--P------------------I-----------------
1794u07  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u08  ------------------------------------N-T-K------------------E--P-------------I----I-----------------
1794u09  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u10  ------------------------------------N-T-K------------------E--P-------------I----I-----------------
1794u11  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u12  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u13  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u14  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u15  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u16  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u17  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u18  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u19  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
1794u20  ------------------------------------N-T-K------------------E--PV-----------------I-----------------

Posttherapy

2000pr  ------------------------------------N-T-K------------------E--PV-----------------I-----------------
2000r01  ------------------------------------T-T-K---------------------P-----R------------------
2000r02  ------------------------------------T-T-K---------------------P-----R------------------
2000r03  ------------------------------------T-T-K---------------------P-----R------------------
2000r04  ------------------------------------T-T-K---------------------P-----R------------------
2000r05  ------------------------------------T-T-K---------------------P-----R------------------
2000r06  ------------------------------------T-T-K---------------------P-----R------------------
2000r07  ------------------------------------T-T-K---------------------P-----R------------------
2000r08  ------------------------------------T-T-K---------------------P-----R------------------
LAI PR PQTLWQRPLVTIKGGQLKEALLDTADTLEEMSLPGWRKPMIGGGFIFKVRQYDLIEICGHRAIGTVLVGPTVPNIIICRNLTQIGCTLN

VSIS03 post therapy continued.

2590003 --I------------------------------T-T-K------------------------P-----R------------L----D-----
2590004 --I-G--F----------------------V-------------------------------N-T-K------------------------E--P----------------------------------------------------------
2590005 -------------------------------T-T-K------------------------E--P--------------D--------------------------
2590006 -------------------------------N-T-K------------------------E---P---------------D--------------------------
2590007 -------------------------------N-T-K------------------------E--P--------------D--------------------------

VSIS04

Pretherapy
8395110                 F-L----H------------------T------------------------RV------------I---------T------------
8395r11                   -F-V----A-----------------IN------------------------VP------------------------------------
8395r12                   -F-V----A-----------------IN------------------------VP------------------------------------
8395r13                   -F-V----H------------------T------------------------M-V------------I----------------------
8395r14                   -F-V----H------------------T------------------------M-V------------I----------------------
8395r16                     A-S-F-V----H------------------T-------------------------RV------------I----------------------
8395r17                   -F------H------------------T------------------------VP------------------------------------
2141pr                     --- -------------------------------IN------------------------VP------------------------------------

Posttherapy
1779u01 ---------------IN------------------------VP-------------------------------R------------------------
1779u02 ---------------IN------------------------VP-------------------------------R------------------------
1779u03 ---------------IN------------------------VP-------------------------------K------------------------
1779u05 ---------------IN------------------------VP-------------------------------K------------------------

VSIS16

Pretherapy
2546pr ---------------------------------N-------------------------------P-------------------------------

VSIS17

Pretherapy
1875pr --- ---------------D-----R------------------------R----------------------I-------------------------------
VSIS20
Pretherapy

LAI PR    PQITLWQRPLVTIKIGGQLKEALLDTGADTVLEEMSLPGRWKPKMIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNF

2784a05 -----------------------------------N---------------------------------------------------------------
2784a06 -----------------------------------N---------------------------------------------------------------
2784a02 -----------------------------------N---------------------------------------------------------------
2784o03 -----------------------------------N---------------------------------------------------------------
2784o04 -----------------------------------N---------------------------------------------------------------
2784o05 -----------------------------------N---------------------------------------------------------------
2784o06 -----------------------------------N---------------------------------------------------------------
2784ca01 -----------------------------------N---------------------------------------------------------------
2784ca02 -----------------------------------N---------------------------------------------------------------
2784ca05 -----------------------------------N---------------------------------------------------------------
2784ca10 -----------------------------------N---------------------------------------------------------------
2784co04 -----------------------------------N---------------------------------------------------------------
2784co05 -----------------------------------N---------------------------------------------------------------
2784co06 -----------------------------------N---------------------------------------------------------------
2784co07 -----------------------------------N---------------------------------------------------------------
2784co08 -----------------------------------N---------------------------------------------------------------
2784co09 -----------------------------------N---------------------------------------------------------------
2784co10 -----------------------------------N---------------------------------------------------------------
2784co11 -----------------------------------N---------------------------------------------------------------
2784co12 -----------------------------------N---------------------------------------------------------------
2784co16 -----------------------------------N---------------------------------------------------------------
2784co17 -----------------------------------G---------------------------------------------------------------
2784co18 -----------------------------------N---------------------------------------------------------------
2784co19 -----------------------------------G---------------------------------------------------------------
2784co20 -----------------------------------N---------------------------------------------------------------
VSIS21

