TARGETING URACIL EXCLUSION MECHANISMS FOR DEVELOPMENT OF ANTI-VIRAL AND ANTI-CANCER THERAPIES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Adam W. Studebaker, B.S.

* * * * *

The Ohio State University
2003

Dissertation Committee:

Dr. M.V. Williams, Adviser

Dr. P.R. Johnson

Dr. W.P. Lafuse

Dr. W.J. Waldman

Approved by

_____________________

Adviser

Department of Molecular Virology

Immunology and Medical Genetics
ABSTRACT

Uracil has the potential to be deleterious when introduced into DNA and because of this therapeutic approaches targeting viruses and cancer cells have been developed to increase the likelihood of uracil formation in DNA. However, organisms have evolved very efficient mechanisms to prevent uracil formation/retention in DNA. Two of the most well characterized mechanisms involve the enzymes uracil DNA glycosylase (UDG) and deoxyuridine triphosphate nucleotidohydrolase (dUTPase). In an attempt to further understand the roles of UDG and dUTPase in viruses and tumor cells, as well as determine whether these enzymes could possibly be targets of future chemotherapy, a protein transduction technique and a small interfering RNA (siRNA) approach were utilized. An inhibitor protein of UDG (UGI) coupled to the protein transduction domain of the human immunodeficiency virus type 1 Tat protein (TaT-UGI) was able to efficiently transduce multiple cell types and inhibit cellular UDG activity. UDG activity was decreased approximately 94.4 ± 4.1% in U-937 cells, 29.6 ± 5.6% in HeLa cells, 69.9 ± 5.0% in HT29 cells, and 78.9 ± 1.0% in SW620 cells. On the other hand, a double stranded siRNA synthesized against domain 3 (nucleotides 339 to 357) of human dUTPase (siRNA-dut3) was able to diminish cellular dUTPase activity in all cell types examined. dUTPase activity was decreased approximately 50.0 ± 0.9% in HeLa and HT29 cells and approximately 27.0 ± 11.0 in SW620 cells. While TaT-UGI decreased
cellular protein concentration, dUTPase levels, and doubling times in cells treated for 72 hours with the highest concentration of TaT-UGI, siRNA-dut3 had no other statistically significant effects on the cells, other then decreasing dUTPase levels, when compared to non-treated controls. These results demonstrate that TaT-UGI and siRNA-dut3 can be used to decrease UDG and dUTPase activities, respectively, in human cells and that such an approach could be useful either alone or in combination with established therapies for anti-viral and/or cancer chemotherapy. Moreover, these reagents can be used to elucidate the functions/roles of UDG and dUTPase in normal cells and viral homeostasis so that rational therapeutic approaches can be developed targeting these enzymes.
This work is dedicated to my

Mom and Dad

Thank You for Everything!
ACKNOWLEDGMENTS

I wish to thank my adviser, Dr. Marshall V. Williams for his intellectual support, encouragement, and enthusiasm injected into these studies, as well as the patience he exhibited as I matured as a researcher. Most importantly, I would like to thank him for the friendship we developed.

I am grateful to Dr. Ganesaratnam K. Balendirian, Dr. Erwann Le Rouzic, Dr. Serge Benichou, and Dr. Maria E. Ariza for providing me with invaluable data, without which I would never have been able to a complete and intellectual story.

I wish to thank Dr. Louis M. Mansky, Dr. W. James Waldman, and Dr. Philip R. Johnson for providing me both intellectual insight and reagents that enabled me to accomplish these studies.

Lastly, I wish to thank Dr. William P. Lafuse for his never-ending optimism, color printer, and the friendship that we have developed.
VITA

December 20, 1971…………………..Born – Columbus, Ohio

1995……………………………….B.S. Biology, University of Illinois

1998-2003……………………….Graduate Teaching and Research Associate, The Ohio State University


PUBLICATIONS

Research Publications


FIELDS OF STUDY

Major Field: Molecular Virology, Immunology and Medical Genetics

Minor Field: DNA Repair and Nucleotide Metabolism
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xii</td>
</tr>
<tr>
<td><strong>Chapters:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>1. INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Human Immunodeficiency Virus Type 1 Disease and Structure</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1 AIDS Epidemic</td>
<td>1</td>
</tr>
<tr>
<td>1.1.2 Genome Structure and Composition</td>
<td>2</td>
</tr>
<tr>
<td>1.1.3 HIV-1 Life-cycle</td>
<td>12</td>
</tr>
<tr>
<td>1.2 Human Immunodeficiency Virus Type 1 Variability and</td>
<td>13</td>
</tr>
<tr>
<td>Genetic Diversity</td>
<td></td>
</tr>
<tr>
<td>1.2.1 Reverse Transcription</td>
<td>13</td>
</tr>
<tr>
<td>1.2.2 Variability and Mutagenesis</td>
<td>19</td>
</tr>
<tr>
<td>1.2.3 Reverse Transcription</td>
<td>23</td>
</tr>
<tr>
<td>1.3 Viral Protein R</td>
<td>26</td>
</tr>
<tr>
<td>1.3.1 Background</td>
<td></td>
</tr>
<tr>
<td>1.3.2 Association with Cellular Proteins</td>
<td>27</td>
</tr>
<tr>
<td>1.3.3 Association with Uracil DNA Glycosylase</td>
<td>28</td>
</tr>
<tr>
<td>1.4 Uracil DNA Glycosylase</td>
<td>31</td>
</tr>
<tr>
<td>1.4.1 Background</td>
<td></td>
</tr>
<tr>
<td>1.4.2 Additional Enzymes with Uracil DNA Glycosylase Activity</td>
<td>32</td>
</tr>
<tr>
<td>1.5 Uracil Formation in DNA</td>
<td>36</td>
</tr>
<tr>
<td>1.5.1 dNTP Pools and dUTP Mis-incorporation</td>
<td>36</td>
</tr>
<tr>
<td>1.5.2 Cytosine Deamination</td>
<td>36</td>
</tr>
<tr>
<td>1.5.3 Evidence of Uracil Formation in HIV-1 DNA</td>
<td>37</td>
</tr>
</tbody>
</table>
1.6 Base Excision Repair
   1.6.1 Role.................................................................40
   1.6.2 Components of the BER Pathway.................................41
   1.6.3 Long-Patch BER ..................................................41
1.7 Uracil DNA Glycosylase Inhibitor....................................42
1.8 Importance of UDG.....................................................46
1.9 Deoxyuridine Triphosphate Nucleotidohydrolase.....................49
1.10 Significance of Research.............................................51

2. PURIFICATION AND FUNCTIONAL PROPERTIES OF A NATIVE
   COMPLEX BETWEEN URACIL DNA GLYCOSYLASE AND HIV-1 Vpr

2.1 Introduction............................................................54
2.2 Materials and Methods................................................56
   2.2.1 Materials..........................................................56
   2.2.2 Bacterial Strains and Plasmids..................................57
   2.2.3 UDG and UGI Assays..............................................57
   2.2.4 Protein Assays....................................................58
   2.2.5 Vpr-UDG Purification............................................58
   2.2.6 SDS-PAGE and Western Blotting...............................60
   2.2.7 Yeast Two-Hybrid Assay.........................................61
   2.2.8 Analysis of UDG Incorporation into HIV-1 Virions............62
   2.2.9 Molecular Modeling..............................................63
2.3 Results..........................................................................64
   2.3.1 Purification of the Vpr-\textit{E. coli} UDG Complex................64
   2.3.2 Biochemical Properties of the Vpr-UDG\textit{coli} Complex ..........70
   2.3.3 \textit{E. coli} UDG Can Support Direct Interaction with Vpr........73
   2.3.4 \textit{E. coli} UDG Can be Incorporated into HIV-1 Virion...........77
   2.3.5 Structural Model of the Vpr-UDG and UGI Complex............82
2.4 Discussion.....................................................................83

3. MECHANISM(S) OF URACIL INCORPORATION INTO
   HUMAN IMMUNODEFICIENCY TYPE 1 DNA

3.1 Introduction....................................................................89
3.2 Materials and Methods................................................93
   3.2.1 Materials..........................................................93
   3.2.2 Cell Culture and Extraction......................................94
   3.2.3 Nucleotide Pool Assays.........................................95
   3.2.4 Nucleotide Insertion/Mismatch Extension Assay..............96
   3.2.5 Effect of UGI on Reverse Transcription.......................98
3.3 Results..........................................................................98
   3.3.1 Ability to Incorporate and Extend an Incorrect Nucleotide ..98
   3.3.2 Pyrimidine Deoxynucleotide Triphosphate Pool Concentrations .99
   3.3.3 Effect of UGI on Reverse Transcription.......................100
3.4 Discussion...................................................................................................... 102

4. PURIFICATION AND CHARACTERIZATION OF TAT-UGI, A NOVEL PROTEIN TRANSDUCTION MOLECULE

<table>
<thead>
<tr>
<th>4.1 Introduction</th>
<th>107</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2 Materials and Methods</td>
<td>110</td>
</tr>
<tr>
<td>4.2.1 Materials</td>
<td>110</td>
</tr>
<tr>
<td>4.2.2 Construction of TaT-UGI</td>
<td>110</td>
</tr>
<tr>
<td>4.2.3 TaT-UGI Purification</td>
<td>110</td>
</tr>
<tr>
<td>4.2.4 SDS-PAGE and Western Blotting</td>
<td>112</td>
</tr>
<tr>
<td>4.2.5 Protein Assays</td>
<td>113</td>
</tr>
<tr>
<td>4.2.6 UDG and UGI Assays</td>
<td>113</td>
</tr>
<tr>
<td>4.2.7 Stability Studies</td>
<td>114</td>
</tr>
<tr>
<td>4.2.8 DUTPase Enzyme Assays</td>
<td>114</td>
</tr>
<tr>
<td>4.2.9 Cell Culture and TaT-UGI Treatment</td>
<td>115</td>
</tr>
<tr>
<td>4.2.10 FITC Labeling</td>
<td>116</td>
</tr>
<tr>
<td>4.2.11 Cell Proliferation/Chemosensitivity Assays</td>
<td>117</td>
</tr>
<tr>
<td>4.3 Results</td>
<td>118</td>
</tr>
<tr>
<td>4.3.1 Purification of TaT-UGI</td>
<td>118</td>
</tr>
<tr>
<td>4.3.2 Biochemical Properties of TaT-UGI</td>
<td>118</td>
</tr>
<tr>
<td>4.3.3 Cellular UDG Activity</td>
<td>119</td>
</tr>
<tr>
<td>4.3.4 Cellular Protein Concentration</td>
<td>123</td>
</tr>
<tr>
<td>4.3.5 Effect of TaT-UGI on Cell Proliferation</td>
<td>124</td>
</tr>
<tr>
<td>4.4 Discussion</td>
<td>136</td>
</tr>
</tbody>
</table>

5. DOWN-REGULATION OF HUMAN DEOXYURIDINE TRIPHOSPHATE NUCLEOTIDOHYDROLASE (dUTPase) USING SMALL INTERFERING RNA (siRNA)

<table>
<thead>
<tr>
<th>5.1 Introduction</th>
<th>142</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2 Materials and Methods</td>
<td>146</td>
</tr>
<tr>
<td>5.2.1 Materials</td>
<td>146</td>
</tr>
<tr>
<td>5.2.2 UDG and UGI Assays</td>
<td>146</td>
</tr>
<tr>
<td>5.2.3 Protein Assays</td>
<td>147</td>
</tr>
<tr>
<td>5.2.4 dUTPase Enzyme Assays</td>
<td>147</td>
</tr>
<tr>
<td>5.2.5 Preparation of siRNA</td>
<td>147</td>
</tr>
<tr>
<td>5.2.6 Transfection of Cells with siRNA</td>
<td>148</td>
</tr>
<tr>
<td>5.2.7 Northern Blot Hybridization</td>
<td>149</td>
</tr>
<tr>
<td>5.2.8 Cell Proliferation/Cytotoxicity Assays</td>
<td>150</td>
</tr>
<tr>
<td>5.3 Results</td>
<td>150</td>
</tr>
<tr>
<td>5.3.1 siRNAdUT3-Mediated Inhibition of dUTPase</td>
<td>150</td>
</tr>
<tr>
<td>5.3.2 siRNAdUT3 Influence on Cell Proliferation</td>
<td>152</td>
</tr>
<tr>
<td>5.4 Discussion</td>
<td>157</td>
</tr>
</tbody>
</table>
6. GENERAL DISCUSSION
   6.1 Introduction.................................................................................................161
   6.2 Review of Results..........................................................................................163
   6.3 Overall Conclusions.....................................................................................166

LIST OF REFERENCES...............................................................................................174
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>HIV-1 Proteins and Reported Functions/Roles</td>
</tr>
<tr>
<td>1.2</td>
<td>A Summary of the Fidelities Associated with Various DNA Polymerases</td>
</tr>
<tr>
<td>1.3</td>
<td>Human DNA Glycosylases that Can Excise Uracil</td>
</tr>
<tr>
<td>2.1</td>
<td>Purification of Vpr-UDGcoli Complex</td>
</tr>
<tr>
<td>2.2</td>
<td>Effect of Monovalent Cations and Polyamines on the Activities of <em>E. coli</em> UDG and the Vpr-UDGcoli Complex</td>
</tr>
<tr>
<td>3.1</td>
<td>Ability of Various Replicative Polymerases to Misincorporate the Incorrect Nucleotide and Extend from this Mismatch</td>
</tr>
<tr>
<td>3.2</td>
<td>Pyrimidine Deoxynucleotide Triphosphate Pool Concentrations</td>
</tr>
<tr>
<td>3.3</td>
<td>Effect of UGI on Reverse Transcription</td>
</tr>
<tr>
<td>4.1</td>
<td>Purification of TaT-UGI</td>
</tr>
<tr>
<td>4.2</td>
<td>Effect of Temperature and Buffer Conditions on the Stability of Purified TaT-UGI</td>
</tr>
<tr>
<td>4.3</td>
<td>Effect of TaT-UGI on Cell Proliferation</td>
</tr>
<tr>
<td>4.4</td>
<td>Summary of the Cellular Effects Associated with TaT-UGI</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>Genome Organization of HIV-1</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>HIV-1 Life-cycle</td>
<td>15</td>
</tr>
<tr>
<td>1.3</td>
<td>The Reverse Transcription Process</td>
<td>18</td>
</tr>
<tr>
<td>1.4</td>
<td>Locations of Vpr’s Structural Domains</td>
<td>29</td>
</tr>
<tr>
<td>1.5</td>
<td>Amino Acid Alignment of the Uracil DNA Glycosylases</td>
<td>34</td>
</tr>
<tr>
<td>1.6</td>
<td>Mammalian Pyrimidine Deoxyribonucleotide Metabolism</td>
<td>39</td>
</tr>
<tr>
<td>1.7</td>
<td>Pathways for Base Excision Repair in Mammalian Cells</td>
<td>44, 45</td>
</tr>
<tr>
<td>2.1</td>
<td>SDS-PAGE and Western Blotting</td>
<td>69</td>
</tr>
<tr>
<td>2.2</td>
<td>Enhanced Stability Imparted on <em>E. coli</em> UDG upon Binding to Vpr</td>
<td>72</td>
</tr>
<tr>
<td>2.3</td>
<td>Inhibition of UDG by Tagged UGI</td>
<td>74</td>
</tr>
<tr>
<td>2.4</td>
<td>Ability of UGI to Interact with a Vpr-UDGcoli Complex</td>
<td>76</td>
</tr>
<tr>
<td>2.5</td>
<td>Specific Interaction of Vpr with UDGcoli in a Yeast Two-Hybrid System</td>
<td>79</td>
</tr>
<tr>
<td>2.6</td>
<td>Incorporation of UDGcoli into HIV-1 Virions</td>
<td>81</td>
</tr>
<tr>
<td>2.7</td>
<td>Model of Vpr-UNG-UGI Interaction</td>
<td>85</td>
</tr>
<tr>
<td>4.1</td>
<td>SDS-PAGE</td>
<td>121</td>
</tr>
<tr>
<td>4.2</td>
<td>Effects of a Single dose of TaT-UGI on U-937 Cells</td>
<td>126</td>
</tr>
</tbody>
</table>
4.3 Effects of a Multiple Dose of TaT-UGI on U-937 Cells..........................128

4.4 Effects of TaT-UGI Dosage on UDG Activity in SW620, HT29, and HeLa Cells.........................................................130

4.5 Cellular Effects of TaT-UGI on SW620, HT29, and HeLa Cells.........132

4.6 Effect of TaT-UGI on Cell Proliferation........................................133

5.1 Homotrimeric dUTPase.................................................................144

5.2 Model for the RNAi Pathway.........................................................145

5.3 Effect of siRNAdUT3 on dUTPase Activity....................................154

5.4 Effect of siRNAdUT3 on DUT mRNA Expression............................155

5.5 Effect of siRNAdUT3 on DUT mRNA Expression............................156

5.6 siRNAdUT3 Influence on Cell Proliferation....................................158

6.1 Hypothetical Model for the Role of dUTPase in Non-Primate Lentivirus Replication.........................................................170

6.2 Hypothetical Model for the Role of UDG in Primate Lentivirus Replication .................................................................171

6.3 Model Depicting the Potential Effects of Inhibiting dUTPase and/or UDG.............................................................173
CHAPTER 1

INTRODUCTION

1.1 Human Immunodeficiency Virus Type 1 Disease and Structure

1.1.1 AIDS Epidemic

In the early 1980’s the human immunodeficiency virus (HIV) was first identified as the causative agent of the acquired immunodeficiency syndrome (AIDS). Since this initial discovery AIDS has become a major worldwide epidemic. AIDS is characterized by severe immunosuppression, opportunistic infections, and neoplasms caused by the depletion of the CD4+ T-lymphocyte population (1). The U.S. Centers for Disease Control and Prevention ascribes a diagnosis of AIDS to HIV-infected individuals whose CD4+ T-cell counts fall below 200 cells/mm³ of blood. A healthy individual usually has CD4+ T-cell counts of 600-1,500 cells/mm³ of blood.

HIV/AIDS statistics provided in a May 30, 2002 world health organization press release indicate that 40 million people in the world are living with the HIV virus, of which 70 % of them live in Sub-Saharan Africa. It is estimated that 15,000 individuals become newly infected with HIV every day, 95 % of these infections are occurring in developing countries. This epidemic has claimed approximately 30 million lives worldwide making it the leading cause of death in the African continent and the fourth
leading cause of death worldwide. Statistics provided by the National Institute of Allergy and Infectious Diseases indicated that worldwide, more than 80% of all adult HIV infections resulted from heterosexual intercourse and 90% of all infant and child infections occurred from mother-to-child transmission. An even more tragic figure is that in 16 African countries, the incidence of infection among adults aged 15 to 49 exceeds 10 percent. These staggering numbers indicate that an effective and, most importantly, affordable treatment needs to be developed if long-term control of the AIDS epidemic is to be achieved.

1.1.2 Genome Structure and Composition

Human immunodeficiency virus type I (HIV-1) is a member of the lentivirus genus of the Retroviridae family. Members of the Retroviridae family have been associated with slow forming malignancies, neurological disorders, and immunodeficiencies (2). Retrovirus members have a similar virion structure, genome organization, and mode of replication. The genome of HIV consists of two identical molecules of positive polarity single-stranded RNA, each about 9.2 kb long (Fig. 1.1). Retroviruses are unique in that they undertake a process termed “reverse transcription” in which their genomic RNA is converted into double-stranded linear DNA (reviewed in sec. 1.2.1). This process is necessary for proper expression of their encoded genes and is quintessential characteristic of retroviruses. The viral RNA is similar to mRNA in that it contains a 5' cap structure and a 3' poly-A tail. In addition there are numerous important noncoding regions found within the genomic RNA (Fig. 1.1), including:
1. The R region located at the 5′ end and at the 3′ end of the genome is important during reverse transcription by permitting DNA strand transfer from one end of the genome to the other.

2. The U5 region is the first region copied during reverse transcription and ultimately becomes the 3′ end of the LTR. The 3′ end of this region contains an att site necessary for integration.

3. A cellular tRNA\(^{\text{Lys}}\) anneals to the primer binding site (PBS) and initiates reverse transcription.

4. The packaging signal (ψ) binds nucleocapsid (NC) and is critical for incorporation of genomic RNA into the virion (3).