Pretherapy

2712pr  ---------V-----------------N---K------------------------A------R------I----------------------
2712u01  ---------V-----------------N---K------------------------A------R------I----------------------
2712u02  ---------V-----------------N---K------------------------A------R------I----------------------
2712u03  ---------F----V-----------------A--N---K------------------------A------R------I----------------------
2712u04  ---------V-----------------N---K------------------------A------R------I----------------------
2712u05  ---------V-----------------N---K------------------------A------R------I----------------------
2712u06  ---------V-----------------N---K------------------------A------R------I----------------------
2712u07  ---------V-----------------N---K------------------------A------R------I----------------------
2712u08  ---------V-----------------N---K------------------------A------R------I----------------------
2712u09  ---------V-----------------N---K------------------------A------R------I----------------------
2712u10  ---------V-----------------N---K------------------------A------R------I----------------------
2712u11  ---------V-----------------N---K------------------------A------R------I----------------------
2712u12  ---------V-----------------N---K------------------------A------R------I----------------------
2712r01  ---------V-----------------N---K------------------------A------R------I----------------------
2712r02  ---------V-----------------N---K------------------------A------R------I----------------------
2712r03  ---------V-----------------N---K------------------------A------R--------------------------
2712r04  ---------V-----------------N---K------------------------A------R--------------------------
2712r05  ---------V-----------------N---K------------------------A------R--------------------------
2712r06  ---------V-----------------N---K------------------------A------R--------------------------

Posttherapy

3097c01  ---------V-----------------N---K------------------------A------R------I----------------------
3097c02  ---------V-----------------N---K------------------------A------R------I----------------------
3097c07  ---------V-----------------N---K------------------------A------R------I----------------------
3097c08  ---------V-----------------N---K------------------------A------R------I----------------------

VSIS25

Pretherapy

2473pr  ---------V-----------------N--------------------------

VFIS10  Posttherapy

1664u01  ------------------I-----------------N-------------------------C-----------------------------L-----
1664u02  ------------------I-----------------N-------------------------C-------------I---------------L-----
1664u03  ------------------I-----------------N-------------------------C-------------I---------------L-----
1664u04  ------------------I-----------------N-------------------------C-----------------------------L-----
1664u05  ------------------I-----------------N-------------------------C-----------------------------L-----
1664u06  ---------I--------I-----------------N-------------------------C-------------I----A-V--------L-----
1664u07  ------------------I-----------------N--------------S----------C-------------I---------------L-----
1664u08  ------------------I-----------------N---------T------V--------C-------V-----I----A-------M--L-----
1664u09  ------------------I-----------------N-------------------------C-----------------------------L-----
1664u10  ------------------I-----------------N-------------------------C-----------------------------L-----
1664o19  ---------I--------I-----------------N----------------V--------CV------V-----I-------------
1664r02  ---------I--------I-----N-E---------N----------------V--------C-------V-----I----A-------M
1664r03  ---------I--------I-----N-----------N----------------V--------C-------V-----I----A-------M
1664r07  ---------I--------I-----N-----------N----------------V--------C-------V-----I----A-------M
1664r09  ---------I--------I-----N-----------N----------------V--------C-------V-----I----A-------M
<table>
<thead>
<tr>
<th>LAI</th>
<th>PR</th>
<th>PQILTWRPLVTIKIGGQLKEALDTGADDTVLEEMSLPGWRKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLGPTPVNIIGRNLLTQIGCTLNF</th>
</tr>
</thead>
<tbody>
<tr>
<td>3241ca01</td>
<td>----I--V--A---R-----------IF--I--------T-----------------R--V--------P--------V--------S--------A------Q----</td>
<td></td>
</tr>
<tr>
<td>3241ca04</td>
<td>----I--V--A---R-----------IF--I--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241ca05</td>
<td>----I--V--A---R-----------IF--I--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241ca06</td>
<td>----I--V--A---R-----------IF--I--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241ca08</td>
<td>----R--I--V--A---R-----------IF--I--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241co01</td>
<td>----I--V--A---R-----------IF--I--------T-----------------V--------P--------V--------A------Q----</td>
<td></td>
</tr>
<tr>
<td>3241co02</td>
<td>----I--V--AR--R-----------IF--I--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241co03</td>
<td>----I--V--A---R-----------IF--I--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241co04</td>
<td>----I--V--A---R-----------IF--I--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241co05</td>
<td>----I--V--A---R-----------IF--I--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241co06</td>
<td>----I--V--A---R-----------IF--I--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241co07</td>
<td>----I--V--A---R-----------IF--IN--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241co08</td>
<td>----I--V--A---R-----------IF--IN--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241co09</td>
<td>----I--V--A---R-----------IF--I--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
<tr>
<td>3241co10</td>
<td>----I--V--A---R-----------IF--IN--------T-----------------V--------P--------V--------A----M-------</td>
<td></td>
</tr>
</tbody>
</table>
VFIS12