5. The donor site (SD) is used to generate all spliced subgenomic mRNAs encoded by the virus.

6. The spliced acceptor (SA) sites allow for the production of spliced products of variable lengths.

7. The polypurine tract (PPT) serves as the primer for synthesis of the plus strand of viral DNA. Lentiviruses are unique in that they also utilize a central PPT (cPPT) for synthesis of plus strand DNA (4,5).

8. The U3 region forms the 5′ portion of the LTR. When converted to the DNA form, this region contains numerous \(cis\)-acting signals necessary for virus replication.

The HIV virion is spherical shaped, about 100 nm in diameter, and enveloped by a lipid bilayer derived from the membrane of the host cell. Lining the inner surface of the viral membrane is a shell consisting of approximately 2000 copies of the matrix (MA) protein.
Figure 1.1: Genome Organization of HIV-1.  (A) + RNA Genome.  (■) Primer binding site; (ψ) encapsidation sequence; (■) splice donor site; (■) splice acceptor site; (●) central polypurine tract (cPPT); (■) polypurine tract (PPT).  (B) Depiction of proviral DNA.  Genes are encoded in all three reading frames.  Splicing patterns of HIV-1 transcripts.  The gag and gag-pol mRNAs are unspliced; mRNAs for the late genes vif, vpr, vpu, and env are singly spliced; mRNAs for the early genes tat, rev, and nef are doubly spliced.  and a conical shaped core particle consisting of 2000 copies of the capsid (CA) protein.
Located within the capsid shell are two copies of the viral genome stabilized by approximately 2000 molecules of the nucleocapsid (NC) protein. The capsid particle also contains protease, reverse transcriptase, and integrase molecules, three virally encoded proteins important in catalytic roles during replication. While the accessory proteins, Nef, Vif and Vpr are packaged into virions; the accessory proteins Rev, Tat and Vpu appear not to be.

HIV-1 encodes three structural proteins, two envelope proteins, three enzymes, and six accessory proteins. The proteins are products of either unspliced, single spliced, and double spliced RNA transcripts (Fig. 1.1). The structural genes, \textit{gag}, \textit{pol}, \textit{env} encode for polyproteins which are subsequently proteolyzed into individual proteins. The Gag polyprotein is cleaved into the MA, CA, and NC proteins. The three Gag proteins along with the two Env proteins, surface or gp120 (SU) and transmembrane or gp41 (TM), are the structural components that make up the core of the virion and envelope. The cleavage of the Pol polyprotein produces the protease (PR), reverse transcriptase (RT), and integrase (IN) molecules. These proteins are incorporated into the virion as part of the Pol polyprotein and provide essential enzymatic functions following cleavage. In addition to the structural genes, HIV-1 also encodes for two regulatory genes, \textit{tat} and \textit{rev}. The proteins expressed from these genes, Tat and Rev, serve as transactivators essential for viral replication. The remaining four proteins Vif, Vpr, Vpu, and Nef are classified as accessory proteins because they are non-essential for virus replication in cell culture (6).
HIV-1 is genetically different than human immunodeficiency virus type II (HIV-2), which was initially isolated from patients in West Africa (7), in that the accessory gene \textit{vpu} is unique to HIV-1 isolates (8), whereas HIV-2 isolates encodes for the accessory gene \textit{vpx} (9,10).

\textbf{Structural Proteins}

The functions of the individual HIV-1 encoded proteins are summarized in Table 1. Briefly, MA is important for targeting Gag and Gag-Pol polyproteins to the cell membrane prior to assembly. It also appears that MA is involved with the incorporation of Env glycoproteins with long cytoplasmic tails into viral particles (11,12). MA has been shown to facilitate infection of non-dividing cells (e.g. macrophages), although the mechanism has yet to be discerned unequivocally.

The CA protein forms a shell that surrounds the RNA-protein complex in a mature virion. NMR studies determined that CA contains two domains (13). Deletion and mutational studies have determined that the C-terminal domain functions primarily in assembly, but it is also important in CA dimerization and Gag oligomerization and the N-terminal domain is important for infectivity (14-20). CA also binds the cellular protein, cyclophilin A (CypA) (21,22). It is thought that CypA increases viral infectivity by destabilizing the capsid during uncoating (23,24).

NC is the third protein proteolyzed from the Gag polyprotein. It is a basic, hydrophilic protein that binds genomic viral RNA inside the virion core (25,26). As a domain within the Gag polyprotein, NC functions in the recognition and packaging of the viral genome (27-29). This process is mediated by the interaction of NC with a region of
the viral RNA known as the packaging signal (ψ) (30-32). This domain is also 
responsible for the packaging of the reverse transcription primer tRNA\textsubscript{Lys} (33).

Following cleavage of the Gag precursor, in addition to coating the genome, NC is 
involved in initiating reverse transcription by annealing the tRNA\textsubscript{Lys} to the viral genome
(34), facilitates synthesis of viral DNA by reducing reverse transcriptase stalling at stem-
loop sites (35), and stabilizes proviral DNA (36).

**Enzymes**

PR provides an essential role in HIV maturation from noninfectious to infectious.

Following particle release PR releases itself from the Gag-Pol precursor by an
autocatalytic mechanism (37-39) allowing PR to cleave the Gag and Gag-Pol
polyproteins at several sites to produce the MA, CA, NC, and p6 proteins from Gag and
PR, RT, and IN proteins from Pol. RT catalyzes both RNA-dependent and DNA-dep-
dendent DNA polymerization reactions (40,41). These activities allow the RNA
genome of HIV to be converted into DNA, a process referred to as reverse transcription,
so that the viral genome can be integrated into the host chromosome. RT also contains an
RNase H domain that cleaves the RNA strand of RNA-DNA hybrids generated during
reverse transcription. Following reverse transcription, IN is involved in a series of
reactions that result in the incorporation of the double-stranded viral DNA into the
chromosomal DNA of the target cell. IN is capable of undertaking this process by
possessing both DNA cleavage and strand transfer activities. The mechanism of
integration can be found in a review by Hindmarsh and Leis (42).
Accessory Proteins

HIV also expresses several additional proteins, referred to as accessory proteins. Although these proteins are not required for viral replication in tissue culture systems they are necessary for HIV to cause disease. HIV negative factor (Nef) is a myristoylated protein that is expressed in high concentrations following viral infection (43). Studies involving a nef mutant strain of simian immunodeficiency virus (SIV) demonstrated that nef was necessary for achieving and maintaining high viral loads (44). A similar finding was subsequently found in HIV-infected individuals containing deletions in the nef gene (45,46). Nef has been shown to have at least two distinct functions that may influence HIV infectivity and pathogenicity. It is responsible for reducing the number of CD4 molecules, the receptor for HIV, on the cell surface (47,48). Down-regulation of the CD4 receptor appears to be important to viral production because it prevents re-infection by budding virions (49-51). Nef contains a consensus SH3 domain binding sequence (PXXP) that allows it to interact with members of the Src family of tyrosine kinases (52). It is thought that Nef may contribute to viral-mediated cellular activation through signal transduction pathways.

The virion infectivity factor (Vif) protein is essential for the establishment of a productive infection in peripheral blood lymphocytes and macrophages (53-56). Vif is abundantly expressed in the cytoplasm of infected cells and can be found at low levels in HIV virions (57,58). Mutational studies on the HIV vif gene indicated that Vif increases the infectivity of cell-free virus as much as 100- to 1000-fold (54,56). Virus particles produced in the absence of Vif can enter cells, but cannot efficiently synthesize proviral
DNA (59). While the steps in the viral life cycle controlled by Vif have yet to be resolved previous studies have shown that Vif enhances viral infectivity during virus particle production, presumably by affecting assembly and/or maturation (60-64).

Vpu is an integral membrane protein that is primarily localized in the perinuclear region (65). Vpu is involved in the degradation of CD4 and the enhancement of virion release. The simultaneous expression of CD4 and newly synthesized Env glycoproteins (gp160) in the endoplasmic reticulum leads to the formation of a complex between the molecules. This complex, in turn, decreases the amount of CD4 expressed on the cell surface and interferes with the cleavage of gp160 to gp120 and gp41 (66). Vpu dissociates gp160 by facilitating the ubiquitin-mediated degradation of CD4 (67). The mechanism by which Vpu facilitates the release of HIV from the membrane of an infected cell is thought to occur through the formation of ion channels in the lipid bilayers (68).

Viral protein R (Vpr) is a virion-associated protein that provides multiple functions for HIV replication. These activities include induction of cell cycle arrest and apoptosis, nuclear import of the HIV genome, and transactivation of viral and cellular gene expression (reviewed in 69 and addressed in more detail in section 1.3.1).

**Envelope Proteins**

HIV entry into a host cell is initiated by the binding of the SU glycoprotein to a specific cell surface receptor. The major receptor recognized by HIV is CD4. SU is a highly glycosylated, hydrophilic protein located on the viral membrane surface and
contains five variable regions (V1-V5) (70,71). The binding of SU to CD4 is not sufficient to allow HIV entry; rather, a group of chemokine receptors serve as essential co-receptors. The interaction between SU and CD4 appears to cause structural changes in Env that allow co-receptor binding and subsequent viral entry (72,73). SU is also the primary antigen recognized by neutralizing antibodies in the infected host.

TM is a transmembrane protein that consists of an N-terminal ectodomain, a transmembrane domain, and a C-terminal segment that interacts with MA inside the virus. The primary function of TM is to mediate fusion between the viral membrane and the cell plasma membrane following receptor binding (74,75).

**Gene Regulatory Proteins**

Expression of the integrated proviral DNA is dependent upon the regulatory protein, Tat (76-79). Tat stimulates transcription of genes linked to the viral long terminal repeat (LTR) promoter (80-82). Unlike typical transcriptional activators, Tat binds to a trans-activating response element (TAR), a stem-loop site located at the 5’ end of all nascent viral RNA transcripts (83). Studies have shown that Tat functions through TAR to increase transcription elongation by blocking an early step that is sensitive to protein kinase inhibitors (84).

The gene regulatory protein, Rev, enhances the transport of unspliced and incompletely spliced viral mRNAs from the nucleus to the cytoplasm. The export of singly spliced and unspliced transcripts is necessary for Gag and Gag-Pol synthesis and packaging. Selective transport is accomplished by Rev recognition of the Rev response element (RRE) located in the env coding region (85,86).
<table>
<thead>
<tr>
<th>Protein</th>
<th>Properties/function of protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>Lines inner surface of virion membrane; targets Gag and Gag-Pol polyproteins to cell membrane</td>
</tr>
<tr>
<td>CA</td>
<td>Forms core of virus particle</td>
</tr>
<tr>
<td>NC</td>
<td>Recognition and packaging of viral genome; promotes annealing of tRNA primer; melting of RNA secondary structures</td>
</tr>
<tr>
<td>PR</td>
<td>Cleaves Gag and Gag-Pol polyproteins; essential for maturation from non-infectious to infectious particles</td>
</tr>
<tr>
<td>RT</td>
<td>Converts genomic RNA into linear double-stranded DNA; RNaseH activity</td>
</tr>
<tr>
<td>IN</td>
<td>Incorporation of double-stranded viral DNA into the chromosomal DNA of infected cells</td>
</tr>
<tr>
<td>Nef</td>
<td>Reduces expression of CD4 receptor and MHC class I; enhances virion infectivity; affects T-cell activation</td>
</tr>
<tr>
<td>Vif</td>
<td>Establishment of productive infection; affects virion assembly and/or viral DNA synthesis; associates with CEM 15</td>
</tr>
<tr>
<td>Vpu</td>
<td>Disrupts Env-CD4 complexes; degradation of CD4 receptor; enhancement of virion release</td>
</tr>
<tr>
<td>Vpr</td>
<td>Causes G2 cell cycle arrest and apoptosis; facilitates nuclear import of double-stranded viral DNA; transactivator</td>
</tr>
<tr>
<td>SU</td>
<td>Binds to CD4 receptor on cell</td>
</tr>
<tr>
<td>TM</td>
<td>Mediates viral and cellular fusion</td>
</tr>
<tr>
<td>Rev</td>
<td>Binds RRE and facilitates nuclear export of unspliced or singly spliced RNAs</td>
</tr>
<tr>
<td>Tat</td>
<td>Transcriptional activator; binds TAR to facilitate initiation and elongation of viral transcription</td>
</tr>
</tbody>
</table>

Table 1.1: HIV-1 Proteins and Reported Functions/ Roles.
1.1.3 HIV-1 Life-cycle

HIV-1 replication is schematically outlined in Fig. 1.2. Briefly, HIV-1 virions bind specifically to cells bearing CD4 molecules in an interaction mediated by specific interactions between SU and the amino-immunoglobulin domain of CD4 and CC or CXC chemokine co-receptors. Following binding, TM undergoes a conformational change that promotes virus cell membrane fusion. Primate lentiviruses, unlike other retroviruses, require co-receptor molecules to promote fusion of the viral and cellular membranes. The CXC chemokine receptor 4 (CXCR4) and the CC chemokine receptor 5 (CCR5) appear to be specific co-receptors for HIV-1 strains that preferentially infect T-cells and macrophages, respectively. The virion core is then deposited into the cytoplasm where the genomic RNA is reverse transcribed by RT into a double-stranded DNA molecule (Fig. 1.3). Once synthesized, the viral DNA is transported to the nucleus as part of a pre-integration complex. After nuclear localization, IN catalyzes the site-specific integration of the viral DNA into the host chromosome. Integrated viral DNA is called the provirus. Using the provirus as a template the host RNA polymerase II synthesizes full-length viral transcripts. Some of the full-length transcripts are spliced within the nucleus while others are exported from the nucleus. The full-length mRNAs either serve as viral mRNAs or become encapsidated as progeny viral genomes. Upon translation in the cytoplasm by the cellular machinery, full-length viral mRNAs form the viral Gag and Gag-Pol polyprotein precursors, while singly spliced mRNAs produce Vif, Vpr, and Vpu proteins and doubly spliced mRNAs produce Tat, Rev and Nef proteins. On the other hand, Env mRNA is translated at the endoplasmic reticulum. Following translation the Env proteins are glycosylated at the Golgi apparatus, cleaved to form the SU and TM molecules, and
delivered to the plasma membrane. The core particle consisting of the Gag and Gag-Pol polyproteins, Vif, Vpr, Nef, and the genomic RNA assembles at the plasma membrane. The particle then buds from the surface of the cell. Subsequent infection of new cells requires a morphologic change known as maturation. This process involves proteolytic processing of the Gag and Gag-Pol polyproteins by PR to produce the mature virion proteins. The mature virion is now infectious.

1.2 Human Immunodeficiency Virus Type 1 Variability and Genetic Diversity

1.2.1 Reverse Transcription

As stated earlier the process of reverse transcription distinguishes retroviruses (e.g. HIV) from other viruses. This process is essential for the survival of the virus and, as will be discussed in subsequent sections, has intrinsic characteristics that provide the virus with survival advantages. Reverse transcription is initiated when the viral core enters the cytoplasm of the target cell. The viral genome enters the cell as part of a nucleoprotein complex, which appears to resemble the virion capsid (87). Within the nucleoprotein complex are the essential components needed for reverse transcription: (a) two copies of the viral single-stranded RNA genome, (b) a host encoded tRNA, and (c) RT.

A schematic of the reverse transcription process is depicted in Fig. 1.3. Briefly, RT uses the 3’ end of a partially unfolded host tRNA\textsubscript{Lys} primer, which is annealed to the PBS in the genomic RNA, as a primer to initiate synthesis of the first DNA strand (e.g. minus strand DNA). Minus strand synthesis proceeds to the 5’ of the viral RNA genome,
Figure 1.2: HIV-1 Life-cycle. **Step 1:** Virions attach to cells expressing CD4 molecules and the appropriate co-receptor. **Step 2:** Following binding fusion between the virus and cell membranes occurs. **Step 3:** The virion core is deposited into the cytoplasm where the genomic RNA is reverse transcribed into linear double stranded DNA. **Step 4:** Viral DNA is then transported to the nucleus where it integrates into the host chromosome. Integrated viral DNA is referred to as the provirus. **Step 5:** The host RNA polymerase II transcribes the provirus into full-length mRNA transcripts. Some of these full-length transcripts are subsequently spliced. **Step 6:** Full-length mRNAs transported into the cytoplasm are either translated into Gag and Gag-Pol polyprotein precursors or end up serving as genomic RNA. **Step 7:** Env mRNA is translated in the ER, glycosylated in the Golgi, and then cleaved to form the SU and TM molecules. **Step 8:** SU and TM molecules are delivered to the plasma membrane. **Step 9:** Gag and Gag-Pol polyproteins, Vif, Vpr, Nef, and the genomic RNA assembles at the plasma membrane. **Step 10:** The nascent virion buds from the surface of the cell. Proteolytic processing of the Gag and Gag-Pol polyproteins produces mature virion proteins. The mature virion is now infectious.
Figure 1.2
minus strand strong-stop DNA (−sssDNA). Subsequently, the RNA portion of the RNA-DNA hybrid is digested by the RNaseH activity of RT, resulting in a single-strand DNA product. Continued minus strand DNA synthesis requires strand transfer of −sssDNA to the 3´ end of the viral genome (88-90). Newly synthesized regions in −sssDNA complementary to the R region at the 5´ end of the viral RNA genome hybridize with the R region at the 3´ end of the same or the second viral RNA genome. Following the first strand transfer RNA-dependent DNA polymerization (RDDP) of the minus strand DNA resumes along with RNaseH digestion of the RNA template strand. Plus strand DNA synthesis is initiated from an RNaseH resistant polypurine tract (PPT) located within the viral genome (91). Lentiviruses differ from other retroviruses in that they also utilize a central PPT (cPPT) as a primer for DNA-dependent DNA polymerization (DDDP) of plus strand DNA (92). Plus strand synthesis proceeds to the end of the minus strand DNA template and terminates after copying the annealed portion of the tRNA (89,93), thus generating the positive strand DNA form of the PBS. The product of this reaction is referred to as plus strand strong-stop DNA (+sssDNA). After removal of the tRNA primer by the RNaseH activity of RT, a second strand transfer occurs allowing the PBS in +sssDNA to anneal to its complement sequence within −sssDNA (94,95). Whereupon RT synthesizes plus strand DNA until it reaches the central termination signal (CTS) located 3´ to the cPPT (96). In addition, plus strand synthesis initiated from the cPPT continues to the end of the minus strand DNA template. Enzymes located in the host cell presumably repair gaps in the newly formed viral DNA due to removal of internal primers and strand displacement (6). Completion of reverse transcription results in a blunt-ended, double-stranded linear DNA with long terminal repeats (LTR) at each end.
Figure 1.3: The Reverse Transcription Process. (Black line) genomic RNA; (blue line) minus strand DNAs; (red line) plus strand DNAs. **Step 1:** minus strand DNA synthesis is initiated from the 3’ end of a host tRNA$_{Lys}$. Minus strand synthesis proceeds to the 5’ end of the genomic RNA, forming the minus strand strong-stop DNA (−sssDNA). **Step 2:** Following RNase H digestion of the RNA portion of the newly formed RNA:-sssDNA product, the first strand transfer causes −sssDNA to be annealed to the complementary R region located at the 3 end of the viral genomic RNA. **Step 3:** Minus strand DNA synthesis continues, accompanied by RNase H digestion of the template strand. **Step 4:** Synthesis of plus strand DNA is initiated from two RNase H resistant primers: a polypurine track (PPT) and a central PPT (cPPT). Plus strand DNA is elongated to the end of the minus strand DNA template to produce plus strand strong-stop DNA (+sssDNA). RNase H degradation of the primer tRNA$_{Lys}$ exposes the PBS sequence located in the +sssDNA. **Step 5:** A second strand transfer allows annealing of the exposed PBS segment to a complementary PBS region on the minus strand DNA. **Step 6:** Plus and minus strand DNA synthesis continues. A strand displacement by RT produces a double stranded linear molecule with long terminal repeats (LTRs).
Figure 1.3
1.2.2 Variability and Mutagenesis

HIV-1 exists as a mixture of nonidentical but closely related virions, known as “quasispecies” (97). This genetic diversity within the HIV-1 population manifests itself in the evolution of immune escape mutants, drug resistant mutants, and variants that have the ability to alter their tropism. In combination, these biological phenomena allow persistent viral replication that ultimately results in immunodeficiency. The high variability seen in HIV-1 can be traced back to the biology of the virus. HIV-1 must carry out reverse transcription in order to propagate itself. Inherent to the process of reverse transcription is the possibility of introducing genetic errors into the newly formed daughter progeny. Genetic errors and/or mutations can be introduced into the virus either by viral or cellular mechanisms. Viral mechanisms include the introduction of mutations by the error-prone RT and recombination between two genetically distinct RNAs. Cellular mechanisms include the DNA polymerases involved in proviral replication, RNA polymerase II, which is needed for RNA synthesis of the proviral genome, and DNA repair enzymes that may be involved in filling gaps (e.g. integration) or completing plus strand synthesis. Although the high error rate associated with replication and recombination form the basis for diversity, the diversity seen in patients is the result of the high rate of HIV-1 replication (98).