Pretherapy

1738pr  ~------------------V-------------------D-N-------------------------P---------------------------IR---V-
1738a02  ~------------------V-------------------D-N-------------------------P---------------------------N--
1738a03  ~------------------V-------------------D-N-------------------------P---------------------------N--
1738a04  ~------------------V-------------------D-N-------------------------P---------------------------N--
1738a05  ~------------------V-------------------D-N-------------------------P---------------------------I--
1738a06  ~------------------V-------------------D-N-------------------------P---------------------------I--
1738a07  ---I-----------------V------------D-N-------------------------P-K-----------I----------------------
1738o01  ~------------------V-------------------D-N-------------------------P---------------------------N--
1738o03  ~------------------V-------------------D-N-------------------------P---------------------------N--
1738o04  ---A----------V------V------------D-N-------------------------P---------------------------I--------

Posttherapy

1943pr  ~---L---I------------------------DIN----------------V---- ---P ------X----------A-X--------
1943o01  ~---Q- -I------------------------DIN----------------V--------P------------------A-----------------
1943r01  ~---------I------------------------DIN----------------V--------P------------------A-V-------------
1943r02  ~---------I---------R-----------A--DIN-----E----------V--------P-------V----------A-V------------
1943r03  ~---------I------------------------DIN----------------V--------P------------------A------M---------
1943r04  ~---------I------------------------DIN----------------V--------P------------------A-V-------------
1943r06  ~------------------V-------------------DIN----------------V--------P-------V----------A-V---LT
1943a01  ~------------------V--V-------------------D-N-------------------------P-------------------I--
1943a03  ~------------------V-------------------D-N-------------------------P-------------------K-------N-V----
1943a05  ~------------------V-------------------D-N-------------------------P-------------------I--K-------N-V--I--
1943a06  ~------------------V-------------------D-N-------------------------P-------------------N--
1943o11  ~------------------V-------------------DVN----------------V--------P-------T----------A-----------------
1943o12  ~------------------V-------------------DVN----------------V--------P------------------A-----------------
1943o13  ~------------------V-------------------DVN----------------V--------P------------------A-----------------
1943o14  ~------------------V-------------------DVN----------------V--------P------------------A-----------------
1943o15  ~------------------V-------------------DVN----------------V--------P------------------A-----------------
VFIS13
Pretherapy
1989pr  ---------I--------I-----------------N--------------------------V----------------------------L------
1989r07  ---------I--------I-----------------N--------------------------V-----------------------------L--
1989r08  ---------I--------I-----------------N--------------------------V-----------------------------L--
1989r09  ---------I--------I-----------------N--------------------------V-----------------------------L--
1989r10  ---------I--------I-----------------N--------------------------V-----------------------------L--
1989r11  ---------I--------I-----------------N--------------------------V-----------------------------L--

Posttherapy
2254a01  ---------I--------I-----------------N--------------------------V-----------------------------L--
2254a02  ---------I--------I-----------------N--------------------------V-----------------------------L--
2254a03  ---------I--------I-----------------N--------------------------V-----------------------------L--
2254a04  ---------I--------I-----------------N--------------------------V-----------------------------L--
2254a05  ---------I--------I-----------------N--------------------------V-----------------------------L--
2254a02  ---------I--------I-----------------N--------------------------V-----------------------------L--
VFIS14
Pretherapy
1329pr  ----------------------------------------K----------------------V-----------------------------------
1329r13 ----------------------------------------K----------------------V--------------------------
1329r23 ----------------------------------------K----------------------V--------------------------
1329r24 ----------------------------------------K----------------------V--------------------------
1329u01 ----------------------------------------K----------------------V-----------------------------------
1329u02 ----------------------------------------K----------------------V-----------------------------------
1329u03 ----------------------------------------K----------------------V-----------------------------------
1329u06 ----------------------------------------K----------------------V-----------------------------------
1329u07 ----------------------------------------K----------------------V-----------------------------------