With numerous steps along the virus life cycle capable of introducing errors where does the high degree of variation seen in HIV-1 come from? The viral RT emerges as a possible determinant of the high diversity seen in HIV-1. HIV-1 RT exhibits poor fidelity when compared with the host DNA polymerases. The base substitution error rate for most RTs in vitro is on the order of $10^{-4}$ errors per base
incorporated, which translates into approximately one error per genome in a reverse transcription cycle (99-102; for reviews see 103-105). In comparison, cellular DNA polymerases (e.g. DNA polymerases δ, γ, and/or ε) have a misincorporation rate of $10^{-6}$ to $10^{-7}$ (106,107). Table 1.2 lists approximate fidelities of DNA polymerases from a wide range of sources. The values reported for HIV-1 RT fidelity support the initial work by Gojobori and Yokoyama that indicated that retroviral genomes evolve at a rate 1 million times higher than the genomes of their hosts (117). Studies performed to measure the fidelity of RT \textit{in vivo} indicated that the mutation rate of HIV-1 was lower than the values predicted by the \textit{in vitro} results. Using an assay that calculates the mutation rate for one round of retroviral replication, Mansky and Temin estimated that the \textit{in vivo} mutation rate for HIV-1 was approximately $3.4 \times 10^{-5}$, which is 20 times lower than the \textit{in vitro} mutation rate (118). Thus, RT fidelity may be enhanced \textit{in vivo}, possibly by viral proteins or host factors. However, the lower rate of mutations observed in the \textit{in vivo} studies compared to \textit{in vitro} studies has been disputed. It has been suggested that the difference between the \textit{in vivo} and \textit{in vitro} mutation rates was due to the RT employed (HIV-1\textsubscript{NL4-3} RT) in the \textit{in vivo} study (119). If the \textit{in vivo} forward mutation frequency is compared to the forward mutation frequency generated from \textit{in vitro} studies using the same RT (HIV-1\textsubscript{NL4-3} RT) the difference is only 2-fold (119). In a separate study measuring the \textit{in vivo} mutation rate, O’Neil at al (120) observed a rate of mutation that was 2.5 times greater than the rate observed by Mansky and Temin (118). They suggest that the technique used to score mutations as an alternative explanation for the discrepancy. In their study they used single-stranded conformation polymorphism (SSCP) to score their mutations where the \textit{in vivo} study used phenotypic scoring (118,120).
Under SSCP conditions ~80% of mutations are scored, whereas the phenotypic assay scores ~40% (121,122). The later two studies suggest that when all the experimental conditions are standardized there is only a modest increase in the \textit{in vitro} mutation rate compared to the \textit{in vivo} rate. In addition to the mutation rate per generation, the number of generations and fixation rate of mutation must be taken into consideration when attempting to calculate the evolution of a virus.

There is evidence suggesting that retroviral proteins can influence the fidelity of reverse transcription. Included are NC, deoxyuridine triphosphate nucleotidohydrolase (dUTPase), and the HIV-1 accessory proteins Vif and Vpr. NC promotes annealing of the tRNA$^{\text{Lys}}$ primer, melting of secondary structures located in the genomic RNA, and strand transfer reactions during reverse transcription (33-36). Although \textit{vif} is essential for HIV-1 replication, the influence of its gene product on fidelity is unknown. However, a recent report indicated that Vif might have a role in the regulation of efficient reverse transcription (123). In contrast to Vif, Vpr has been shown to have an impact on the mutation rate. Mutations in \textit{vpr} increased the HIV-1 mutation rate approximately threefold (124 and this is addressed in more detail in section 1.3.2). Non-primate lentiviruses (e.g. feline immunodeficiency virus (FIV) (125,126), caprine arthritis encephalitis virus (CAEV) (125,127), equine infectious anemia virus (EIAV)(125,128) encode and express a dUTPase gene as a Pol polyprotein precursor which is a structural component of the virion. Studies have shown that non-primate lentiviruses lacking a functional dUTPase have an increased mutation rate (129,130).
<table>
<thead>
<tr>
<th>Polymerase</th>
<th>Error rate&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Viral</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMV-RT</td>
<td>$10^{-5}$</td>
<td>101</td>
</tr>
<tr>
<td>MMLV-RT</td>
<td>$10^{-5}$</td>
<td>101</td>
</tr>
<tr>
<td>HIV-RT</td>
<td>$10^{-4}$</td>
<td>99,100</td>
</tr>
<tr>
<td><strong>Bacterial</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pol I</td>
<td>$10^{-3}$ to $10^{-4}$</td>
<td>108</td>
</tr>
<tr>
<td>Pol II</td>
<td>$10^{-4}$</td>
<td>109</td>
</tr>
<tr>
<td>Pol III</td>
<td>$10^{-4}$ to $10^{-5}$</td>
<td>110</td>
</tr>
<tr>
<td>Pol IV</td>
<td>$10^{-4}$</td>
<td>111,112</td>
</tr>
<tr>
<td>Pol V</td>
<td>$10^{-3}$</td>
<td>112,113</td>
</tr>
<tr>
<td><strong>Eucaryotic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pol α</td>
<td>$10^{-4}$</td>
<td>106</td>
</tr>
<tr>
<td>Pol β</td>
<td>$10^{-3}$</td>
<td>106</td>
</tr>
<tr>
<td>Pol δ</td>
<td>$10^{-4}$ to $10^{-5}$</td>
<td>114</td>
</tr>
<tr>
<td>Pol γ</td>
<td>$10^{-4}$</td>
<td>106</td>
</tr>
<tr>
<td>Pol η</td>
<td>$10^{-2}$</td>
<td>107</td>
</tr>
<tr>
<td>Pol ι</td>
<td>$10^0$</td>
<td>115</td>
</tr>
<tr>
<td>Pol κ</td>
<td>$10^{-2}$ to $10^{-3}$</td>
<td>116</td>
</tr>
</tbody>
</table>

<sup>a</sup>Values were generated from separate individual experiments that measured initial insertion fidelity without 3’ exonucleolytic activity. Since fidelity studies are dependent upon context, conditions and methods the values are only approximations.

Table 1.2: A Summary of the Fidelities Associated with Various DNA Polymerases.
Recombination is another characteristic associated with RT that can lead to the increased infidelity observed with this polymerase. Under certain conditions two genetically distinct RNAs can be co-packaged into a virion. If portions of each RNA are used to generate a single viral dsDNA molecule a recombinant is formed (131-133). Two models of retroviral recombination have been proposed, the forced copy choice (134) and the strand displacement/assimilation model (135). In the copy choice model, recombination occurs during synthesis of the first, or minus strand of DNA when a newly synthesized DNA from one viral RNA template is transferred to the second viral RNA. The strand displacement/assimilation model, recombination occurs during plus strand DNA synthesis. This model is predicated on the conversion of each copy of the viral RNA into DNA and that plus strand DNA synthesis is initiated at multiple sites. The strand displacement/assimilation model proposes that the 5´ ends of the plus strand DNA initiated from one polypurine tract primer are displaced by the 3´ ends of the plus strand DNA initiated from an upstream primer. The displaced product is then able to base pair with the other minus strand DNA in the virion.

1.2.3 Reverse Transcriptase

Multiple components influence the fidelity of an individual DNA polymerase and RT. Possible sources include: proofreading functions, nucleotide selectivity, and the ability of the polymerase to extend mismatched primer termini. RTs differ from cellular DNA polymerases in that they are devoid of 3´→5´ exonuclease proofreading activity (136). Without this editing function RT is not able to remove a misincorporated nucleotide. Numerous in vitro studies have been performed to address the
misincorporation frequency associated with various RTs; HIV-1 (137-139), human immunodeficiency virus type 2 (HIV-2) (140), avian myeloblastosis virus (141), EIAV (142), mouse mammary tumor virus (143). However, numerous parameters, such as: the type of polymerase, the nature of the template (e.g. DNA or RNA), the concentration of deoxyribonucleoside triphosphates (dNTPs), and the analyzed repair influence in vitro assays that measure the misincorporation frequency and make it difficult to compare polymerases.

In general the misincorporation efficiency of the HIV-1 RT is greater than other members of the Retroviridae family (99,139). Fidelity estimates are obtained by measuring the incorporation rate ($V_{\text{max}}$ or $k_{\text{cat}}$) of correct or incorrect dNTPs and their corresponding nucleotide binding affinities ($K_M$). The incorporation rate ($V_{\text{max}}/k_{\text{cat}}$ value) is expected to be higher for a correct nucleotide than for an incorrect one. On the other hand, incorrect nucleotides have a higher $K_M$ value than correct nucleotides, indicating that they bind the RT/template-primer complex with less affinity. The misincorporation frequency ($f_{\text{inc}}$) of a polymerase can be calculated by comparing the kinetic efficiencies when an incorrect nucleotide is used as a substrate relative to a correct nucleotide:

$$f_{\text{inc}} = \frac{k_{\text{cat}}/ K_M (\text{incorrect})}{k_{\text{cat}}/ K_M (\text{correct})}$$
Therefore, lower misincorporation efficiency is indicative of higher polymerase fidelity. The misincorporation efficiencies of retroviral RTs have been estimated to range from $10^{-3}$ to $10^{-6}$, depending on the experimental conditions used (e.g. type of mispair and template) (99, 137, 138, 140-144). The most common misincorporations have been reported to be G:T and G:U mismatches (141,145).

In addition to RT’s high propensity to insert the wrong nucleotide, RT is highly efficient at extending this newly formed mispair. This ability serves a major determinant of the infidelity associated with HIV-1 RT. Assays utilized to determine the frequencies of mispair extension indicated that HIV-1 RT had a 50-fold higher rate of extension than the mammalian replicative enzyme DNA polymerase α (146). The highest mispair extensions are seen with purine:pyrimidine or pyrimidine:purine mismatches (147).

Hypermutation, an accumulation of large blocks of a specific type of mutation, is characteristic of the HIV-1 mutation spectra. The majority of mutations isolated are base substitutions, with G-to-A transitions being the most common (148,149). These substitutions most commonly arise from –1 frameshifts resulting from template-primer slippage, which lead to runs of T’s or A’s in the HIV-1 genome (149,150). However, a study performed by Ji and Loeb indicated that the preponderance of G-to-A substitutions observed in vivo could not be attributed solely to the HIV-1 RT (151). Aberrant RNA editing (152,153) and low nucleotide pools (154) are additional factors implicated in causing the hypermutation seen in HIV-1.
1.3 Viral Protein R

1.3.1 Background

In addition to the structural genes gag, pol, and env, HIV-1 encodes a 14-kDa, 96 amino acid non-structural protein that accumulates in the nucleus of an infected cell (155-157). Viral protein R (Vpr) is expressed in primate lentiviruses (e.g. simian immunodeficiency virus (SIV) and HIV-1) (158). The ability to replicate in non-dividing cells is a hallmark of the primate lentiviruses (159). While Vpr is not essential for the replication of HIV-1 in cultured cell lines or activated T-cells, it is required for the efficient replication of the virus in quiescent cells such as macrophages (160-162). It is believed that Vpr interacts with the cellular nuclear transport machinery to facilitate nuclear import of the viral preintegration complex (PIC) (163-168). There exist differing reports on the amount of Vpr packaged into virus particles, ranging from ~50-1,200 particles/virus depending on the technique and system used (157,169,170). Packaging of Vpr into the virion requires an interaction with the p6 domain of the Gag precursor (Pr55Gag) (171,172). Numerous phenotypes have been attributed to Vpr, including influencing the viral mutation rate (124,173,174), regulation of apoptosis (175-178), and increasing viral transcription (175,179,180). It has been suggested that the increased viral transcription is due to Vpr’s ability to arrest a cell at the G2/M phase of the cell cycle (181-183). Vpr’s influence on HIV-1 mutation rate was originally demonstrated in a study using a model system that measures the in vivo mutation rate. In this study a vpr mutant increased the mutation rate associated with HIV-1 (124). In a separate study, vpr− HIV-1 shuttle vectors displayed a four-fold increase in the mutation rate compared to vpr+ HIV-1 vectors (173). The aforementioned phenotypes associated
with Vpr are also characteristic of HIV-2 and SIV. However, unlike HIV-1 the nuclear import and the cell cycle manipulation properties of these viruses are isolated on two separate proteins, Vpx and Vpr2/VprSIV (184,185).

### 1.3.2 Association with Cellular Proteins

While Vpr’s role(s) in some of these processes (e.g. nuclear localization, cell cycle arrest, influencing mutation rate) remains unclear, it may be related to Vpr’s ability to interact with and/or recruit different cellular proteins. Vpr is reported to interact with the transcriptional activators Sp1 (186) and TFIIB (187), and the human homologue of MOV34 (188), which is involved in transcriptional and cell cycle regulation. In addition Vpr has been shown to associate with the nuclear transport factors importin-α (165,168) and nucleoporin Pom21 (168) as well as the glucocorticoid receptor (189,190) and a cytosolic protein that associates with the glucocorticoid receptor (180). Lastly, Vpr has been reported to bind to two proteins that are involved with DNA repair processes, hHR23A (191,192) and uracil DNA glycosylase (UDG) (174,193,194). While the function of hHR23A has not been determined, it is the homologue of the yeast DNA repair protein Rad23 (192,195). hHR23A has been reported to associate with the xeroderma pigmentosum complementation group C (XPCC) protein (195,196) which functions in global nucleotide excision repair (197,198). It has also been reported that Vpr may mediate apoptosis through its interaction with hHR23A (199).
1.3.3 Association with Uracil DNA Glycosylase (UDG)

In a study attempting to identify cellular proteins that may contribute to Vpr’s role in previously demonstrated cellular functions, UDG was shown to associate with Vpr (193). The interaction between Vpr and UDG was initially identified using a yeast-2-hybrid screen (193). The yeast-2-hybrid results were confirmed both in vitro by protein-protein binding studies demonstrating that purified UDG interacts with Vpr expressed as a GST-Vpr fusion protein and ex vivo by co-immunoprecipitation experiments in HeLa cells transiently expressing Vpr (193). Subsequent analysis of the interaction between UDG and Vpr demonstrated that UDG enzymatic activity was not diminished nor was the interaction disrupted by the addition of a specific inhibitor of UDG (e.g. uracil DNA glycosylase inhibitor (UGI)) (193). However, UDG enzymatic activity was not detected when wildtype HIV-1 were analyzed (193). Using N- and C-terminal Vpr deletion mutants, it was determined that the core of Vpr between residues 15 and 77 was involved in the UDG-Vpr interaction. More rigorous investigation revealed two domains spanning amino acids 15 to 27 and 63 to 77, and a Trp residue at position 54 of Vpr critical for this association (193). The Vpr binding site within UDG was shown to contain a WXXF motif (200). This motif is found in other proteins shown to associate with Vpr, as well as many members of the conserved UDG family (Fig. 1.5) (200,201).

While the role of UDG in the HIV-1 replication cycle is unclear it has been proposed that the Vpr-UDG interaction influences the HIV-1 mutational frequency (174). Using a model system developed to measure the in vivo forward mutation rate of HIV-1, recombinant virus containing mutant vpr had a slightly greater than a three-fold increase in its mutation rate compared to recombinant virus containing wild-type vpr (174).
Figure 1.4: Locations of Vpr’s Structural Domains. Putative functions attributed to each region are shown. Mutagenesis experiments were utilized to ascertain the region(s) of Vpr critical for each proposed function. Adapted from Gaynor and Chen (187).
Further investigation determined that the packaging of the nuclear form of UDG into HIV-1 particles was required for the influence seen on the \textit{in vivo} mutation rate and the incorporation of UDG into particles was dependent on Vpr. However, the observation that Vpr increases the fidelity of HIV-1 RT by four-fold has been questioned (120). In a study that analyzed the mutation rates of two different HIV-1 strains, HXB2 and NL4-3, similar mutation rates were observed, $9.2 \times 10^{-5}$ and $7.9 \times 10^{-5}$, respectively (120). These results were unexpected because HXB2 does not express functional Vpr (202). Since it had previously been reported that Vpr modulated the HIV-1 the authors had expected a higher mutation rate in the HXB2 strain. The authors offer as a possible explanation a higher fidelity associated with the HBX2 RT compared to the NL4-3 RT (120).

There exists contradictory data concerning the mechanism by which HIV-1 incorporates UDG into virus particles. It has been reported that UDG incorporation was dependent upon its association with the HIV-1 encoded IN protein (203). Similarly to Vpr, this interaction was initially discovered in a yeast-2-hybrid screen. \textit{In vivo} experiments looking at Vpr mutants, IN mutants, Vpr/IN double mutants, and wild-type HIV-1 determined that the incorporation of UDG was dependent on the presence of IN (203). Furthermore, IN mediated the incorporation of the precursor form of UDG, which has little, if any enzymatic activity, whereas Vpr mediated the incorporation of the nuclear isoform of UDG (174,203).
1.4 Uracil DNA Glycosylase

1.4.1 Background

The structure and function of UDG has been characterized extensively from numerous organisms. Analysis has shown a high degree of sequence similarity between UDGs from yeast, *Escherichia coli* (*E. coli*), members of the *Herpesviridae* and *Poxviridae* families, as well as human (201,204-210), arguing for a strong evolutionary conservation of the gene. As one example, the enzymes from the *E. coli ung* gene and the human *ung* gene have 55.7 % identity at the amino acid level and 73.3 % similarity when conservative amino acid substitutions are taken into consideration. UDG, which is a member of the base excision repair (BER) complex, is responsible for initiating the removal of uracil residues found in the DNA (reviewed in 211 and 212 and addressed in more detail in section 1.6). The human *ung* encodes a mitochondrial (UDG1) and a nuclear (UDG2) enzyme that are produced by alternative transcription start sites and alternative splicing (213-215). UDG1 (Mr: 24 kDa) and UDG2 (Mr: 35 kDa) have unique N-terminal amino acid sequences that are required for mitochondria and nuclear import respectively. The N-terminal region of UDG2 also contains a motif for binding of proliferating cell nuclear antigen (PCNA), and two different motifs that bind replication protein A (RPA). The two enzymes share the same C-terminal 269 amino acids. Within this common sequence is the catalytic domain that contains a conserved DNA binding groove and a tight-fitting uracil-binding pocket. Mammalian UDGs have been shown to remove uracil residues from single-stranded and double-stranded DNA substrate (211,216,217). The rate of uracil hydrolysis occurs almost twice as fast on single-stranded versus double-stranded DNA (217-220), and both U:A basepairs and U:G
mismatched bases are recognized (217,221). The rates of removal of uracil from U:G mismatches are usually higher, however the sequence context surrounding the uracil residue is a major determinant of this rate (217,221). In some sequence contexts the rate of uracil removal is faster in U:A basepairs (217,221). Consensus sequences for good and poor removal are 5′-(A/T)UA(A/T) and 5′-(G/C)U(T/G/C), respectively (222,223). Uracil is not excised from dUMP, dUTP, RNA, or RNA contained in a RNA-DNA hybrid; however, it has been reported that a UDG from calf thymocyte nuclei was able to excise uracil residues from the DNA strand of a RNA-DNA hybrid (224).

UDG activity is cell-cycle dependent and is increased in response to cell proliferation (225-227). UDG mRNA levels increase 8-12-fold during late G1 phase, and enzyme activity increases prior to S phase and reaches a maximum early in S phase (228,229). The enzyme has a half-life of approximately 30 hours (229). UDG is an efficient enzyme with a turnover number of 500-1000 per minute, depending on the reaction conditions (230). Similar to all other known DNA glycosylases, UDG retains enzymatic activity in the presence of EDTA and does not require divalent cations for activity. However, human UNG2 activity is stimulated by Mg$^{2+}$, whereas the core catalytic domain is not. Uracil acts as a noncompetitive inhibitor of UDG whereas apurinic/apyrimidinic (AP) sites inhibit competitively (220).

1.4.2 Additional Enzymes with Uracil DNA Glycosylase Activity

In addition to the mitochondrial (UDG1) and nuclear (UDG2) isoforms of UDG, there are three other enzymes in human cells capable of removing uracil residues from the DNA. Table 1.3 summarizes mammalian enzymes capable of releasing uracil from
Figure 1.5: Amino Acid Alignment of the Uracil DNA Glycosylases. The organisms are listed with GenBank accession numbers. VZV, varicella zoster virus, (X04370); *E. coli*, (J03725); *S. cerevisiae*, *Saccharomyces cerevisiae*, (J04470); human, UDG2, (X15653). (■) Residues critical for cleaving the glycosylic bond; (■) residues important for the exclusion of uracil bound to ribose in RNA; (■) most critical residue for specific recognition of uracil in DNA; (■) motif demonstrated to be important for Vpr-UDG interaction; (■) conserved leucine loop, critical for UGI binding. The leucine residue in this motif is involved in uracil base flipping, whereas the histidine residue is involved in cleavage of the glycosylic bond.
Figure 1.5
DNA. Thymine DNA glycosylase (TDG), a mismatch repair DNA glycosylase, specifically removes uracil from U:G mismatches or thymine residues from G:T mismatches (231-234). Neither uracil nor thymine residues in single-stranded DNA nor U:A mispairs are substrates for TDG. MBD4 is a DNA glycosylase that contains a methyl-CpG binding domain and can remove thymine and uracil from a CpG mismatch (235-237). This enzyme fails to remove uracil residues in single-stranded DNA (238). Lastly, a single-stranded monofunctional uracil DNA glycosylase (SMUG1) has been recently identified and shown to preferentially excise uracil residues from single-stranded DNA (238-240). Subsequent analysis revealed that SMUG1 is the previously characterized 5-hydroxymethyluracil DNA N-glycosylase (HMUDG) (241). It has been postulated that SMUG1 is important during replication and transcription when single-stranded regions of DNA (240), which are more susceptible to cytosine deamination than double-stranded DNA (242).

<table>
<thead>
<tr>
<th>Name</th>
<th>Size (aa)</th>
<th>Cellular location</th>
<th>Known substrates</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDG1</td>
<td>304</td>
<td>Mitochondria</td>
<td>ssU &gt; U:G &gt; U:A &gt; 5-FU</td>
</tr>
<tr>
<td>UDG2</td>
<td>313</td>
<td>Nucleus</td>
<td>ssU &gt; U:G &gt; U:A &gt; 5-FU</td>
</tr>
<tr>
<td>TDG</td>
<td>410</td>
<td>Nucleus</td>
<td>U:G &gt; T:G</td>
</tr>
<tr>
<td>MBD4</td>
<td>580</td>
<td>—</td>
<td>U:G, T:G; U/TpG:m5CpG</td>
</tr>
<tr>
<td>SMUG1</td>
<td>270</td>
<td>—</td>
<td>ssU &gt; U:A, U:G</td>
</tr>
</tbody>
</table>

Table 1.3: Human DNA Glycosylases that Can Excise Uracil.
1.5 Uracil Formation in DNA

1.5.1 dNTP Pools and dUTP Mis-incorporation

Uracil incorporation into DNA has been demonstrated in a number of organisms. Incorporation occurs because DNA polymerases have the ability to utilize dUTP as a substrate for DNA synthesis in place of dTTP (243). The replacement of dTMP by dUMP results in U:A rather than T:A basepairs. However, this substitution has the same coding capacity, therefore, a U:A basepair is not mutagenic. The amount of uracil introduced into DNA is dependent on the intracellular pool levels of dUTP relative to that of dTTP. The levels of available dUTP are limited because of the activity of deoxyuridine triphosphate nucleotidohydrolase (dUTPase), which catalyzes the conversion of dUTP to dUMP and pyrophosphate (244). This hydrolysis provides the cell with a valuable source of dUMP for de novo dTTP biosynthesis (Fig. 1.6). It has been estimated that in human lymphoid cells, the concentration of dUTP is less than 0.3 fmol/10^6 cells, whereas the concentration of dTTP is 40 pmol/10^6 cells (245). These values are indicative of this cell type and probably do not represent the levels in all type of cells. Under normal physiological conditions approximately 2000 uracil residues are incorporated into human DNA per cell per replication cycle (246,247).

1.5.2 Cytosine Deamination

Uracil may also be introduced into DNA by deamination of dCMP residues. The deamination of cytosine results in a G:U mismatch that, unless repaired before the next round of replication, will result in a G:C to A:T transition mutation. It has been estimated that approximately 400 cytosine residues are deaminated per day in a diploid cell.
Studies investigating the spectrum of spontaneous mutations in *E. coli* and in Chinese hamster ovary cells disclosed that G:C to A:T transition mutations were responsible for 62% and 77% of the total mutations respectively (250,251). The preponderance of this type of transition mutation is indicative of the enhanced cytosine deamination in single-stranded DNA (250).

### 1.5.3 Evidence of Uracil Formation in HIV-1 DNA

Analysis of the mutation spectrum associated with HIV-1 shows a preponderance of base substitutions, the majority of which are G to A and C to T transition mutations (118). These types of mutations are characteristic of cytosine deamination of dCMP residues in the viral DNA and then failure to accurately remove and repair this damage by a DNA glycosylase (e.g. UDG). If cytosine deamination occurred in the minus strand DNA during reverse transcription then a G-to-A transition mutation would be present in the plus strand genomic viral RNA. On the other hand, if cytosine deamination occurred in the plus strand DNA during reverse transcription, then the likelihood of a mutation occurring in the plus strand genomic viral RNA is remote. Alternatively, a G-to-A transition mutation could occur if uracil was incorporated opposite a guanine during minus strand synthesis. The probability of this event occurring is very unlikely, but is possible if the enol forms of uracil or guanine are present. Under these conditions guanine is equivalent to adenine and uracil is equivalent to cytosine. However, the enol forms of these bases occur infrequently (<0.01%) under physiological conditions (252).
Figure 1.6: Mammalian Pyrimidine Deoxyribonucleotide Metabolism. Enzymes are represented by italicized print. *de novo* dTTP biosynthesis utilizes the dUMP intermediate. Alternatively, thymine can be scavenged to synthesis dTTP in a “salvage pathway”. The major source of dUMP is generated from dCMP deamination. dTTP acts as an allosteric regulator of ribonucleotide reductase.
Figure 1.6
It appears that members of the lentivirus genus of the \textit{Retroviridae} family have evolved mechanisms to address the potential problem of uracil residues in viral DNA. Non-primate lentiviruses encode and package dUTPase into viral particles, whereas primate lentiviruses incorporate the cellular enzyme UDG. Interestingly enough, primate and non-primate lentiviruses lacking UDG and dUTPase respectively, exhibited similar increases in G to A transition mutations (124,130). Thus, it appears that lentiviruses have evolved different mechanisms to secure the integrity of its genetic material.

1.6 Base Excision Repair

1.6.1 Role

Mammalian cells are constantly exposed to a variety of endogenous and environmental DNA-damaging agents. In order to maintain genomic integrity cells have evolved DNA repair mechanisms to excise and replace the damaged nucleotides. These lesions may be caused by a variety of agents, such as spontaneous deamination of bases, radiation, oxidative stress, alkylating agents or replication errors. Damaged bases may be miscoding, cytotoxic or both. The most frequently occurring lesions are repaired by BER. BER mainly repairs small lesions in DNA that do not strongly distort the structure of the DNA backbone. Usually this repair involves removal of a single nucleotide from one strand of the DNA. In contrast nucleotide excision repair removes bulky adducts from DNA and usually results in the excision of an oligonucleotide consisting of 12-30 residues. The components of the BER pathway have, for the most part, precise lesion specificity, but lack versatility.
1.6.2 Components of the BER Pathway

Five enzymatic activities are required for functional BER (Fig. 1.7A). Initiation of BER involves the cleaving of the N-glycosyl bond linking the base to the deoxyribose by a DNA glycosylase. When a uracil residue is formed in the DNA, UDG2 is the DNA glycosylase predominantly used. Removal of uracil results in the formation of an apyrimidinic (AP) site (253). This resulting AP site is highly cytotoxic and mutagenic, and must be further processed by a 5`-AP endonuclease (APEX) (254-256). UDG remains bound to the AP site until it is displaced by APEX, thus reducing the likelihood that these sites will be cytotoxic or mutagenic (257,258). APEX cleaves the sugar-phosphate backbone 5` to the apyrimidinic site leaving a free 3`-hydroxyl terminus capable of priming DNA repair synthesis and a deoxyribose phosphate (dRP) group at the 5` terminus. Excision of the dRP group is catalyzed by the 2-deoxyribose-5-phosphate lyase activity of polymerase β (β-pol) (259). Completion of BER involves filling in the gap by β-pol and sealing the nick with either DNA ligase I or III (260).

1.6.3 Long-Patch BER

In addition to the aforementioned “short-patch” BER pathway involving replacement of one nucleotide, there exists a “long-patch” BER pathway involving the gap-filling of several nucleotides. Exposure of DNA to oxidative agents and UV radiation can cause the formation of AP site adducts that are refractory to the 2-deoxyribose-5-phosphate lyase activity of β-pol (261). These adducts, which are not substrates for the lyase activity of β-pol can be repaired through the long-patch BER pathway.
A model for the reaction mechanism for long-patch BER repair consists of four enzymatic reactions and is shown in Fig. 1.7B. Similar to short-patch BER the AP site is cleaved at its 5′ side by APEX. A multiprotein complex consisting of replication factor C (RF-C), proliferating cell nuclear antigen (PCNA) and a DNA polymerase δ (or ε) is then recruited to the incised AP site where DNA synthesis occurs creating a 5′ flap (262,263). The displaced flap containing the dRP residue is recognized and removed by flap Endonuclease I (FEN-1) (262-264). The resulting nick is sealed by a DNA ligase.

1.7 Uracil DNA Glycosylase Inhibitor

The *Bacillus subtilis* bacteriophage PBS1 and PBS2 are unique organisms since their double-stranded DNA genome contains uracil in place of thymine (265,266). The genome of PBS2 is also unique in that it has an overall base composition of 28% G+C and 72% A+U (265). To overcome the obstacle of replicating in an environment non-conducive to uracil formation in DNA, PBS2 has evolved mechanisms to ensure the incorporation and retention of uracil in its genome. One of these mechanisms is the expression of a UDG inhibitor protein (UGI) (267,268). Inhibition of UDG prevents the excision of uracil residues in the PBS2 genome once they have been incorporated.

UGI is a small 84 amino acid protein (Mr: 9.5 kDa) that is acidic and stable to heat and 8M urea (269). The interaction between UGI and UDG occurs in a 1:1 stoichiometry that does not dissociate under physiological conditions (269). Because of the high degree of conservation amongst the UDG family of enzymes, UGI is able to inhibit UDGs from varying biological organisms, including *E. coli*, *S. cerevisiae*, human, and herpes simplex virus type 1 (HSV-1) (201,270-276). However, TDG, MBD4, and
Figure 1.7: Pathways for Base Excision Repair in Mammalian Cells. (A) Short-patch BER. A damaged base (e.g. uracil) is excised by UDG to generate an AP site. APEX displaces UDG from the AP site and hydrolyzes the phosphodiester bond immediately 5´ to the AP site. The dRPase activity of β-Pol excises the 5´ deoxyribose phosphate group creating a single-nucleotide gap. β-Pol fills this gap and the nick is sealed by a DNA ligase. (B) Long-patch BER. Similar to short-patch BER the damaged base is excised and the resulting AP site is recognized and cleaved at its 5´ side by APEX. A repair complex consisting of RF-C, PCNA and Pol δ/ε forms at the free 3´-OH. Subsequent displacement synthesis displaces the 5´ deoxyribose phosphate group generating a 5´ flap. FEN-1 removes the 5´ flap strand. The resulting nick is sealed by DNA ligase.
Figure 1.7A
Figure 1.7B continued

5'--p p p p p p p p--3'

UDG

A G U T G A

5'--p p p p p p p p--3'

APEX

A G OH T G A

5'--p p p p p p p p--3'

PCNA, RF-C, Pol δ/ε

A G OH OH T G A

5'--p p p p p p p p--3'

FEN-1 + ligase

A G C T G A

5'--p p p p p OH p--3'

45
SMUG1, which have the ability to excise uracil residues from DNA, are insensitive to UGI activity (235,236,239,277). The structure of UGI and the UGI-UDG complex has been demonstrated with several UDGs (273-275,278-280). UGI inhibits UDGs by acting as a transition-state DNA mimic (278-280). The major interactions between UGI and UDG occur between eight hydrophobic residues of UGI and the conserved leucine loop (HPSPLS), residues 187 to 192 and 268 to 273 of the *E. coli* and human UDG2, respectively (280). The conserved leucine residue in the UDG motif is involved in base flipping and the binding of UGI effectively inhibits the UDG activity (280,281). Thus, under physiological conditions, the association of UGI with UDG prevents UDG from binding to DNA. UGI also can dissociate UDG bound to DNA (270).

1.8 Importance of UDG

It has been very difficult to fully understand the importance and/or precise function of UDG in mammalian cells. These obstacles are associated with the inability to isolate or construct cells deficient in UDG. Krokan et al (282) reported analyzing over 100 different cell lines for mutations in the *udg* gene without finding any mutations that abolished or even reduced UDG activity significantly. It is inferred from the inability to isolate UDG mutants and the high degree of evolutionary conservation among *udg* sequences from varying biological organisms that UDG activity is vital to a cell. However, UDG mutants have been engineered in *E. coli, S. cerevisiae* and mouse cells and these cells/organisms grew normally suggesting that the *udg* gene is nonessential (239,283-287). Further analysis demonstrated an approximately 30-fold and 50-fold increase in the G:C to A:T transition mutation frequency in *E. coli* and *S. cerevisiae*,
respectively when compared to their wild-type UDG isogenic strains (284,287). Since there was little change in the frequencies of other mutations, the results strongly suggested that UDG was responsible for the excision of uracil from DNA. In contrast to the \textit{udg} mutant bacteria and yeast, there was only a modest increase in the G:C to A:T transition mutation frequency in the \textit{udg} knockout mice (239). The low frequency of G:C to A:T mutations found in mice may have been due to other enzymes capable of excising uracil. Ultimately, it was determined that the mice cells contained SMUG1, an enzyme shown to be capable of excising uracil (238-240).

An alternative approach to creating UDG deficient cells is by constitutively expressing UGI in human cells (288). This approach was initially utilized in an attempt to define the role of UDG in varicella zoster virus replication. A recent study using human glioma cells employing this approach failed to detect UDG activity in transfected cells but demonstrated that the cells were viable in the presence of UGI (289), further confirming the nonessential nature of UDG. Similar to previous data generated on bacterial and yeast \textit{udg} mutants, the glioma cells exhibited an increased frequency of G:C to A:T base pair substitution mutations (284,287,289). Cell extracts were analyzed for the presence of TDG and MBD4, enzymes capable of excising uracil from DNA but insensitive to UGI. TDG was not detected while MBD4 was (289).

In the study that constructed glioma cell lines over-expressing UGI, the effect of an UDG deficiency on the long-term viability of the host cells was not addressed (289). Previous data generated from UGI overexpression in a transgenic Chinese hamster ovary cell line indicated that the cell doubling time and the relative cloning efficiency was
decreased (Dr. M. V. Williams, The Ohio state University, unpublished results). Thus, there exists a controversy over the cell viability in the presence of overexpressed UGI.

Even with these recent studies there still exists uncertainties concerning the function of UDG. Little is known about the role of UDG in DNA replication or of its association with other cellular proteins. It is inferred from previous studies that UDG is responsible for preventing uracil accumulation in DNA because of the increase in G:C to A:T transition mutations. However, the effect of continual accumulation of uracil in the DNA has not been closely analyzed in mammalian cells. The researchers that performed the studies in mammalian cells lacking UDG activity (e.g. knockout mice and glioma cells expressing UGI) offer a reason for the nonessential nature of UDG as being the redundant mechanisms of uracil excision in mammalian cells (239,289). Nevertheless, in the case of the knockout mouse study, it was reported that approximately 2000 uracil residues were introduced into the DNA (239) and in the study constructing the glioma cell line the mutation frequency was similar to \textit{E. coli} and \textit{S. cerevisiae} UDG mutants (288). These results indicate that although other enzymes exist in cells that remove uracil from DNA, there still is an introduction and retention of uracil into the DNA. As alluded to earlier, these two studies did not report on the long-term affects of UDG deficiency on cell viability.

With respect to HIV-1 and the incorporation of UDG2 into the virion, other than the report demonstrating an increase in the forward mutation frequency there has not been any work indicating the long-term viability of UDG2 deficient HIV-1. There are many complications when trying to address the role of UDG in the HIV-1 life-cycle and comparisons to UDG involvement in mammalian cells are equally as difficult. First of
all, polypeptides encoded by HIV-1 have only been shown to interact with UDG2 or its precursor. There are no reports indicating association with any of the other human enzymes capable of excising uracil or reports demonstrating HIV-1 association with the enzymes shown to be critical for completion of BER. These enzymes are usually located in the cell’s nucleus. Therefore, if the nuclear environment contains multiple enzymes with uracil excising ability, including UDG2, as well as the enzymes necessary for BER, UDG2 must play a vital role in HIV-1 replication within the cytoplasmic core prior to nuclear localization.

1.9 Deoxyuridine Triphosphate Nucleotidohydrolase

Under normal physiological conditions, there is no significant incorporation of dUTP into DNA, even though DNA polymerases and reverse transcriptases can utilize dUTP as a substrate. The lack of significant incorporation of dUTP into DNA is due to the action of deoxyuridine triphosphate nucleotidohydrolase (dUTPase), which cleaves the α-β phosphodiester bond of dUTP to form dUMP and pyrophosphate, thus providing the cell with dUMP, which is required for TMP synthesis. With the exception of some members of the cell wall-less Mollicutes (290), all free-living unicellular and multicellular organisms, that have been examined, possess dUTPase activity. Replicating cells contain high levels of dUTPase activity when compared to terminally differentiated and/or non-dividing cells (291-295). dUTPase activity is cell cycle regulated, as well as developmentally regulated (291-295). Attempts to construct null dUTPase mutants in Escherichia coli (296) and Saccharomyces cerevisiae (297) have not been successful nor have any higher eukaryotic cells been isolated or constructed that are deficient in this
enzyme. These results suggest that dUTPase may play an essential role in maintaining cellular stability. Some eukaryotic viruses, such as herpesviruses (298), poxviruses (299), and a subset of retroviruses also encode dUTPases (125-130). The viral encoded dUTPases are suggested to be essential for efficient viral replication in vivo and dUTPase deficient mutants are attenuated for virulence (127-130,300).

The human dUTPase gene encodes for distinct mitochondrial (DUT-M) and nuclear (DUT-N) isoforms that arise through alternate use of the first two exons of the structural gene (301). The DUT-M open reading frame encodes a protein of 252 amino acids with a predicted molecular weight of 26,704, while the DUT-N open reading frame encodes a protein of 164 amino acids with a predicted molecular weight of 17,748. Expression of DUT-M is constitutive and independent of the cell cycle phase or proliferation status of the cell (reviewed in 302). In contrast, DUT-N expression is cell-cycle regulated, with increased expression coinciding with DNA replication (302).

All known dUTPases can be classified based upon their subunit composition into three classes: monomeric, homodimeric or homotrimeric. The monomeric dUTPases are encoded by mammalian herpesviruses. The homotrimeric form is found in most prokaryotes, eukaryotes, and a large majority of viruses. The dimeric dUTPases were initially identified in the parasites Trypanosoma cruzi and Leishmania major (303,304), but since then a gene encoding a homologous protein was found in Campylobacter jejuni (305). Analysis of the amino acid sequences of the dUTPases revealed five conserved motifs (domains) within the monomeric and homotrimeric but not the homodimeric dUTPases. However, the order of these motifs is different in the homotrimeric dUTPases (1-2-3-4-5) and monomeric dUTPases (3-1-2-4-5). In the homotrimeric dUTPases the
five conserved motifs in each monomer contribute to the formation of three identical active sites. To form the active site motifs 1, 2 and 4 of one monomer, motif 3 of a second monomer, and motif 5 of the third monomer are contributed. Motif 3 has been shown to be critical for catalytic activity (reviewed in 306), whereas motif 5 has been suggested to be involved in recognition/binding of dUTP (306). The active site in the monomeric dUTPases is formed when motif 3 contributed by the N-terminal half of the protein associates with motifs 1, 2 and 4 from the C-terminal half (306). Completion of the active site occurs when motif 5 located on the C-terminal end is contributed.