Posttherapy
1576pr  ----------------------------------------K----------------------V-----------------------------------
1628u01 ----------------------------------------K----------------------V-----------------------------------
1628u02 ----------------------------------------K----------------------V-----------------------------------
1628u03 ----------------------------------------K----------------------V-----------------------------------
1628u04 ----------------------------------------K----------------------V-----------------------------------
1628u07 ----------------------------------------K----------------------V-----------------------------------
1628u13 ----------------------------------------K----------------------V-----------------------------------
1628u14 ----------------------------------------K----------------------V-----------------------------------
1628u17 ----------------------------------------K----------------------V-----------------------------------
1628r13 ----------------------------------------K----------------------V-----------------------------------
1628r16 ----------------------------------------K----------------------V-----------------------------------
1628r19 ----------------------------------------K----------------------V-----------------------------------
1628r21 ----------------------------------------K----------------------V-----------------------------------
1628r24 ----------------------------------------K----------------------V-----------------------------------
1628r25 ----------------------------------------K----------------------V-----------------------------------
<table>
<thead>
<tr>
<th>LAI</th>
<th>PR</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3188c01</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A--M---M---</td>
<td></td>
</tr>
<tr>
<td>3188c02</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A--M---M--F---</td>
<td></td>
</tr>
<tr>
<td>3188ca01</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188ca02</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188ca06</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188ca08</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188co01</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A--M---M--F---</td>
<td></td>
</tr>
<tr>
<td>3188co02</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A--M---M--F---</td>
<td></td>
</tr>
<tr>
<td>3188co03</td>
<td>--I--V--Q--VN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188co05</td>
<td>--I--V--Q--VN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188co06</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188co10</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188co12</td>
<td>--I--V--Q--VN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188co13</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188co15</td>
<td>--I--V--Q--VN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3188co18</td>
<td>--I--V--Q--VN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3357ca03</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---R-------------</td>
<td></td>
</tr>
<tr>
<td>3357ca04</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---R-------------</td>
<td></td>
</tr>
<tr>
<td>3357ca06</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---R-------------</td>
<td></td>
</tr>
<tr>
<td>3357ca07</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---R-------------</td>
<td></td>
</tr>
<tr>
<td>3357ca08</td>
<td>--I--V--Q--IN--T---L---V--E--E--VP---R-------------</td>
<td></td>
</tr>
<tr>
<td>3357co06</td>
<td>--I--V--Q--VN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3357co13</td>
<td>--I--V--Q--VN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td>3357co22</td>
<td>--I--V--Q--VN--T---L---V--E--E--VP---A-------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*(20)</td>
<td>*(40)</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>LAI</td>
<td>PR</td>
<td>PQITLWQRPLVTIKIGGQLKEALLTVDGTVLEEMSLPGRKPKMIGGGFIKVRQYDQLIEICGHKAIGTVLVGPTFVNIIGRNLLTQIGCTLN</td>
</tr>
<tr>
<td>*</td>
<td>*(20)</td>
<td>*(40)</td>
</tr>
</tbody>
</table>
| **VFIS18**

**Pretherapy**

| 2550pr | ---------I--V-----------------------N------------------------VNV----------------------------L------ |
| 2550u02 | ---------II-V-----------------------N------------------------VNV----------------------------L------ |
| 2550u03 | ---------I--V-----------------------N------------------------VNV----------------------------L------ |
| 2550u04 | ---------I--V-----------------------N--T-E-E----------VNVK----------------------------L--S-- |
| 2550u05 | ---------I--V-----------------------N------------------------VNV----------------------------L------ |
| 2550u06 | ---------I--V-----------------------N------------------------VNV----------------------------L------ |
| 2550u07 | ---------I--V-----------------------N--E-E----------VNVK----------------------------L------ |
| 2550u08 | ---------I--V-----------------------N------------------------VNV----------------------------L------ |
| 2550u09 | ---------I--V-----------------------N------------------------VNV----------------------------L------ |
| 2550u10 | ---------I--V-----------------------N------------------------VNV---------------------L------ |
| 2550u11 | ---------I--V-----------------------N------------------------VNV----------------------------L------ |
| 2550u12 | ---------I--V-----------------------N------------------------VNV----------------------------L------ |