1.10 **Significance of Research**

The genetic blueprint for life is found in the DNA of all eukaryotes and most other organisms. Unfortunately, endogenous and exogenous agents continuously challenge the structural integrity of DNA. A common form of DNA damage involves the modification and/or replacing the bases that make up DNA. These processes adversely affect the genetic integrity of organisms and can result in miscoding, cytotoxicity or both. To counteract these threats, cells have evolved defense mechanisms to prevent or repair base damage and/or eliminate the cell if the damage is too severe.

A major perturbation to DNA that has the potential of becoming deleterious is the formation of the RNA base uracil in DNA. Uracil can inappropriately arise in DNA via two processes: (i) the misincorporation of deoxyuridine monophosphate (dUMP) into DNA by DNA polymerases/reverse transcriptases and (ii) the spontaneous deamination of dCMP in DNA. Uracil is able to base pair with adenine and mispair with guanine, the latter process being mutagenic. Likewise, deamination is one of the major pro-mutagenic
events in DNA. The potential mutagenic and ultimately lethal nature of these lesions demonstrates the fundamental problem of preventing uracil accumulation within DNA that organisms are faced with.

As alluded to above organisms have evolved mechanisms to prevent the incorporation and retention of uracil in DNA. These mechanisms are highly conserved among organisms from varying biological sources. As discussed earlier, two enzymes in particular, deoxyuridine triphosphate nucleotidohydrolase (dUTPase) and uracil DNA glycosylase (UDG), are critical to ensure genetic integrity. In normal repair-proficient mammalian cells ~1 uracil residue can be found in the genome, thus demonstrating the efficiency of both dUTPase and UDG (260,307,308).

While the role(s) of dUTPase and UDG in normal cellular metabolism are fairly well defined, their function(s) in virus replication are unclear. Likewise, while several studies have demonstrated that both dUTPase and UDG may have critical roles in regulating the chemotherapeutic efficiency of specific antiviral and cancer therapeutic agents, there are presently no agents available clinically that specifically target these enzymes. This latter feature is the focus of my research. Can specific novel approaches be developed which can modulate UDG and dUTPase activity? To this end, I have developed a fusion protein (TaT-UGI) that can be introduced into cells by the process of protein transduction. Furthermore, using small interfering RNA technology I have developed a system for down-regulating dUTPase activity. I believe that these reagents will allow me to determine the roles of dUTPase and UDG in normal cellular metabolism as well as virus replication. Moreover, I believe the use of these reagents constitute a novel and effective therapeutic approach for modulating dUTPase and UDG. The studies
described in this manuscript detail the experiments undertaken and the results generated concerning the role(s) of these enzymes and the efficacy and feasibility of targeting dUTPase and UDG.
CHAPTER 2

PURIFICATION AND FUNCTIONAL PROPERTIES OF A NATIVE COMPLEX BETWEEN URACIL DNA GLYCOSYLASE AND HIV-1 Vpr

2.1 Introduction

Uracil-DNA glycosylase (UDG) is the major enzyme responsible for preventing the retention of potentially detrimental uracil residues that occur in DNA (217,282). Uracil residues are formed in DNA either by the misincorporation of dUTP by DNA polymerases or by the deamination of dCMP residues. Excision of a uracil residue by UNG initiates a base excision repair system to repair the resulting abasic site (282). UDG is expressed from a highly conserved family of genes that have been identified in a wide variety of organisms (282). The human ung gene located on chromosome 12q24.1 encodes for mitochondrial (UDG1) and nuclear (UDG2) isoforms of UDG through the use of different promoters and alternative splicing (215,227,309). UDG2 interacts with proliferating cell nuclear antigen (PCNA) and with replication factor A (RPA) and is localized within replication foci (310). Thus, it has been proposed that UDG2 is the primary enzyme responsible for initiating the removal of uracil residues from replicating DNA (217, 282). In addition to its role in post-replicative DNA repair processes, UDG has been demonstrated to be involved with immunoglobulin class switching.
recombination and somatic hypermutation (311-314). This process involves the
deamination of dC residues in the immunoglobulin variable genes by the activation-
induced cytidine deaminase followed by the excision of the uracil residue by UDG2.

Several poxviruses and herpesviruses also encode for UDGs. While these viral-
encoded UNGs are members of the highly conserved UNG gene family (282), their
role(s) in virus replication are not completely understood. Herpes simplex virus type 1
(HSV-1) UDG is not essential for replication in tissue culture, but it is required for
efficient replication, establishment of latency and reactivation in a mouse model
(315,316). Likewise, either host or virus-encoded UDG is required for the transition
from early to late phase replication in cytomegalovirus (CMV) (317). Conversely, the
replication of UDG deficient mutants of varicella-zoster virus (VZV) is not impaired in
cells depleted of cellular UDG, suggesting that UDG is not required for replication of
VZV in tissue culture (288). In the case of members of the poxviridae, the virus encoded
UDG is essential for virus replication (318). A recent study demonstrated that point
mutations in the catalytic site of the vaccinia encoded UDG allowed replication of the
virus in a variety of cultured cells, whereas UDG deletion mutants could not replicate
under the same conditions (319). However, these catalytic site mutants exhibited
decreased virulence in vivo. These data support the premise that viral UDGs may not be
functioning in a classical DNA repair process (320).

Studies on human immunodeficiency virus type 1 (HIV-1) and simian
immunodeficiency virus (SIV) have demonstrated that these lentiviruses incorporate the
nuclear isoform of human UDG (UDG2) into the virion through a Vpr and/or integrase-
mediated interaction (174,193,194,203,321-323). While these studies reported that the
virion associated UDG is functional, there has not been any studies to demonstrate whether the interaction of UDG with Vpr or integrase modifies UDG’s functional activity. Furthermore, structural and functional studies on Vpr have been limited because of the inability to obtain adequate quantities of the purified protein using either classical biochemical or recombinant DNA approaches. This is due in part to the relative toxicity of Vpr when expressed in eukaryotic cells, its instability and tendency to spontaneously aggregate at high concentrations (324-326).

The purpose of the present study was to determine whether native Vpr could be purified based upon its interaction with UDG and whether the complexing of Vpr with UDG altered the structural and/or functional activities of UDG. To accomplish these studies, a novel system was developed for the purification of Vpr complexed with the *Escherichia coli* UDG (UDGcoli). Using the Vpr-UDGcoli complex, it was possible to demonstrate that the interaction of Vpr with UDG did not alter the biochemical or functional properties of UDG. The ability of Vpr to interact directly with *E. coli* UDG was confirmed using a yeast two-hybrid assay. Data is also presented demonstrating that the *E. coli* UDG can be efficiently incorporated into HIV-1 virions.

### 2.2 Materials and Methods

#### 2.2.1 Materials

Unless otherwise indicated all chemicals and chromatography matrixes were obtained from Sigma Chemical Company (St. Louis, MO). All chemicals were of the highest purity available. [5-3H]dUTP was purchased from Moravek Biochemicals Inc., Brea, CA. Centriprep-30 concentrators were obtained from Amicon (Beverly, MA).
2.2.2 Bacterial Strains and Plasmids

The strains of *E. coli* used in this study include: BL21 (DE3) pLysS, CJ236
(*ung1, dut 1*) and Top 10. The HIV-1 encoded protein Vpr was expressed as a fusion
protein containing a N-terminal His(6)-tag using a pET16b-vpr construct kindly provided
by Dr. L. M. Mansky, The Ohio State University, Columbus, Ohio. The pTrc99C
construct containing the *E. coli* UDG gene was kindly provided by Dr. U. Varshney,
Indian Institute of Science, Banglore, India. The pGEX-2T constructs containing either
the varicella-zoster virus (VZV) encoded UDG (pGEX-VZVUNG) or the bacteriophage
PBS2 gene encoding the uracil-DNA glycosylase inhibitor protein (UGI) (pGEX-UGI)
have been described previously (288). pTrchisUGI was constructed by cloning a PCR
generated fragment containing the UGI gene into *BamH1- EcoRI* digested pTrchisA
(Invitrogen, Carlsbad, CA) and subsequently transforming the ligated DNA into Top 10.
CJ236 containing pGEX-VZVUNG, pGEX-UGI, or pTrchisUGI was grown and
maintained in 2XYT medium containing ampicillin (50 mg/ml) and chloroamphenicol
(25 mg/ml), while Top 10 containing pTrchisUGI was grown and maintained in LB
medium containing ampicillin (50 mg/ml). For the induction of the VZV encoded
UDG, UGI and Vpr, exponentially growing cells were treated with 1 mM isopropylthio-
β-D-galactoside (IPTG) for 2 hr at 37°C.

2.2.3 UDG and UGI Assays

UDG activity was determined as previously described (271,288). Briefly, the
reaction mixture contained the following in a total volume of 0.2 ml: 50 mM Tris-HCl
(pH 7.5), 100 µg of bovine serum albumin per ml, 3-5 µg of calf thymus double-stranded
DNA containing [3H]uracil residues (specific activity 99µCi/µmole, 40 ρmol of dUMP/µg DNA), 10 mM EDTA and the enzyme sample. After incubation at 37°C, the reactions were terminated and the acid-soluble radioactivity was determined as described previously (271,288). A unit of UDG activity was defined as the amount of enzyme required to release 1 ρmol of uracil as trichloroacetic acid soluble material per minute at 37°C.

UGI activity was measured as previously described (271,288) using the UDG assay described above. A unit of UGI was defined as the amount of inhibitor required to neutralize a unit of E. coli BL21 UDG under our assay conditions.

2.2.4 Protein Assays

Protein was estimated using the Coomassie blue dye binding assay as described by Bio-Rad Laboratories using bovine serum albumin (BSA) as the standard.

2.2.5 Vpr-UDGcoli Purification

BL21 cells transduced with pET16b-νpr were routinely grown and maintained in LB medium containing ampicillin (50 mg/ml) and chloroamphenicol (25 mg/ml). For the induction of Vpr, exponentially growing cells were treated with 1 mM IPTG for 2 hrs at 37°C. Cells were collected by centrifugation at 4°C using a Sorvall RC-5B centrifuge (5000 x g for 10 min). The cell pellets were suspended in extraction buffer containing 10 mM Tris-HCl (pH 8.0), 20 % (v/v) glycerol, 2 mM β-mercaptoethanol, 1 mM MgCl₂ and a protease inhibitor cocktail containing antipain, benzamidine, chymostatin, leupeptin, peptstatin, TLCK, and TPCK (final concentration 1 µg/ml). Cells were lysed by
sonication (Branson sonifier Model 350, microtip setting 4) for 1 min in a series of 15-sec pulses separated by 30-sec cooling periods on ice. The resulting homogenate was clarified by centrifugation at 5000 x g for 10 min 4°C. The resulting supernatant (designated crude extract) was used immediately for purification of UDG. Unless otherwise stated all remaining purification procedures were performed at 4°C.

The crude extract was applied to a Blue Sepharose affinity column (1.2 x 9 cm) equilibrated in 10 mM Tris-HCl (pH 7.5) (equilibration buffer). The column was washed with 5 bed volumes of equilibration buffer and UDG was eluted by increasing the ionic strength using a linear gradient of NaCl from 0.05 to 2 M in equilibration buffer. Fractions of 10 mls were collected and those fractions containing UNG activity were pooled.

The pooled fraction following Blue Sepharose chromatography was subjected to a series of concentration steps using a Centriprep-30 concentrator (Amicon; Beverly, MA). The concentration consisted of a series of centrifugations: (i) 20 minutes at 5000 x g, (ii) 15 minutes at 5000 x g, (iii) 10 minutes at 5000 x g. After each centrifugation step the filtrate, which corresponds to proteins and/or complexes with a molecular weight of less than 30,000 was collected. Likewise, the retentate which represents proteins having a molecular weight of greater than 30,000 was collected following the three centrifugations. Upon completion of the centrifugation, the combined filtrates and the retentate were examined for UDG activity using the standard assay.

The retentate from above was applied to a nickel charged iminodiacetic acid (Ni-IAA) affinity column (1.2 x 9cm) that was equilibrated in 20 mM sodium phosphate buffer (pH 7.8) containing 0.5 M NaCl. The matrix was washed with 5 bed volumes of
20 mM sodium phosphate buffer (pH 6.0) containing 0.5 M NaCl. The His-tagged Vpr-UDG complex was eluted by increasing the ionic strength using a linear gradient of 50 mM imidizole and 350 mM imidizole in 20 mM sodium phosphate buffer (pH 6.0) containing 0.5 M NaCl. Fractions were collected and assayed for UDG using our standard assay.

2.2.6 SDS-PAGE and Western Blotting

For the purification studies purity was determined using SDS-PAGE on 15% gels as previously described (327). SigmaMarker™ low molecular weight range kit (M3913) was used for protein standards. Proteins were identified using the Sigma Stain Silver Kit (AG-25) as described by the manufacturer. For Western blotting, proteins were transferred to a Hybond™ ECL™ nitrocellulose membrane (Amersham Pharmacia Biotech; Little Chalfont, Buckinghamshire, England) using the protocol established by Bio-Rad (Richmond, CA). The membrane was incubated in blocking buffer (20 mM Tris-HCl (pH 7.6), 137 mM NaCl, 0.1% (v/v) Tween and 5% (w/v) nonfat dry milk) for 1 hr at room temperature on an orbital shaker. The membrane was subsequently washed twice with TBS-T (20 mM Tris-HCl (pH 7.6), 137 mM NaCl, and 0.1% (v/v) Tween) for 20 minutes at room temperature on an orbital shaker. A rabbit polyclonal anti-Vpr (NL4-3) antibody (antiserum to HIV-1 Vpr (1-46)) obtained from Dr. Jeffery Kopp through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health was diluted 1:500 in blocking buffer and added and incubated overnight at 4°C on an orbital shaker. The membrane was subsequently washed twice with TBS-T for 20 minutes at room
temperature on an orbital shaker. An anti-rabbit goat antibody coupled to horseradish 
peroxidase (SC2004, Santa Cruz Biotechnology; Santa Cruz, CA) was diluted 1:5000 in 
blocking buffer and incubated with the membrane at room temperature for 1 hr on an 
orbital shaker. The membrane was subsequently washed twice with TBS-T for 20 
minutes at room temperature. The protein was detected by chemiluminescence using ECL 
horseradish peroxidase substrate (Amersham Pharmacia Biotech) and autoradiography. 
Additionally, immunoblotting was performed with an anti-His rabbit polyclonal antibody 
(SC803, Santa Cruz Biotechnology; Santa Cruz, CA) diluted 1:1000 in blocking buffer. 
The immunoblotting protocol described above was then followed.

2.2.7 Yeast Two-Hybrid Assay

The vectors used for expression of HIV-1 wild-type Vpr and the VprW54R 
variant fused to the Gal4 DNA binding domain (Gal4BD), and the vectors used for 
expression of the human UDG2 fused to the Gal4 activation domain (Gal4AD) have been 
previously described (194). Vectors used for expression of UDGs from E. coli and VZV 
fused to the Gal4AD were constructed by inserting the corresponding fragments obtained 
by PCR into pGAD-GE (194). Site-directed mutagenesis was used to generate a single-
point mutation into the WxxF sequence motif located in UDG2 and UDGcoli. The 
phenylalanine (F) residues in position 225 and 144 of UDG2 and UDGcoli, respectively 
were replaced by glutamine (Q) residues to generate the UDG2F225Q and 
UDGcoliF144Q mutants. The HF7c yeast reporter strain containing the two LexA-
inducible genes, HIS3 and LacZ, was co-transformed with the indicated Gal4AD
and Gal4BD hybrid expression vectors, plated on selective medium and double transformants were then assayed for histidine auxotrophy as previously described (194).

2.2.8 Analysis of UDG Incorporation into HIV-1 Virions

Incorporation of Vpr and UDG was analyzed using a packaging assay in which UDG was expressed in trans and incorporated into virions (171). Vectors for expression of HA-tagged UDGs from E. coli and VZV were constructed in the pAS1B plasmid as described previously for the HA-tagged UDG2 expression vector (174). The HIV-1 based packaging vectors pCMVΔR8.9 (lacking the env and auxillary genes) and pCMVΔ8.2 (lacking only the env gene) have been described (328). For incorporation analysis of UDGs, cells were transfected with 10 µg of either pCMVΔR8.2 or pCMVΔR8.9 in combination with either 5 µg of pAS1B-UDG2, -UDGcoli or –UDGvzv. Cell culture supernatants were collected 48 hr after transfections, filtered through 0.45 µm-pore-size filters. Virions were collected by ultracentrifugation at 4°C for 1h at 100,000 x g and suspended in ice-cold lysis buffer (10 mM Tris-HCl (pH 7.6), 150 mM NaCl, 2 mM EDTA, 0.5% TritonX-100). For preparation of cell lysates, cells were trypsinized, collected by centrifugation, and suspended in ice-cold lysis buffer. Proteins from cell (50 µg of total protein) and virion (50 ng of Cap24) lysates were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and analyzed by Western blotting with anti-HA 3F10 (Roche) or anti-CAp24 antibodies (NIH).

To determine virion-associated UDG activity, 293T cells were transiently transfected with a HIV-1 gag-pol expression vector pSVgagpolMPMV (10 µg), pAS1BVprwt (10 µg), and pAS1B UDGcoli (10 µg). Forty-eight hours post-transfection,
culture supernatants were collected, filtered and virions purified by ultracentrifugation at 40,000 x g for 1 hr. The virions were resuspended and lysed in RIPA buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5mM EDTA, 0.5% deoxycholate, 0.1% SDS and 1% IGEPAL). Virion-associated UDG activity was determined using a modification of a protocol described previously (193). Briefly, a 25-mer poly-T DNA oligonucleotide containing a uracil residue at position 13 was used as the substrate. The oligonucleotide was end-labeled using the DNA 5′-end labeling kit (Roche). The radiolabeled DNA oligonucleotide was incubated with the crude virus particle extract (1 µg protein) in a 10 µl reaction buffer containing 20 mM Tris-HCl (pH 8.0), 1mM EDTA and 1 mM dithiothreitol) at 37°C for 1 hr. Reactions were terminated by adding one-half volume 0.5 M NaOH and boiling, which cleaves apyrimidinic (AP) sites. Samples were applied to a non-denaturing 20% polyacrylamide gel with electrophoresis at 60 V for 3.5 hr. UDG activity would result in the conversion of the 25-mer oligonucleotide substrate to a 12-mer oligonucleotide product.

2.2.9 Molecular Modeling

A complete three-dimensional model for the Vpr molecule was constructed using sequence alignment with the partial NMR structures (PDB entries 1ESX, 1VPC, 1BDE and 1DSJ (329-333). Atomic coordinates corresponding to partial structures of Vpr were extracted from the above entries and a 96 residue Vpr model was built on Silicon Graphics Workstation using O (334). This model along with the crystal structure of the UDG-UGI complex (PDB entry 1UHGH) (278) was used to form the Vpr-UDG-UGI structural assembly using the following criteria. The domains of Vpr (residues 15 to
25 and 63 to 77) that show interaction with UDG according to the functional studies were oriented in order to maintain atomic interactions. Similarly, the domain of Vpr (residues 25 to 40) that is not reported to have interaction was placed away from UDG residues. In addition regions away from UGI were chosen to locate the binding site for Vpr because functionally Vpr binding has no effect on the availability of UGI. Furthermore, The UGI binding region corresponds to the DNA binding site of UDG.