**Posttherapy**

| 2930pr | ---------I--V-----------------------F--XN----------------X-------VNV-----------------A----------L------ |
| 2930c01 | ---------I--V-----------------------F--DIN----------------V-------VNV-----------------A----------L------ |
| 2930c02 | ---------I--V-----------------------F--DIN----------------V-------VNV-----------------A----------L------ |
| 2930c03 | ---------I--V-----------------------F--DIN----------------V-------VNV-----------------A----------L------ |
| 2930c04 | ---------I--V-----------------------F--IN----------------V-------VNV-----------------A----------L------ |
| 2930c05 | ---------I--V-----------------------F--DIN----------------V-------VNV-----------------A----------L------ |
| 2930c06 | ---------I--V-----------------------F--DIN----------------V-------VNV-----------------A----------L------ |
| 2930c07 | ---------I--V-----------------------F--DIN----------------V-------VNV-----------------LSTY |
| 2930c08 | ---------I--V-----------------------F--DIN----------------V-------VNV-----------------A---------- |
VFIS19

Pretherapy
1838pr  -------------------------G----------------------------------------VP-------------I---------------L------
1838r11  ---------------------------------------- --------------------VP-------------I---------------L---
1838r12  ----------------------------------------K--------------------VP-------------I------------------
1838r13  ----------------------------------------K--------------------VP-------------I------------------
1838r14  ----------------------------------------K--------------------VP-------------I---------------L----
1838r15  ----------------------------------------K--------------------VP---------S---I---------------L----
1838r16  ----------------------------------------K--------------------VP-------------I---------------L----

Posttherapy
2268pr  ---------------X-----------------------------------X---------------VP-------------I----------S----L------
2268c01  ---------------------------------------- --------------------VP-------------I----------S----L------
2268c02  ----------------------------------------R--------------------VP-------------I----------S----L------
2268c03  ----------------------------------------I----------------------I----------------------------VP-------------I----------S---HLP
2268c04  ----------------------------------------I----------------------I----------------------------VP-------------I----------S---L------
2268c05  ----------------------------------------I----------------------I----------------------------VP-------------I----------S---L------
2268c06  ----------------------------------------I----------------------I----------------------------VP-------------I----------S---L------
2268c07  ----------------------------------------I----------------------I----------------------------VP-------------I----------S---L------
2268c08  ----------------------------------------I----------------------I----------------------------VP-------------I----------S---S
VFIS22

Pretherapy
2041pr: R---------IN---K---------P
2041r11: R---------IN---K---------P
2041r12: R-E---------IN---K---------P
2041r13: R---------IN---K---------P
2041r14: R---------IN---K---------P
2041r15: R---------IN---K-------VP
2041r16: ---------IN---K-------VP

Posttherapy
2411pr: I-R-X---------X-IN---K---------V---------XP---------A
2411c01: R---------IN---K---------V---------VP---------A
2411c02: R---------IN---K---------V---------VP---------A-----------V
2411c03: R---------IN---K---------V---------VP---------A
2411c04: R---------IN---K---------V---------VP---------A
2411c05: R-TKGSF---------IN---K---------V---------VP---------A
2411c06: R---------IN---K---------V---------VP---------A
2411c07: R---------IN---K---------V---------VP---------A
2411c08: R---------IN---K---------V---------VP---------A
VFIS23
Pretherapy
1937pr  ----------------------------------D---------------------------P-------------I----------------------
1937o01 ----------------------------------D---------------------------P-------------I----------------------
1937o02 ----------------------------------D---------------------------P-------------I----------------------
1937o04 ----------------------------------D-N-------------------------P-------------I----------------------
1937o05 ----------------------------------D---------------------------P-------------I-----D----------------
Posttherapy
2282a01 ----------------------------------D---------------------------P-------------I----------------------
2282a02 ----------------------------------D-N-------------------------P-------------I----------------------
2282a03 ----------------------------------D---------------------------P-------------I----------------------
2282a04 ----------------------------------D---------------------------P-------------I----------------------
2282a05 -L--------------------------------D---------------------------P-------------I----------------------
2282o02 ----------------------------------D-N-------------------------P-------------I----------------------
2282o03 ----------------------------------D---------------------------P-------------I----------------------
2282o04 ----------------------------------D---------------------------P-------------I----------------------
2282o05 ----------------------------------D-N-------------------------P-------------I-------T----------------
VFIS24
Pretherapy
2098u01 -------------R---------------------ID-------------------------P------------------------------------
2098u03 -------------R---------------------ID-------------------------P------------------------------------
2098u04 -------------R---------------------ID-------------------------P------------------------------------
2098u05 -------------R---------------------ID-------------------------P--------M---------------------------
Posttherapy
2375u01 -------------R---------------------IN-------------------------P-------------------D----------------
2375u02 -------------R---------------------ID-------------------------P
2375u03 -------------R---------------------IN-------------------------P
2375u04 -------------R---------------------G-------------------------ID-------------------P
2375u05L -----------------------------------------------------VP----------------------R-S----------------
VFIS26
2996pr  ------------------V---------------------N--------L----------------P------ -----------A-----------------
VFIF05
Pretherapy
1877pr
1877u01
1877u03
1877u04
1877u05
Posttherapy
2308u02
2308u04
2308u05
2308u06
2308u07
VFI0F06

Pretherapy
1944r01 ----------------------------------D-N-------------------------------PL--------------------------
1944r03 ----------------------------------N------------------K--E--P---------------------------
1944r04 -------------A------------------N------------------K--E--P---------------------I-----------
1944r05 -----------------------------------IN------------------K--E--P---------------------------
1944r06 ----------------------------------N------------------K--E--P---------------------------
1944r07 ----------------------------------N------------------K--E--P---------------------------
1944r08 ----------------------------------R------------------N------------------K--E--P-------------------
1944r09 ----------------------------------N------------------K--E--P---------------------------
1944r10 ----------------------------------N------------------K--E--P---------------------------
1944r11 ----------------------------------N------------------K--E--P---------------------------
1944r12 ----------------------------------N------------------K--E--P---------------------------

Posttherapy
2292pr ----------------------------------D-N-------------------------------K--E--P---------------------------
2292u01 ----P------------------------------IN------------------K--E--P---------------------------
2292u02 -----------------------------------IN------------------K--E--P---------------------------
2292u03 ---------I------------------------D-N------------------K--E--P-------------------M---------
2292u04 ------------------------------------N------------------K--E--P---------------------------
2292u05 -----------------------------------IN------------------K--E--P---------------------------
2292r11 ----------------------------------D-N------------------K--E--P---------------------------
2292r12 ----------------------------------N------------------K--E--P---------------------------
2292r13 ----------------------------------N------------------K--E--P---------------------------
2292r14 ----------------------------------D-N------------------K--E--P---------------------------
2292r15 ----------------------------------N------------------K--E--P---------------------------
VF1F07
Pretherapy
2154pr  ---------I--------------------------N---K-R------------------VT-------T-----I---------------L------
794582  ---------I--------------------------N---K-R------------------VT-------T-----I---------------L------
794583  ---------I--------------------------N---K-R------------------VT-------T-----I---------------L------
794584  ---------I--------------------------N---K-R------------------VT-------T-----I---------------L------
7945811 ---------I--------------------------N---K-R------------------VT-------T-----I---------------L------

PostTherapy
1678pr  ---------I--------------------------N---K-R------------------VT-------T-----I---------------L------
1678r04         -I--------------------------N---K-R------------------VP-----Y-T-----I------------M--L------
1678r07         -I--------------------------N---K-R------------------VT-------T-----I---------------L------
1678r08        --I--------------------------N---K-R------------------VT-------T-----I---------------L------
1678u02 -X-------I-A----------------G-------N---K-R------------------VT-------T-----I---------------L------
1678u03 ---------I---------R----------------N---K-R------V-S---------VT-------T-----I---------------L------
1678u04 ---------I--------------------------N---K-R------------------VT-------T-----I---------------L------
1678u05 ---------I-P----------------G-------N---K-R------------------VT-------T-----I---------------L------
1678u06 ---------I--------------------------N---K-R------------------VT-------T-----I---------------L------
1678u07 ---------I--------------------------N---K-R------------------VP-------T-----I----I----------L------

VF1F08
Pretherapy
1934pr  -----------------------------------IN------------------------VP------------------------------------
1934u01 --------------V--------------------IN--------------------------------------------------------------
1934u03 --------------V--------------------IN--------------------------------------------------------------
1934u04 ---------I--------I-----------------N---------------------------------------I---------------L------
1934u06 --------------V---I----------------IN--------------------------------------------------------------
1934u08 --------------V--------------------IN--------------------------------------------------------------