2.3 Results

2.3.1 Purification of the Vpr-\textit{E. coli} UDG Complex

While various approaches have been employed to purify Vpr none have been successful. This has generally been attributed to Vpr’s toxicity, which is associated with its ability to induce G2, arrest and/or its instability. Previous studies have demonstrated, however, that Vpr forms a stable complex with the human nuclear isoform of UDG (UDG2) (174,193,194,203). Therefore I hypothesized that it should be possible to purify Vpr as a complex with UDG. Preliminary studies were performed to determine whether Vpr could be purified based upon its affinity for the VZV encoded UDG co-expressed in \textit{E. coli}. During the course of these studies it became apparent that Vpr exhibited very low or no significant affinity for the VZV encoded UDG (see Fig. 2.5A), but that all preparations of Vpr contained UDG activity. Further studies demonstrated that the expression of Vpr in \textit{E. coli} resulted in the binding of Vpr to the endogenous bacterial UDG (UDG coli) (data not shown). These results should not have been surprising considering the high degree of homology between UDGs (Fig. 1.5).
Since Vpr was apparently binding to the endogenous UDGcoli, I decided to purify this complex. The Vpr-UDGcoli complex was purified using affinity chromatography. A summary of a typical purification is shown in Table 2.1. Briefly, the Vpr-UDGcoli complex bound to Blue Sepharose matrix and eluted at a NaCl concentration of approximately 0.5 M. Fractions containing UDG activity were combined and subjected to concentration using a Centriprep-30 concentrator. Based upon the known molecular weights of Vpr (Mr: 11-15 kDa) and the *E. coli* UDG (Mr: 25.6 kDa), both proteins should be in the filtrate unless a complex of > 30 kDa was formed. Neither Vpr nor UDG could be detected in the filtrate, but both proteins were detected in the retentate (data not shown). The retentate from the concentration step was applied to a Ni-IAA affinity matrix and the bound Vpr-UDGcoli complex was eluted at approximately 200 mM imidazole. The elution profiles of the Vpr-UDG complex off of the Blue Sepharose and Ni-IAA columns were similar to that obtained with crude extracts of UDG from BL21 not containing Vpr (Blue Sepharose: 0.5 M NaCl; Ni-IAA: 200 mM) supporting the concept that the complex was being purified based upon the interaction of UDG with the matrixes. SDS-PAGE analysis demonstrated that the preparation following Ni-IAA separation contained two protein species with molecular weights of approximately 12 and 26 kDa, which correspond to the reported molecular weights of Vpr and the *E. coli* UDG, respectively (Fig 2.1A). Western blotting using a monoclonal antibody against the His-tag of the recombinant Vpr, as well as a polyclonal anti-Vpr (NL4-3) antibody confirmed that the protein exhibiting a molecular weight of 12 kDa was Vpr (Fig 2.1B,C). These results were also confirmed using the plasmid construct pET16b-vpr W54. The W54
The Vpr-UDGcoli complex was purified as described in Materials and Methods from exponentially growing IPTG (1 mM) induced BL21 cells containing the pET16b-vpr construct. The Vpr-UDGcoli complex was purified based upon the specific activity (units/mg) of the *E. coli* UDG associated with Vpr.

<table>
<thead>
<tr>
<th>Step</th>
<th>Specific Activity (x10^2)</th>
<th>Percent Recovery</th>
<th>Fold-Purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude Extract</td>
<td>2.55</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>Blue Sepharose Chromatography</td>
<td>17.09</td>
<td>110</td>
<td>6.70</td>
</tr>
<tr>
<td>Centricon -30K</td>
<td>117.08</td>
<td>59</td>
<td>45.91</td>
</tr>
<tr>
<td>Nickel-Affinity Chromatography</td>
<td>153.86</td>
<td>39</td>
<td>60.34</td>
</tr>
</tbody>
</table>

Table 2.1: Purification of Vpr-UDGcoli Complex.\(^a\)

\(^a\) The Vpr-UDGcoli complex was purified as described in Materials and Methods from exponentially growing IPTG (1 mM) induced BL21 cells containing the pET16b-vpr construct. The Vpr-UDGcoli complex was purified based upon the specific activity (units/mg) of the *E. coli* UDG associated with Vpr.
**Figure 2.1: SDS-PAGE and Western Blotting.** (A) SDS-PAGE was performed on 15% gels as described in Materials and Methods. *Lane 1.* Protein Standards: aprotinin (6.5 kDa), lactalbumin (14.2 kDa), soybean trypsin inhibitor (20 kDa), trypsinogen (24 kDa), carbonic anhydrase (29 kDa), glyceraldehyde-3-phosphate dehydrogenase (36 kDa), ovalbumin (45 kDa), bovine serum albumin (66 kDa). *Lane 2.* UDGcoli purified as described in Materials and Methods from IPTG induced BL21 cells containing the pET16b-vpr construct. Approximately 5 µg of the protein from the UDG active fractions following nickel affinity chromatography was analyzed. Proteins were visualized by silver staining as described in Materials and Methods. *Lane 3.* UDGcoli purified as described in Materials and Methods from IPTG induced BL21 cells containing the pET16b-vprW54R construct. Approximately 6 µg of the protein from the UDG active fractions following nickel affinity chromatography was analyzed. Proteins were visualized by silver staining as described in Materials and Methods. (B) Western blotting with a rabbit polyclonal anti-Vpr (NL4-3) antibody. UDGcoli was purified from IPTG induced BL21 cells containing the pET16b-vpr vector as described in Materials and Methods. Approximately 5 µg of the protein from the UDG active fractions following nickel affinity chromatography was subjected to SDS-PAGE and the proteins were transferred to a Hybond™ECL™ nitrocellulose membrane. The membrane was immunoblotted using a rabbit polyclonal anti-Vpr (NL4-3) antibody as the primary antibody and SC2004, an anti-rabbit goat antibody coupled to horseradish peroxidase, as the secondary antibody. The protein was detected by chemiluminescence. (C) Western blotting using an anti-His rabbit polyclonal antibody. The immunoblot described above
was stripped and re-probed using an anti-His rabbit polyclonal antibody (SC803) as the primary antibody and SC2004 as the secondary antibody. The protein was detected by chemiluminescence.
Figure 2.1

A.  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

- UDG
- Vpr

B.  

- **anti-Vpr**

C.  

- **anti-His**
mutation produces a Vpr molecule that does not bind to human UDG2 (193). While the Vpr W54R variant was expressed as efficiently as wild-type Vpr, it was not possible using our purification protocol to demonstrate the co-purification of VprW54R and UDGcoli (Fig. 2.1A, lane 3).

2.3.2 Biochemical Properties of the Vpr-UDGcoli Complex.

To determine whether the binding of Vpr to UDGcoli resulted in changes in the enzymatic activity of bound UDGcoli. The standard UDG assay was used to assess changes in UDGcoli activity upon addition of either monovalent cations (K\(^+\) and Na\(^+\); 50-200 mM) or polyamines (putrescine, spermine, and spermidine; 50-200 mM). Neither the addition of the monovalent cations nor polyamines altered UDGcoli activity bound to Vpr when compared to that of purified BL21 UDGcoli not bound to Vpr (Table 2.2). Furthermore, the interaction of UDGcoli with Vpr did not alter the specificity of UDGcoli nor did this interaction result in the acquisition of a new enzymatic activity, specifically deoxyuridine triphosphate nucleotidohydrolase (dUTPase) (data not shown). The binding of Vpr to UDGcoli, did enhance the stability of UDGcoli when applied to a Ni-IAA column. There was a 58-62% recovery of the BL21 UDGcoli activity when applied either alone or in combination with Vpr-W54 (Fig 2.2). Conversely, 94% of the UDGcoli activity was recovered when applied to the Ni-IAA column in the presence of wild-type Vpr. Additionally, the binding of Vpr to UDGcoli did not affect the ability of the PBS2 uracil-DNA glycosylase inhibitor (UGI) to inhibit UDG activity (Fig. 2.3A,B). As shown in Fig. 2.4A,B, the addition of purified his-tagged UGI (1.4 x 10\(^4\) units) to a Blue Sepharose column preloaded with Vpr-UDG (1.4 x 10\(^4\) units) resulted in the
<table>
<thead>
<tr>
<th>Additive</th>
<th>% Activity&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mM KCl NaCl Spermidine Spermine Putrescine</td>
</tr>
<tr>
<td>E. coli UDG</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>25</td>
<td>ND</td>
</tr>
<tr>
<td>50</td>
<td>46</td>
</tr>
<tr>
<td>100</td>
<td>28</td>
</tr>
<tr>
<td>200</td>
<td>10</td>
</tr>
<tr>
<td>Vpr-UDGcoli</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>25</td>
<td>ND</td>
</tr>
<tr>
<td>50</td>
<td>52</td>
</tr>
<tr>
<td>100</td>
<td>29</td>
</tr>
<tr>
<td>200</td>
<td>8</td>
</tr>
</tbody>
</table>

<sup>a</sup> Assays were performed as described in Materials and Methods with the addition of the monovalent cation or the polyamine. The amount of enzyme used per assay was 0.04 units for the E. coli UDG and 0.08 units for the Vpr-UDGcoli complex. Values represent the average of two experiments. ND means not done.

Table 2.2: Effect of Monovalent Cations and Polyamines on the Activities of E. coli UDG and the Vpr-UDGcoli Complex.
<table>
<thead>
<tr>
<th>% UDG Recovery</th>
<th>BL21-UDG</th>
<th>His-Vpr/BL21-UDG</th>
<th>His-Vpr (W54R)/BL21-UDG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wash</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Wash</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>200 mM</td>
<td>45.8</td>
<td>80</td>
<td>36.6</td>
</tr>
<tr>
<td>350 mM</td>
<td>14.5</td>
<td>9.1</td>
<td>12</td>
</tr>
<tr>
<td>350 mM</td>
<td>1.9</td>
<td>4.6</td>
<td>9.3</td>
</tr>
<tr>
<td>Total Recovery</td>
<td>62.2</td>
<td>93.7</td>
<td>57.9</td>
</tr>
</tbody>
</table>

**Figure 2.2: Enhanced Stability Imparted on *E. coli* UDG upon Binding to Vpr.** Cell extracts from BL21 cells, BL21 cells expressing Vpr, and BL21 cells expressing Vpr (W54R), each containing 10,000 units/mg of UDG activity, were applied to a Nickel-charged Iminodiacetic acid column. Columns were washed with equilibration buffer and UDG was eluted by increasing the ionic strength of NaCl (200-350 mM). The amount of UDG activity (units/mg) in each fraction was determined using the assays described in Materials and Methods. Each value represents the percentage of the initial UDG activity added to the column recovered. The values are the average of two independent studies.
inability to recover UDG activity from the column, but subsequent elution of the column and SDS-PAGE demonstrated the presence of Vpr, UDG and UGI (Fig. 2.4). This suggests that the binding of Vpr to UDG does not prevent the binding of UGI to UDG.

2.3.3  *E. coli* UDG Can Support Direct Interaction with Vpr

Since the biochemical studies demonstrated that HIV-1 Vpr co-purified in a complex with UDGcoli when expressed in *E. coli*, a yeast two-hybrid approach was implemented to confirm whether UDGcoli could bind directly to Vpr. The Vpr protein was fused to the Gal4BD and analyzed for interaction with UDGcoli fused to the Gal4AD in the HF7c yeast strain containing the Gal4-inducible reporter gene, *HIS3*. Vpr binding to human UDG2 was used as a positive control, while the VZV UDG protein was tested as a negative control. Interactions between Vpr and both UDGcoli and UDG2 were indicated by the ability of the HF7c strain co-transformed with the hybrid expression plasmids to grow in the absence of histidine (Fig. 2.5A). These interactions were specific, since there was no transcriptional activation of the reporter gene in yeast cells expressing Gal4BD-Vpr or Gal4AD-UDGs alone. In agreement with the purification studies, Vpr did not interact with VZV UDG in this assay. As expected, the Trp54 residue of Vpr is also crucial for binding to UDGcoli, since the VprW54R variant failed to interact with both UDGcoli and UDG2 (Fig. 2.5B). It has been reported that the interaction of HIV-1 Vpr with UDG2 is mediated by a WXXF motif located within the UDG2 sequence. A single-point mutation of the phenylalanine (F) residue of this motif was sufficient to abolish Vpr binding (200). Since there is an equivalent motif conserved
Figure 2.3: Inhibition of UDG by Tagged UGI. Assays were conducted as described in Materials and Methods containing either 0.0035 units of *E. coli* UDG or 0.0042 units of purified *E. coli* Vpr-UDG. Reactions were incubated at 37°C for 20 min. Under these conditions approximately 50% of the substrate was hydrolyzed in reactions lacking UGI. (A) GST-tagged UGI; (B) His-tagged UGI.
Figure 2.4: Ability of UGI to Interact with a Vpr-UDGcoli Complex. (A) *E. coli* extract containing His-Vpr was added to a Ni-IAA column. Subsequently, an *E. coli* extract-containing GST-UGI was added to the column preloaded with His-Vpr. (+) Indicates presence of activity, (-) indicates lack of activity. UDG and UGI assays were performed as previously described (UDG, 271; UGI, 288). (B) UGI is able to associate with a preformed Vpr-UDGcoli complex. SDS-PAGE analysis performed on individual fractions eluted off a Blue Sepharose affinity column preloaded with an *E. coli* protein extract containing over-expressed Vpr. Next, a protein extract prepared from *ung* *E. coli* over-expressing UGI was added to this preloaded affinity column. The column was eluted with increasing concentrations of NaCl. *Lane 1*, Molecular weight markers; *Lane 2*, 0.5 M NaCl; *Lane 3*, 1 M NaCl; *Lane 4*, 1 M NaCl; *Lane 5*, 2 M NaCl. UDG (Mr: ~26 kDa); UGI (Mr: 9.5 kDa); Vpr (11-15 kDa).
A.

<table>
<thead>
<tr>
<th>NaCl</th>
<th>UDG Recovery</th>
<th>UGI Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 mM</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>350 mM</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>350 mM</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

B.

<table>
<thead>
<tr>
<th>kDa</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.4
in UDGcoli (see Fig. 1.5), the phenylalanine residues in positions 144 of UDGcoli and 225 of UDG2 were mutated to generate the UDGcoliF144Q and UDG2F255Q variants. As indicated in Fig. 2.5B, both UDGcoliF144Q and UDG2F255Q failed to interact with Vpr, confirming the importance of the WxxF motif within UDGs in the binding of UDGs to HIV-1 Vpr. The VZV UDG, which did not interact with Vpr, contains a WxxL motif at the expected position (see Fig. 1.5) supporting the critical nature of the WxxF motif in the binding of HIV-1 Vpr to UDGs.

2.3.4  

**E. coli UDG Can be Incorporated into HIV-1 Virion**

Since it was previously demonstrated that the incorporation of Vpr into virus particles mediates virion packaging of UDG2 (174), I wanted to explore whether UDGcoli could also be recruited into virions. The incorporation of UDG proteins into virions was analyzed using a packaging assay in which HA-tagged UDG2, -UDGcoli and –UDGVZV were constructed in the plasmid pAS1B as described previously (174) and expressed in *trans* in virus-producing cells. 293T cells were transfected with one of the HA-tagged UDG expression vectors in combination with a HIV-1-based packaging vector either pCMVΔ8.2 (containing an intact *vpr* gene) or pCMVΔ8.9 (lacking an intact *vpr* gene). Cells and virions were analyzed for the presence of UDG by Western-blot analysis using an anti-HA monoclonal antibody. As shown in Fig. 2.6 (lower panel), both HA-tagged UDG2 and UDGcoli were detected in virions released into the supernatant of transfected cells. In contrast, UDGVZV was not incorporated despite detectable expression of the protein in transfected cell lysate (Fig. 2.6, upper panel). Similar to the UDG2 protein, UDGcoli was incorporated in a Vpr-dependent manner
Figure 2.5: Specific Interaction of Vpr with UDGcoli in a Yeast Two-Hybrid System. (A) Binding of Vpr to both *E. coli* and human UDGs. The HF7c reporter strain expressing Gal4BD-Vpr hybrid in combination with each of the Gal4AD-UDG hybrid indicated was analyzed for histidine auxotrophy. Double transformants were patched on selective medium with histidine (+His) and then replica plated on medium without histidine (-His). Growth in the absence of histidine indicates an interaction between hybrid proteins. (B) The WxxF motif of UDGcoli and the Trp54 residue of Vpr are required for the Vpr-UDGcoli interaction. HF7c expressing either the Gal4BD-Vpr or Bal4BD-VprW54R hybrid in combination with either wild-type or mutated UDG2 and UDGcoli fused to Gal4AD was analyzed for histidine auxotrophy. Data provided by E. Le Rouzic and S. Benichou.
### Figure 2.5

#### A

<table>
<thead>
<tr>
<th>Cal4BD-hybrid</th>
<th>Gal4AD-hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV-1 Vpr</td>
<td>+ His</td>
</tr>
<tr>
<td></td>
<td>- His</td>
</tr>
<tr>
<td>Gal4BD</td>
<td>hUNG2</td>
</tr>
<tr>
<td></td>
<td>E. coli UNG</td>
</tr>
<tr>
<td></td>
<td>VZV UNG</td>
</tr>
<tr>
<td></td>
<td>Gal4AD</td>
</tr>
<tr>
<td></td>
<td>hUNG2</td>
</tr>
<tr>
<td></td>
<td>E. coli UNG</td>
</tr>
<tr>
<td></td>
<td>VZV UNG</td>
</tr>
</tbody>
</table>

#### B

<table>
<thead>
<tr>
<th>Cal4BD-hybrid</th>
<th>Gal4AD-hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vpr</td>
<td>+ His</td>
</tr>
<tr>
<td>Vpr W54R</td>
<td>- His</td>
</tr>
<tr>
<td>Vpr</td>
<td>hUNG2</td>
</tr>
<tr>
<td>Vpr W54R</td>
<td>E. coli UNG</td>
</tr>
<tr>
<td>Vpr</td>
<td>hUNG2</td>
</tr>
<tr>
<td></td>
<td>E. coli UNG</td>
</tr>
<tr>
<td></td>
<td>hUNG2F234Q</td>
</tr>
<tr>
<td></td>
<td>E. coli UNGF144Q</td>
</tr>
</tbody>
</table>
Figure 2.6: Incorporation of UDGcoli into HIV-1 Virions. 293T cells were co-transfected with either pCMVΔR8.9 or pCMVΔR8.2 in combination with each HA-tagged UDG expression vector. Lysates from transfected cells and virions released in the cell culture medium were separated by SDS-PAGE and analyzed by Western blotting using anti-HA (cells and virions, upper panels) and anti-CAp26 monoclonal antibodies (cells and virions, lower panels). Data provided by E. Le Rouzic and S. Benichou.
Figure 2.6

<table>
<thead>
<tr>
<th></th>
<th>human UDG2</th>
<th>E. coli UDG</th>
<th>VZV UDG</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.2 (wtVpr)</td>
<td>- + -</td>
<td>- + -</td>
<td>- + -</td>
</tr>
<tr>
<td>Cells</td>
<td>[Image]</td>
<td>[Image]</td>
<td>[Image]</td>
</tr>
<tr>
<td>Virions</td>
<td>[Image]</td>
<td>[Image]</td>
<td>[Image]</td>
</tr>
</tbody>
</table>

- UDGs
- anti-HA
- Gag
- anti-p24
- UDGs
- anti-HA
- p24
- anti-p24
since it was not detected in virions when 293T cells were transfected with a HIV-1 packaging vector lacking the vpr gene (pCMVΔ8.9). The same blots were probed with an anti-Cap24 monoclonal antibody to verify that similar amounts of virions were produced in each transfection. Furthermore, UDG activity was detected using the previously described method in lysates of purified virions produced from 293T cells overexpressing UDGcoli (data not shown).

2.3.5 Structural Model of the Vpr-UDG and UGI Complex

The biochemical data demonstrate that the association of Vpr with UDGcoli does not alter the enzymatic function of UDG or interfere with the binding of UGI to UDGcoli. These results suggest that the overall structure of UDG remains intact following its association with Vpr. While the crystal structure of native Vpr is unknown, NMR studies on synthetic domains and on the intact synthetic Vpr molecule have provided insight into the structural organization of Vpr (326,329-332). Using the NMR data, data from mutational studies on Vpr, and structural data on UDG and UDG complexed to UGI, it was possible to develop a model depicting the interaction of Vpr with UDG2 and UDGcoli (Fig. 2.7). The structural model confirms the functional observations that residue 54W in Vpr and residues 141W and 144F in UDGcoli and residues 222W and 225F in UDG2 are involved in binding. The model suggests that binding of Vpr to UDG involves aromatic interactions between the 54W of Vpr and residues 141W and 144F and 222W and 225F of the E. coli and human UDGs, respectively. Replacing one of these residues with a non-aromatic residue would result in
the loss of stacking interactions and a weaker binding affinity between Vpr and UDGs. Furthermore, the model demonstrates the high-degree of structural similarity between UDG2 and UDGcoli.

UGI inhibits UDGs by acting as a transition-state DNA mimic. The major interactions between UGI and UDG occur between eight hydrophobic residues located in UGI and the conserved leucine loop (HPSPLS), residues 187 to 192 and 268 to 273 of the E. coli and human UDG, respectively (333). The conserved leucine residue in the UDG motif is involved with uracil base flipping and the binding of UGI effectively inhibits UDG activity (282,335). Thus, the model indicates that the binding of Vpr to UDG does not alter the structure of UDG and this confirms the biochemical data demonstrating that there is no significant alteration in UDG activity when bound to Vpr. The model corroborates the in vitro UDG activity and indicates that the lack of UDG activity in reactions where UGI is added is due to the binding of UGI to Vpr-UDG complex through UGI interaction with UDG.

2.4 Discussion

There is controversy concerning the mechanism(s) by which the human nuclear isoform of UDG (UDG2) is incorporated into HIV-1. One mechanism proposes that UDG is incorporated into HIV-1 by a Vpr-dependent interaction (174,193,194,321), while the second mechanism involves an integrase-dependent interaction (203,322,323). The results of the present study demonstrate conclusively that Vpr can interact directly with UDG and not only with human UDG2, but also with UDGcoli. The binding of UDGcoli to Vpr and the subsequent incorporation of the Vpr-UDGcoli complex into
Figure 2.7: Model of Vpr-UDG-UGI Interaction. Vpr is shown in red; residues 15 to 25 and 63 to 77, which interact with UDG, are shown in brown; residues 25 to 40 which interact with HHR23A but not UDG are shown in gray. The human UDG-UGI complex is shown in blue; a solid blue line represents human UDG residues 82 to 304 whereas UGI residues are indicated by the broken blue line. The *E. coli* UDG-UGI complex is shown in purple; a solid purple line represents *E. coli* UDG residues 3 to 226 whereas UGI residues are indicated by a broken purple line. Residue 54W of Vpr is shown with black bonds and atoms while residues 222W and 225F of human UDG is shown with gray atoms and bonds. Data provided by G. K. Balendiran.
Figure 2.7
HIV-1 virions should not be surprising since UDG is evolutionary conserved (282). However, as demonstrated by the results generated in this study, the binding affinity of Vpr with UDGs is not universal. A key determinant in the interaction of Vpr with various cellular proteins is the presence of a WxxF motif in the cellular protein (200). Although other regions of the UDG molecules may participate in the binding affinity of Vpr (SB and ELR, unpublished results), the data reported here confirm that the WxxF motif found within the sequences of both human and *E. coli* UDGs play a crucial role for the maintenance of the interaction. While the motif is conserved in the human and *E. coli* UDGs, it is altered in the VZV encoded UDG (WxxL). A leucine/isoleucine substitution for phenylalanine is commonly found in viral encoded UDGs. However, some viruses such as human herpesvirus 8 (HHV-8) and Epstein-Barr virus contain the WxxF motif in their UDGs (336,337). Such an interaction may be important since it has recently been reported that reciprocal regulatory interactions occur between HIV-1 and HHV-8 (338). However, studies to address this possibility have not been performed.

The initial observations reporting an interaction of Vpr (193) and/or integrase (203) with UDG2 suggested that UDG2 retained enzymatic activity. The data in the present study expand upon these initial observations, at least in the case of Vpr, and demonstrate that the interaction of Vpr with UDGcoli does not have any apparent affect on the function or structure of UDG. Binding of Vpr did not alter UDG’s substrate specificity, response to monovalent cations or polyamines, nor did it have any affect on the ability of UDG to interact with UGI. These results are consistent with the structural model depicting the interaction between Vpr, UDG and UGI. However, I cannot rule out, at this time, the possibility that Vpr binding is altering kinetic properties (\(K_M, K_D, V_{\text{max}}\)).
Conversely, UDG2 is reported to interact with several cellular proteins (PCNA and RPA) and the interaction of Vpr with UDG may alter the ability of UDG to interact with these cellular proteins. Thus, the Vpr-UDGcoli complex can serve as a model for Vpr-UDG2 complex that naturally occurs in HIV-1 infected cells and virus particles.

The results of this study also demonstrate a novel method for purifying native Vpr complexed to UDGcoli. It has not been possible to purify native Vpr using classical biochemical or recombinant DNA approaches. This is due in part to the relative toxicity of Vpr, its instability and tendency to spontaneously aggregate at high concentrations (320,325,326). While there are recent reports describing the synthesis of the entire Vpr molecule, the synthesis remains difficult and expensive (324-326,333). Using the relatively simple protocol described in this study, sufficient quantities of the Vpr-UDG complex are easily obtained for use in functional and structural studies. Prior difficulties with purifying native Vpr and the expense and difficulties associated with manufacturing synthetic molecules understate the significance of the results demonstrated in this study.

The lentiviruses have evolved differently to ensure that uracil residues are not incorporated or retained in viral DNA. The non-primate lentiviruses encode for a dUTPase, which hydrolyzes dUTP thus, preventing it from being incorporated into the viral DNA by reverse transcriptase (125-128). Conversely, the primate lentiviruses incorporate cellular UDG (174,193,194,203,321-323). While the role of the virion incorporated UDG in the life cycle of the primate lentiviruses has not been fully elucidated, it has been reported that UDG may be altering the rate at which mutations occur during reverse transcription (174). However, given the high rate of mutation associated with HIV-1, it does not seem likely that the virus would have specifically
evolved the ability to incorporate UDG to lower these rates. An alternative role for UDG in the primate lentiviruses’ life cycle may simply involve removing uracil residues from the viral DNA that otherwise would disrupt efficient virus replication. Mazumder and Pommier (339) demonstrated that deoxyuridine residues within and near the conserved CA dinucleotide of the U5 end of the HIV-1 long terminal repeat (LTR) altered 3′-processing and strand transfer. Moreover, a recent study demonstrated that substitution of dUTP for dTTP in an in vitro reverse transcriptase assay had no effect on minus strand DNA synthesis. However, it significantly decreased plus strand DNA synthesis and this decrease correlated with an alteration in RNaseH cleavage specificity (340).

While dUTPase and UDG may be required for efficient replication of lentiviruses in non-dividing cells, their precise role in lentivirus replication remains unclear. Furthermore, while it has been suggested that intracellular dUTP pools are elevated in non-dividing cells (323,127), experimental data, although limited, indicates that dUTP levels are not increased in macrophages and that the dUTP to dTTP ratio in macrophages is similar to that observed in rapidly dividing cells (341,342). It is apparent that further studies are necessary to determine the effect that lentivirus infection has on cellular nucleotide pools, in particular dUTP and dTTP, as well as the relationship between the viral encoded and cellular dUTPase, UDG, and reverse transcriptase. Such studies are essential to determine the roles of these enzymes in lentivirus replication.
CHAPTER 3

MECHANISM(S) OF URACIL INCORPORATION INTO HUMAN IMMUNODEFICIENCY TYPE 1 DNA

3.1 Introduction

Uracil residues can arise in DNA either by dCMP deamination or incorporation of dUTP by a DNA polymerase/reverse transcriptase during DNA replication or repair. Organisms, including lentiviruses, have evolved mechanisms to prevent the incorporation and/or retention of uracil in their DNA. The two primary methods involve the enzymes deoxyuridine triphosphate nucleotidohydrolase (dUTPase) and uracil DNA glycosylase (UDG). dUTPase is responsible for reducing the pool of dUTP. Since DNA polymerases/reverse transcriptases demonstrate very little selectivity between dTTP and dUTP (243), reducing the dUTP pool eliminates dUTP as a substrate for these enzymes during DNA synthesis and/or repair. UDG initiates a base excision repair process that excises uracil residues located in DNA. Non-primate lentiviruses encode a dUTPase as a structural component of the virus (125-128), while primate lentiviruses incorporate into progeny virus the cellular UDG from an infected cell (174,193,194,203,321-323).
Several lines of evidence suggest that dUTPase and UDG may have important functions in the replication of lentiviruses. The mutation spectra associated with HIV-1 not able to incorporate UDG and dUTPase mutants of non-primate lentiviruses depicted an increased frequency of G-to-A transition mutations (174). Moreover, dUTPase deficient non-primate lentiviruses displayed slower replication kinetics (129,130). Recent studies in HIV-1 have demonstrated that uracil residues in HIV-1 DNA inhibit critical steps in the HIV-1 life-cycle. Mazumder and Pommier (339) demonstrated that deoxyuridine residues located within and near the conserved CA dinucleotide of the U5 end of the HIV-1 LTR altered 3′-processing and strand transfer. Additionally, a recent study demonstrated that uracil incorporation into HIV-1 minus-strand DNA altered the specificity of plus-strand DNA synthesis initiation during reverse transcription (340). The lack of specificity was created by a cascade of events initiated from uracil incorporation into minus-strand DNA resulting in decreased initiation from polypurine tracts (PPTs) and increased initiation of plus-strand DNA synthesis from non-PPT sites.

The incorporation of cellular UDG into the virion suggests that primate lentiviruses have evolved a mechanism to reduce the retention of uracil residues in their DNA. However, the process(es) by which uracil residues are formed in HIV-1 DNA are poorly understood. As mentioned previously, uracil can arise in the viral DNA either by cytosine deamination or by incorporation of dUTP by RT during reverse transcription. It has been estimated that approximately 400 uracil residues are formed by cytosine deamination per day in the human genome (248,249). Using this value, I estimated that approximately 0.001 cytosine deamination events could occur in HIV-1 DNA per round of reverse transcription. Based upon these calculations it appears that cytosine
deamination is not a significant source of uracil formation in HIV-1 DNA. Conversely, the amount of uracil introduced into the viral DNA by incorporation of dUTP by RT has the potential to be considerably higher. Kornberg initially estimated that one out of every 1200 nucleotides introduced into newly synthesized DNA would be dUMP (243). That would indicate that approximately 16 molecules of dUMP would be incorporated into HIV-1 DNA during reverse transcription. This number represents a 16,000-times greater probability of uracil being introduced by RT than by cytosine deamination. However, the likelihood of this event occurring also has some limitations, specifically, the presence of a dUTP pool and its size relative to the dTTP pool and the ability of RT to insert dUMP into DNA. Cells usually have a very low level of dUTP because of the action of dUTPase.

HIV-1 exhibits an extremely high mutation rate when compared to other organisms, with G-to-A transitions accounting for a large percentage of the total mutational spectrum. This type of mutation is increased in UDG deficient HIV-1 and dUTPase deficient non-primate lentiviruses (129,130,174). The high frequency of G-to-A mutations observed in these UDG and dUTPase deficient mutants is most likely the result of dUTP mispairing with guanine during minus-strand DNA synthesis. Analysis of nucleotide base-pairing chemistry indicates that the frequency of this event occurring is very low. dUTP is able to base-pair with guanine when uracil is in the enol form, however this form only represents 0.01% of the total dUTP population, the remainder being in the keto form (343). Even though the amount of the enol form is very low I estimated that approximately 0.16 dU:G events could occur in HIV-1 DNA during one round of reverse transcription. This amount is significantly greater (160-times) than the
likelihood of the mutation being caused by dCMP deamination. Another factor that may actually increase the amount of uracil incorporation is the level of the dCTP pool. It has been reported that this pool is diminished in HIV-1 infected cells (154,344,345). A low level of the dCTP pool would increase the likelihood of incorporation of the wrong nucleotide (e.g. dUTP) across from guanine during reverse transcription. Thus it is important to look at nucleotide pools to determine whether alterations have occurred.

The poor nucleotide specificity of RT which would allow the incorporation of dUMP across from guanine and the ability of RT to extend this mispair may allow this event to occur more frequently in HIV-1 infected cells than in other organisms. Thus warranting the incorporation of UDG.

The purpose of this study was to gain insight into the mechanisms by which uracil may be introduced into the viral DNA. I hypothesized that the poor nucleotide selectivity of HIV-1 RT and its ability to extend a mismatched dU:G is responsible for the uracil found in the viral DNA. This process can be further perpetuated by modification of nucleotide pool levels. The results of this study demonstrate that none of the DNA polymerases or reverse transcriptases examined were effective in incorporating a dUTP residue opposite a dCMP residue in the template DNA. Secondly, that while differences existed in the overall amounts of dCTP and dTTP present in these pools; the dTTP was always elevated when compared to dCTP. None of the cells examined contained significant levels of dUTP. Finally, while all the purified polymerases could incorporate dUTP into DNA, dUTP could only be incorporated into SIV RNA when UDG was inhibited.
3.2 Materials and Methods

3.2.1 Materials

Enzymes

Recombinant Epstein–Barr virus dUTPase (75 units/mg; 1 unit is defined as the amount of enzyme required to hydrolyze 1 nmol of dUTP to dUMP per min at 37°C under our assay conditions) was purified as described (Williams, unpublished results). *Escherichia coli* DNA polymerase I and the large subunit of DNA polymerase I (Klenow) (11.8 units/µg and 5 units/µg, respectively; 1 unit is defined as the amount of enzyme required to incorporate 10 nmol of dNTP into acid-soluble material in 30 min at 37°C) were obtained from Invitrogen-GIBCO (Carlsbad, CA). Avian myeloblastosis virus (AMV) RT and Moloney murine leukemia virus (M-MLV) RT (10 units/µl and 200 units/µg, respectively; a unit is defined as the amount of enzyme that converts 1 nmol of dNTP into acid-soluble material in 10 min at 37°C) were obtained from Promega (Madison, WI). Human immunodeficiency virus type 1 (HIV-1) RT (30 units/µg; 1 unit is defined as the amount of enzyme that incorporates 1 nmol of dNTP into acid-soluble material in 10 min at 37°C) was obtained from Dr. Stuart LeGrice through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health #2897.

Primer-Templates

Synthesized oligonucleotides were obtained from Sigma-Genosys. (St. Louis, MO). To measure dTTP and dUTP nucleotide pools the DNA template: 5’-TTTATTTATTTATTTATTTAGGCGGTGGAGGCGG-3’ (dT/U) was used.
To measure the dCTP nucleotide pool the DNA template:

5’-TTTGTGTGTTTGTGGTGGGCGGTGAGGCGG-3’ (dC) was used. The nucleotide pool DNA templates were annealed to the primer:

5’-CCGCCTCCACCGCC-3’ for the nucleotide pool studies and to the primer:

5’-CCGCCTCCACCGC-3’ for the nucleotide insertion/mismatch extension assay.

**Chemicals and Materials**

2’-Deoxyadenosine 5-triphosphate, tetraammonium salt, [2,8-3H] (specific activity 20 Ci/mmol) was purchased from Moravek Biochemicals (Brea, CA). Non-radioactive high performance liquid chromatography-purified dNTPs were purchased from Sigma Chemical Co. Unless stated all other chemicals were obtained from Sigma. Microcon-3 microcentrators were purchased from Amicon (Beverly, MA). Whatman exchange filters were purchased from Fisher. Cell culture media and medium supplements were obtained from Invitrogen-GIBCO.

**3.2.2 Cell Culture and Extraction**

SW620 and HT29 human colon adenocarcinoma cells and HeLa human cervical epitheloid carcinoma cells were grown and maintained in modified McCoy’s 5A medium supplemented with 10% (v/v) fetal bovine serum and 1% penicillin (10 units)-streptomycin (10 µg) at 37°C in a CO2 atmosphere. For pool determinations cells were plated at a density of 1.25 X 10⁵ cells/well (6-well plates). After 48 hr the medium was discarded and the cells extracted in 400 µl of cold 60% methanol for 1 hr at 4°C.
Cell extracts were centrifuged in a microcon-3 microconcentrators to remove precipitated protein and DNA. The filtrate was divided into paired aliquots and lyophilized to remove the methanol.

### 3.2.3 Nucleotide Pool Assays

Deoxynucleotide triphosphate pool sizes were determined by an enzymatic method described by Williams et al (346). dTTP and dUTP pools were determined independently by pre-incubating cellular extracts with or without recombinant Epstein-Barr virus dUTPase. Briefly, the paired aliquots described above, were resuspended in 40 µl of buffer containing 34 mM Tris-HCl (pH 8.3), 10 mM MgCl₂, 1 mM EDTA and 0.5 mg/ml bovine serum albumin (BSA). 5 units recombinant EBV dUTPase was added to one aliquot. Samples were then incubated for 20 min at 37°C. Following incubation 60 µl of 100% cold methanol was added to give a final methanol concentration of 60% and the samples were centrifuged in a microcon-3 microconcentrator. The filtrate was again lyophilized to remove the methanol.

The DNA polymerase assay utilized to measure deoxynucleotide pools was modified from that of Williams et al (346). Briefly, the reaction mixture contained in a total volume of 100 µl: 34 mM Tris-HCl (pH 8.3), 10 mM MgCl₂, 0.2 mM EDTA, 0.25 mg/ml BSA, 0.01 µCi/µl [³H]-dATP, 16 nM of each oligonucleotide (template and primer), 0.3 units of the 3′→5′ exonuclease deficient large fragment of DNA pol I (Klenow fragment) and the dNTP sample. The reaction mixture was incubated for 45 min at 42°C. Reactions were terminated by spotting 30 ul of the reaction mixture on DE81 filter discs followed by washing the discs for 5 min with 1% sodium
pyrophosphate (NaPPi) and 5% trichloroacetic acid (TCA) six times and dried. The radioactivity incorporated into acid-insoluble material was determined by liquid scintillation spectrophotometry using a Beckman LS 6000IC. The amount of dNTP in the sample was determined using standard curves established in the laboratory. Results are expressed as pmoles/10^6 cells.

Standard curves were obtained by assaying serial dilutions of dCTP, as well as dUTPase treated and untreated dUTP and dTTP. Pretreatment of the nucleotides with dUTPase was necessary to ensure complete dUTP digestion. dTTP concentrations were determined by comparing the results of the dUTPase-treated sample to the standard curve generated from dTTP treated with dUTPase. dUTP concentrations were determined by subtracting the counts associated with dTTP from the counts attributable to the sample not treated with dUTPase, which is equal to dUTP + dTTP. Samples used to generate the standard curve and the experimental samples were assayed in triplicate, and the experiment was performed twelve times.

3.2.4 Nucleotide Insertion/Mismatch Extension Assay

Annealing Conditions

Annealing of the DNA primer to the DNA template was carried out at 42°C for 2 hr and the samples were allowed to cool slowly at 4°C overnight. For assays utilizing DNA polymerase I and Klenow, annealing was performed in a reaction containing 50 mM Tris-HCl (pH 7.5), 1.0 mM 2-mercaptoethanol (BME), 6.6 mM MgCl2, and 36 nM each of template and primer. For assays utilizing HIV-1 RT, annealing was performed in a reaction containing 50 mM Tris-HCl (pH 7.5), 75 mM KCl, 2.0 mM dithiothreitol
(DTT), 5.0 mM MgCl₂, 0.05% NP40, 0.5 mM EGTA, and 36 nM each of template and primer. For assays utilizing M-MLV RT, annealing was performed in a reaction containing 50 mM Tris-HCl (pH 8.0), 40 mM KCl, 10 mM DTT, 5.0 mM MgCl₂, 0.01% NP40, 0.1 mg/ml BSA, and 36 nM each of template and primer. For assays utilizing AMV RT, annealing was performed in a reaction containing 50 mM Tris-HCl (pH 8.0), 40 mM KCl, 10 mM DTT, 5.0 mM MgCl₂, 0.1 mg/ml BSA, and 36 nM each of template and primer.

**Nucleotide Reactions**

The annealed primer-template (20 µl) was mixed with the enzyme (5 µl) and allowed to incubate for 5 min at 37°C. The reaction was started by adding 25 µl of a solution containing 20 µM of various dNTPs in 50 mM Tris-HCl (pH 7.5), 1.0 mM BME, 6.6 mM MgCl₂ for the DNA polymerase I and Klenow reactions; for the HIV-1 RT reactions the mixing conditions consisted of 50 mM Tris-HCl (pH 7.5), 75 mM KCl, 2.0 mM DTT, 5.0 mM MgCl₂, 0.05% NP40, 0.5 mM EGTA; for the M-MLV RT reactions the mixing conditions consisted of 50 mM Tris-HCl (pH 8.0), 40 mM KCl, 10 mM DTT, 5.0 mM MgCl₂, 0.01% NP40, 0.1 mg/ml BSA; and for the AMV RT reactions the mixing conditions consisted of 50 mM Tris-HCl (pH 8.0), 40 mM KCl, 10 mM DTT, 5.0 mM MgCl₂, 0.1 mg/ml BSA. Reaction mixtures were incubated at 37°C for various time periods depending on the enzyme. Reactions were terminated by adding 25 µl of 20 mM EDTA and immediately spotting 50 µl of each sample onto DE81 filter discs. The discs were washed twice for 5 min in a solution containing 1% sodium pyrophosphate (NaPPi)
and 5% trichloroacetic acid (TCA) and then washed once in 100% ethanol. The radioactivity incorporated into acid-insoluble material was determined by liquid scintillation spectrophotometry using a Beckman LS 6000IC.