Posttherapy
2234u02 -V---I----------------IN------------------------VP------------------------------------
2234u03 -V---I----------------IN------------------------VP------------------------------------
2234u04 -V---I----------------IN------------------------VP------------------------------------
2234u06 -V---I----------------IN------------------------VP------------------------------------
2234u08 -V---I----------------IN------------------------VP------------------------------------
APPENDIX B
ALIGNMENT OF GAG SEQUENCES
<table>
<thead>
<tr>
<th>B Site</th>
<th>C Site</th>
<th>D' Site</th>
<th>D Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>2272c01</td>
<td>---A---K---KG---</td>
<td>Q------E--G------------H---------N--------</td>
<td>---F-E--- ------K-------</td>
</tr>
<tr>
<td>2272c03</td>
<td>---A---A----K---KG----</td>
<td>I---------------------------</td>
<td>Q---------</td>
</tr>
<tr>
<td>2272c04</td>
<td>---A---A----K---KG-------------------</td>
<td>K--------</td>
<td>Q------E--G------------H---------N--------</td>
</tr>
<tr>
<td>2272c05</td>
<td>---A---A----KD--KG-----------------------</td>
<td>K--------</td>
<td>Q------E--G------------H---------N-------</td>
</tr>
<tr>
<td>2774a03</td>
<td>---A---A----K---KG--------------------------------</td>
<td>Q------E--G------------H---------N--------</td>
<td>EES-RVG-A-----------</td>
</tr>
<tr>
<td>2774a04</td>
<td>---A---A----K---KG--------------------------------</td>
<td>R------E--G------------H---------N--------</td>
<td>EES-RVG-A-----------</td>
</tr>
<tr>
<td>2774a05</td>
<td>---A---A----K---KG--------------------------------</td>
<td>Q------E--G------------H---------D--------</td>
<td>EES-RVG-A-----------</td>
</tr>
<tr>
<td>2774o01</td>
<td>---A---A----K---KG--------------------------------</td>
<td>Q------E--G------------H---------N--------</td>
<td>EES-RVG-T-----------</td>
</tr>
<tr>
<td>2774o03</td>
<td>---A---A----K---KG--------------------------------</td>
<td>Q------E--G------------H---------N--------</td>
<td>EES-RVG-T-----------</td>
</tr>
<tr>
<td>2774o05</td>
<td>---A---A----K---KG--------------------------------</td>
<td>Q------E--G------------H---------D--------</td>
<td>EES-RVG--------A----</td>
</tr>
<tr>
<td>1722o01</td>
<td>---H- -----V------------T----------------------------R----------------------H------------------</td>
<td>-----F-E-----------</td>
<td>M--D----A--K----------</td>
</tr>
<tr>
<td>1722o02</td>
<td>---H- -----V------------T----------------------------R----------------------</td>
<td>H--------------A---</td>
<td>-----F-E-----------</td>
</tr>
<tr>
<td>1722a02</td>
<td>---h---V------------T----------------------------R----------------------H------------------</td>
<td>-----F-E-----------</td>
<td>M--D----A--K----------</td>
</tr>
<tr>
<td>2088o01</td>
<td>---h---V------------T----------------------------R----------------------</td>
<td>H------------------</td>
<td>-----F-E-----------</td>
</tr>
<tr>
<td>2088o02</td>
<td>---V- ---T----------------------------R----------------------H------------------</td>
<td>-----F-E-----------</td>
<td>M--D----A--K----------</td>
</tr>
<tr>
<td>2088o04</td>
<td>---V- ---T----------------------------R----------------------H------------------</td>
<td>-----F-E-----------</td>
<td>M--D----A--K----------</td>
</tr>
<tr>
<td>2088o06</td>
<td>---V- ---T----------------------------R----------------------H------------------</td>
<td>-----F-E-----------</td>
<td>M--D----A--K----------</td>
</tr>
<tr>
<td>2088a01</td>
<td>------V------------T----------------------------R----------------------H------------------</td>
<td>-----F-E-----------</td>
<td>M--D----A--K----------</td>
</tr>
<tr>
<td>2088a02</td>
<td>---V- ---T----------------------------R----------------------H------------------</td>
<td>-----F-E-----------</td>
<td>M--D----A--K----------</td>
</tr>
<tr>
<td>2088a04</td>
<td>------V------------T----------------------------R----------------------H------------------</td>
<td>-----F-E-----------</td>
<td>M--D----A--K----------</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VS053</th>
</tr>
</thead>
<tbody>
<tr>
<td>1794u02</td>
</tr>
<tr>
<td>1794u03</td>
</tr>
<tr>
<td>1794u04</td>
</tr>
<tr>
<td>1794u05</td>
</tr>
<tr>
<td>1794u06</td>
</tr>
<tr>
<td>1794u07</td>
</tr>
<tr>
<td>1794u08</td>
</tr>
<tr>
<td>1794u09</td>
</tr>
<tr>
<td>1794u11</td>
</tr>
<tr>
<td>1794u12</td>
</tr>
<tr>
<td>2590a03</td>
</tr>
<tr>
<td>2590a04</td>
</tr>
<tr>
<td>2590a05</td>
</tr>
<tr>
<td>2590a06</td>
</tr>
<tr>
<td>2590a07</td>
</tr>
<tr>
<td>2590a08</td>
</tr>
<tr>
<td>2590a09</td>
</tr>
<tr>
<td>2590a10</td>
</tr>
<tr>
<td>2590a11</td>
</tr>
<tr>
<td>2590a12</td>
</tr>
</tbody>
</table>

Note: The table includes amino acid sequences and positions for various sites, showing variations in amino acid sequences at different positions across the sites.
<table>
<thead>
<tr>
<th>VT1009</th>
<th>gagAI</th>
<th>KARVLAXAMIQV TNSATIMBMFQHMQKIVCFNCGKEGHIARNCRAPRKKCGCNCGKEGQKDCETERQANFLISKRGPGNLQSRQPP</th>
<th>FLQSRQPPPTAPPPEFRSGVEETTTPSQKREPIDEKELVPLTSASLFHDPSSQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1955c01</td>
<td>---I-</td>
<td>-----</td>
<td>--PT----P-T---K----E------V-</td>
</tr>
<tr>
<td>1955c02</td>
<td>H-T-</td>
<td>-IQ-</td>
<td>--------T---------K------I------E------V-</td>
</tr>
<tr>
<td>1955c03</td>
<td>---</td>
<td>-T-</td>
<td>--------T---------K------I------E------V-</td>
</tr>
<tr>
<td>1955c08</td>
<td>---</td>
<td>-PT---T---K---S</td>
<td>--------K------I------E------V-</td>
</tr>
<tr>
<td>2686c02</td>
<td>--TN---K---T</td>
<td>--------K------I------E------S---V-</td>
<td>-----F-E---A-P------A------A------</td>
</tr>
<tr>
<td>2686c03</td>
<td>---------</td>
<td>-TN---K---T---K</td>
<td>--------K------I------E------S---V-</td>
</tr>
<tr>
<td>2686c04</td>
<td>---------</td>
<td>-TN---K---T---K</td>
<td>--------K------I------E------S---V-</td>
</tr>
<tr>
<td>2686c05</td>
<td>---------</td>
<td>-TN---K---T---K</td>
<td>--------K------I------E------S---V-</td>
</tr>
<tr>
<td>2686c06</td>
<td>---------</td>
<td>-TN---K---T---K</td>
<td>--------K------I------E------S---V-</td>
</tr>
<tr>
<td>2686c07</td>
<td>---------</td>
<td>-TN---K---T---K</td>
<td>--------K------I------E------S---V-</td>
</tr>
<tr>
<td>2686c08</td>
<td>---------</td>
<td>-TN---K---T---K</td>
<td>--------K------I------E------S---V-</td>
</tr>
<tr>
<td>2686u04</td>
<td>---</td>
<td>-T-</td>
<td>--------T---------K------I------E------S---V-</td>
</tr>
</tbody>
</table>
REFERENCES


179. **Peden, K., M. Emerman, and L. Montagnier.** 91 A.D. Changes in growth properties on passage in tissue culture of viruses derived from infectious molecular clones of HIV-1LAI, HIV-1MAL, and HIV-1ELI. Virology 185:661-672.


BIOGRAPHICAL SKETCH

Stephanie Leigh Rose was born in Hampton, Virginia, in 1971 to Rose and Russell Rose. She graduated from Menchville High School, Newport News, VA, in 1989. She then received her Bachelor of Science Cum Laude degree (majoring in biology) from Christopher Newport University in 1993. She worked for SRA Technologies Inc, a small biotechnology company in Rockville, MD before entering graduate school. She entered the University of Florida Interdisciplinary Program in Biomedical Sciences in 1997. After completion of her degree, she will pursue postdoctoral training in the laboratory of Maureen Goodenow.