3.2.5 Effect of UGI on Reverse Transcription

The reverse transcriptase reaction mixture contained in a total volume of 0.075 ml: 33 mM Tris-HCl (pH7.8), 50 mM KCl, 1.3 mM dithiothreitol, 3.3 mM MgCl₂, 2.7 μg poly (rA)-oligo(dT)₁₂₋₁₈, 0.03% NP40, 0.33 mM EGTA and either 0.5 μM dTTP (specific activity 250 μCi/nmol) or 0.5 μM dUTP (specific activity 140 μCi/nmol) and the enzyme. Reaction mixtures were incubated at 37°C and terminated by spotting 50 μl of the reaction mixture on a glass fiber disc and immediately washing the disc twice for 5 min each in an ice cold solution of 5% TCA and 1% sodium pyrophosphate followed by a single wash in ethanol.

3.3 RESULTS

3.3.1 Ability to Incorporate and Extend an Incorrect Nucleotide

In an attempt to elucidate a mechanism by which uracil incorporation into viral DNA might result in the G- to-A hypermutagenesis observed in lentiviruses, I assessed the ability of various polymerases to incorporate an incorrect nucleotide and extend from this mismatch. The following polymerases were used: *E. coli* DNA polymerase I, the Klenow fragment of *E. coli* DNA polymerase I, and the reverse transcriptases (RT) encoded by HIV-1 and Moloney murine leukemia virus (M-MLV). Similar studies looking at misincorporation and the ability to extend a mismatch have been performed
previously, but there has been no reports looking at dUTP misincorporation. As described in Materials and Methods these polymerases were analyzed for their ability to incorporate the incorrect nucleotide (dTTP or dUTP) and the correct nucleotide (dCTP) across from a template guanine. Misincorporation of dUTP across from guanine would create a G-to-A transition mutation in the next round of replication if not repaired. As shown in Table 3.1, none of the polymerases were able to incorporate dUTP across from guanine. Conversely, all of the polymerases were able to incorporate dCTP, the correct nucleotide. All four enzymes were able to incorporate dTTP, but at greatly reduced level compared to dCTP.

3.3.2 Pyrimidine Deoxyribonucleoside Triphosphate Pool Concentrations

Current knowledge of the enzymes involved in the metabolism of deoxyribonucleoside triphosphate (dNTP) pools, as well as previous studies measuring dNTP pools indicate that dUTP pools are significantly lower compared to dTTP in mammalian cell (347,348). To begin to address the impact of HIV-1 infection on dNTP pool levels the cell lines SW620, HT29 and HeLa were analyzed for dUTP, dTTP, and dCTP levels. The results shown in Table 3.2 demonstrate that dUTP pools were significantly lower than dTTP pools, but dTTP pools were higher than dCTP pools. Moreover, the assay used in these studies (sensitivity: 1 μmole/10⁶ cells) was not able to generate any value for dUTP pools.
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Correct Nucleotide</th>
<th>Incorrect Nucleotide</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dCTP</td>
<td>dTTP</td>
</tr>
<tr>
<td>Pol I</td>
<td>104.62 ± 28.11</td>
<td>17.64 ± 4.48</td>
</tr>
<tr>
<td>Klenow</td>
<td>123.48 ± 20.71</td>
<td>40.87 ± 13.52</td>
</tr>
<tr>
<td>HIV-1 RT</td>
<td>65.24 ± 3.25</td>
<td>3.98 ± 1.46</td>
</tr>
<tr>
<td>M-MLV RT</td>
<td>41.05 ± 8.57</td>
<td>0.98 ± 1.26</td>
</tr>
</tbody>
</table>

*a Assays were performed as described in Materials and Methods. Values represent pmole of dNTP incorporated into the synthetic template. For *E. coli* DNA polymerase I and Klenow fragment reaction mixtures were incubated for 2.5 minutes. For HIV-1 and MMLV reactions mixtures were incubated 20 min. Values represent the average ± standard deviation of at least two experiments.

**Table 3.1: Ability of Various Replicative Polymerases to Misincorporate the Incorrect Nucleotide and Extend from this Mismatch.**

### 3.3.3 Effect of Ugi on Reverse Transcription

While several studies have reported that functionally active UDG2 or its precursor is incorporated into HIV-1 or SIV virions (193,203,322), none of these studies have demonstrated whether the virion incorporated UDG functions during reverse transcription. Furthermore, no studies have demonstrated whether RT can use dUTP in place of dTTP as a substrate during reverse transcription in the presence of UDG. To address these questions an *in vitro* assay was utilized to determine whether dUTP was incorporated into an oligo(dT)$_{12-18}$-primer, poly rA-template at similar levels as dTTP.
Virions collected from the supernatants of infected cells were used as a source of reverse transcriptase. The results, which are summarized in Table 3.3, demonstrate that dUTP was not incorporated/retained at the same efficiency as dTTP when supernatants from SIV infected cells were used. However, dUTP incorporation was enhanced when reactions were performed in the presence of UGI. There was not a significant difference between the incorporation and retention of dUTP and dTTP in the reactions using purified Avian myeloblastosis virus (AMV) and M-MLV RTs.

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>dUTP</th>
<th>dTTP</th>
<th>dCTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>HeLa</td>
<td>&lt;1</td>
<td>37.8 ± 3.4</td>
<td>28.1 ± 5.3</td>
</tr>
<tr>
<td>SW620</td>
<td>&lt;1</td>
<td>76.4 ± 18.3</td>
<td>42.4 ± 11.7</td>
</tr>
<tr>
<td>HT29</td>
<td>&lt;1</td>
<td>91.7 ± 10.6</td>
<td>33.7 ± 8.1</td>
</tr>
</tbody>
</table>

* Concentrations are expressed as pmole/10^6 cells. Values represent the average ± the standard deviation of two experiments.

Table 3.2: Pyrimidine Deoxyribonucleoside Triphosphate Pool Concentrations.
3.4 DISCUSSION

It has been proposed that UDG functions as an anti-mutator either by specifically removing uracil residues from the viral DNA or by other non-enzymatic mechanisms such as altering the access of deoxyribonucleotide triphosphates to reverse transcriptase (RT) or directly interacting with RT to increase its fidelity and/or processivity (174).

<table>
<thead>
<tr>
<th>Source</th>
<th>pmole dUTP incorporated</th>
<th>pmole dTTP incorporated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purified AMV RT</td>
<td>1.62 ± 0.45</td>
<td>1.56 ± 0.33</td>
</tr>
<tr>
<td>Purified AMV RT + UGI</td>
<td>1.55 ± 0.33</td>
<td>1.77 ± 0.22</td>
</tr>
<tr>
<td>Purified M-MLV RT</td>
<td>1.20 ± 0.18</td>
<td>1.08 ± 0.11</td>
</tr>
<tr>
<td>Purified M-MLV RT + UGI</td>
<td>0.99 ± 0.34</td>
<td>1.02 ± 0.09</td>
</tr>
<tr>
<td>Sup. (non-infected cells)</td>
<td>0.01 ± 0.01</td>
<td>0.01 ± 0.01</td>
</tr>
<tr>
<td>Sup. (non-infected cells) + UGI</td>
<td>0.01 ± 0.01</td>
<td>0.01 ± 0.01</td>
</tr>
<tr>
<td>Sup. (SIV)</td>
<td>0.14 ± 0.03</td>
<td>8.70 ± 1.41</td>
</tr>
<tr>
<td>Sup. (SIV) + UGI</td>
<td>8.42 ± 0.74</td>
<td>9.14 ± 1.23</td>
</tr>
</tbody>
</table>

*Approximately 56 and 110 units as defined by the manufacturer of AMV and M-MLV respectively were employed per assay. Approximately 0.002 units of purified Tat-UGI were employed per assay. Reactions were started by the addition of the substrate and incubated at 37°C for 2 (AMV, M-MLV) or 3.5 hrs for the supernatants from non-infected and SIV infected cells.

Table 3.3: Effect of Ugi on Reverse Transcription*
Thus, it was concluded by these investigators that UDG influences the fidelity associated with the reverse transcription process. However, this premise not only presents several problems, but it also raises additional questions. (i) Although several DNA viruses encode a UDG (209,210), there is no evidence to support the concept that viruses repair their DNA during their replicative processes. This is further complicated by the feature that repair of uracil residues in DNA requires five enzymatic activities. Only the genomes of the viruses in the family Poxviridae, genus Entomopoxvirus have been reported to encode for all the enzymes necessary to repair a uracil residue in viral DNA (349,350). (ii) My analysis of the data employing a single cycle HIV-1 replication model (118,124,174), demonstrated that there was no statistical difference in the frequency of G-to-A transition mutations in HIV-1 Vpr mutants, which could not bind UDG and the wild-type virus. Most mutations in the Vpr-mutants occurred in regions of the recorder gene (lacZα) that had previously been identified as mutational hotspots using in vitro assays with HIV-1 reverse transcriptase (293-295). Interpretation of these studies suggests that neither Vpr nor UDG influences the formation and/or retention of G-to-A mutations in HIV-1. (iii) If UDG was involved in removing uracil from the DNA then the percentage of G-to-A transition mutations would be expected to be lower in wild-type HIV-1, but this is not the case. (iv) Lastly, if the success of HIV-1 and lentiviruses is due in part to their ability to generate variants, the packaging of UDG, a so-called anti-mutator would appear paradoxical. Taken together there is very little evidence to suggest the UDG, which is incorporated into the virion of primate lentiviruses, is being used in a
DNA repair related process. Conversely, UDG maybe incorporated to remove uracil residues from the viral DNA that could potentially inhibit critical steps in the HIV-1 lifecycle.

In addition to their potentially mutagenic properties (reviewed in 242,351,352), uracil residues in DNA also alter the structure of DNA (353-357) and interfere with DNA-protein interactions (358-360). Mazumder and Pommier (339) demonstrated that deoxyuridine residues within and near the conserved CA dinucleotide of the U5 end of the HIV-1 LTR altered 3’-processing and strand transfer. Moreover, a recent study demonstrated that substitution of dUTP for dTTP in an in vitro reverse transcriptase assay had no effect on (-) strand DNA synthesis. However, it significantly decreased (+) strand DNA synthesis and this decrease correlated with an alteration in RNase cleavage specificity (340). Thus, in the case of the primate lentiviruses UDG may be incorporated into the virion to remove uracil residues that could be detrimental to the HIV/SIV replication process.

An essential question that remains unanswered is what mechanism(s) is responsible for uracil formation in HIV DNA during reverse transcription. While it is not possible to accurately estimate at this time the amount of dUTP that may be incorporated into HIV-1 DNA by reverse transcriptase or from dCMP deamination, I have estimated that at a maximum there would be 16 molecules of dUTP incorporated into HIV-1 DNA during reverse transcription and 0.001 molecule of uracil formed from deamination of dCMP per round of HIV-1 replication. While these studies did not address the issue of cytosine deamination, the results of the in vitro DNA polymerization assays indicate that SIV RT incorporated dUTP as efficiently as dTTP (Table 3.3). However, these assays
were performed in the presence on only one nucleotide, either dUTP or dTTP. A more accurate determination of RTs capability of using dUTP as a substrate in DNA polymerization should include dTTP in the reaction mixture, acting as a competitor. Alternatively, $K_m$ values for each nucleotide could be determined.

Further analysis of the results generated in Table 3.3 indicated that supernatants from SIV infected cells were not able to polymerize DNA when dUTP was used in the assay but were able to when dTTP was used. The inability to use dUTP as a substrate was lost when UGI was added to the reaction mixture. These results suggest that while RT can incorporate dUTP very efficiently in the absence of UDG, an enzyme that specifically excises uracil from DNA, the presence of UDG prevents DNA polymerization from proceeding. As one would have predicted the presence of UGI did not affect the amount of dUTP incorporated by either AMV or M-MLV RT. The difference in the results from SIV and AMV, M-MLV was probably due the presence of UDG in the SIV reaction.

The results presented in Table 3.3 demonstrated that RT could introduce uracil into viral DNA very efficiently when base-pairing with adenine. On the other hand the results generated in Table 3.1 demonstrated that RT could not incorporate dUTP across from guanine. Although this finding was not surprising because the enol form of uracil, which represents only 0.01% of the total uracil population, is required for base-pairing with guanine to occur, an alternative mechanism to explain the G-to-A hypermutagenesis observed in HIV-1 needs to be identified. It has been suggested that an alternative way to explain the G-to-A hypermutagenesis, characteristic of lentiviruses, is a disruption in the host dNTP pool levels (154,361,362). In particular, researchers have argued that an
increase in the dUTP pool compared to the dTTP and/or a decrease in the dCTP pool could explain the increased incidence of uracil incorporation. Although the studies shown in Table 3.2 were performed on non-infected cells they demonstrate that in actively dividing cells there was no detectable amount of dUTP, thus indicating the efficiency of dUTPase to maintain very low levels of dUTP. It would be beneficial to determine the levels of dNTP pools in primary macrophages, a natural site of HIV-1 infection, to determine if infection alters the pools.
CHAPTER 4

PURIFICATION AND CHARACTERIZATION OF TAT-UGI:
A NOVEL PROTEIN TRANSDUCTION MOLECULE

4.1 Introduction

Uracil-DNA glycosylase (UDG) is the major enzyme responsible for preventing the retention of potentially detrimental uracil residues that occur in DNA (reviewed in 211,212). UDG is expressed from a highly conserved family of genes that have been identified in a wide variety of organisms (201,204-210). The human udg gene, located on chromosome 12q24.1, encodes for mitochondrial (UDG1) and nuclear (UDG2) isoforms of UDG through the use of different promoters and alternative splicing (213-215). UDG2 interacts with proliferating cell nuclear antigen (PCNA) and with replication factor A (RPA) and is localized within replication foci (310,363). Thus, it has been proposed that UDG2 is the primary enzyme responsible for initiating the removal of uracil residues from replicating DNA (211,212). In addition to its role in post-replicative DNA repair processes, UDG has been demonstrated to be involved with immunoglobulin class switching recombination and somatic hypermutation (311-313).
Studies to determine the role of UDG in various cellular processes have been limited because of the inability to isolate mutants defective in UDG activity. Recently UDG knock-out mice have been generated and these mice, which are viable, and established cell lines from these mice have provided information concerning the role of this enzyme in DNA replication and repair (239) and immunoglobulin switching (314). Directly relevant to our studies is the demonstration that UDG is the primary enzyme for uracil excision from DNA. In the absence of UDG activity the steady-state level of uracil increased to approximately 2000 residues per genome (239). Furthermore, there was no increase in the spontaneous mutation rate exhibited by these UDG knock-out mice. Thus, the investigators concluded that the nuclear UDG (UDG2) was not involved in post-replicative DNA repair to remove uracil residues that were formed in DNA by dCMP deamination but rather had evolved into a specialized role in mammalian cells, to specifically remove uracil residues that were formed in DNA during replication by the misincorporation of dUTP by DNA polymerases (239,240). A subsequent study using nuclear extracts from mammalian cells suggested, however, that UDG2 was, also, the major enzyme for removal of deaminated dCMP residues (364). This is consistent with data generated using UDG mutants of *E. coli* (284) and *Saccharomyces cerevisiae* (287) and a recent study in a human glioma cell line in which UDG was depleted using a retroviral vector expressing Ugi (289). Further studies to examine why there is a low spontaneous mutation frequency in UDG knock-out mice have not been performed.

The bacteriophage PBS2 encodes for a protein that specifically inhibits UDGs from a variety of sources (267,268). The uracil-DNA glycosylase inhibitor protein (UGI) is a heat stable, acidic, low molecular weight (84 amino acids, 9.5 kDa) protein that
interacts with UDGs in a stoichiometry of 1:1 to form a complex that does not dissociate under physiological conditions (269). UGI does not bind to the catalytic site, but rather to the conserved DNA binding groove of the UNGs (see Fig. 1.5), preventing the UDGs from binding to DNA (278-280). UGI has been employed in many in vitro studies to distinguish UDG activity from other enzymes that remove uracil residues from DNA and several studies have demonstrated the ugi gene can be successfully delivered to cultured cells and its expression results in the inhibition of UDG activity (271,289). However, because of the inherent problems associated with gene delivery systems and the problems of developing somatic cell knock-out mutants, I wanted to develop a mechanism that could be used to rapidly deplete a variety of cells of UDG activity.

The purpose of the present study was to determine whether intracellular UDG activity could be modulated with UGI introduced into cells using protein transduction technology. To accomplish these studies a novel system was developed for the production and purification of recombinant UGI coupled to the TaT- protein transduction domain. Using a three-step procedure, I have successfully purified a functionally active TaT-HA-UGI fusion protein. The purified protein is highly stable, with no loss of activity following storage for at least 15 days at room temperature. The results of this study demonstrate that TaT-Ugi is rapidly incorporated into a variety of human cells and that this results in a decrease in UDG activity.
4.2 Materials and Methods

4.2.1 Materials

Unless otherwise indicated all chemicals and chromatography matrixes were obtained from Sigma Chemical Company (St. Louis, MO). All chemicals were of the highest purity available. [5-3H]dUTP was purchased from Moravek Biochemicals Inc., Brea, CA. Centriprep-30 concentrators were obtained from Amicon (Beverly, MA). Whatman exchange filters were purchased from Fisher. Cell culture media and medium supplements were obtained from Invitrogen-GIBCO.

4.2.2 Construction of TaT-UGI

A 280 bp PCR generated double-stranded DNA fragment containing the ugi gene from bacteriophage PBS2 was sub-cloned in frame into the Ncol/EcoR1 site of the pTaT-HA vector, kindly provided by Dr. Steven Dowdy (Washington University School of Medicine). This construct was transfected into the E. coli strain BL21(DE3) pLysS following initial screening and DNA sequencing, a clone (pTaT-HA-Ugi) was selected for further studies. BL21 containing pTaT-HA-Ugi was grown and maintained in LB medium containing ampicillin (50 mg/ml) and chloroamphenicol (25 mg/ml).

4.2.3 TaT-UGI Purification

BL21 cells transduced with pTaT-HA-Ugi were routinely grown and maintained in LB medium containing ampicillin (50 mg/ml) and chloroamphenicol (25 mg/ml). For the induction of TaT-Ugi, exponentially growing cells were treated with 1 mM IPTG for 2 hrs at 37°C. Cells were collected by centrifugation at 4°C using a Sorvall RC-5B
centrifuge (5000 x g for 10 min). The cell pellets were suspended in extraction buffer containing 10 mM Tris-HCl (pH 8.0), 20 % (v/v) glycerol, 2 mM 2-mercaptoethanol, 1 mM MgCl2 and a protease inhibitor cocktail containing antipain, benzamidine, chymostatin, leupeptin, pepstatin, TLCK, and TPCK (final concentration 1 µg/ml). Cells were lysed by sonication (Branson sonifier Model 350, microtip setting 4) for 1 min in a series of 15-sec pulses separated by 30-sec cooling periods on ice. The resulting homogenate was clarified by centrifugation at 5000 x g for 10 min at 4°C. The resulting supernatant (designated crude extract) was used immediately for purification of TaT-UGI. Unless otherwise stated all remaining purification procedures were performed at 4°C.

The crude extract was applied to a Blue Sepharose affinity column (1.2 x 9 cm) equilibrated in 10 mM Tris-HCl (pH 7.5) (equilibration buffer). The column was washed with 5 bed volumes of equilibration buffer and TaT-UGI was eluted by increasing the ionic strength using a linear gradient of NaCl from 0.05 to 2 M in equilibration buffer. Fractions of 10 ml were collected and those fractions containing UGI activity were pooled. The pooled fraction following Blue Sepharose chromatography was boiled for 20 min. Following boiling the sample was centrifuged. The resulting supernatant was collected and assayed to confirm the presence of TaT-UGI using the standard assay.

The sample was then subjected to a series of concentration steps using a Centriprep-30 concentrator (Amicon; Beverly, MA). The concentration consisted of a series of centrifugations: (i) 20 minutes at 5000 x g, (ii) 15 minutes at 5000 x g, (iii) 10 minutes at 5000 x g. After each centrifugation step the filtrate, which corresponds to proteins and/or complexes with a molecular weight of less than 30,000 was collected.
Likewise, the retentate which represents proteins having a molecular weight of greater than 30,000 was collected following the three centrifugations. Upon completion of the centrifugation, the combined filtrates and the retentate were examined for UGI activity using the standard assay.

### 4.2.4 SDS-PAGE and Western Blotting

For these purification studies, purity was determined using SDS-PAGE on 15% gels as previously described (327). SigmaMarker™ low molecular weight range kit (M3913) was used for protein standards. Proteins were identified using the Sigma Stain Silver Kit (AG-25) as described by the manufacturer. For Western blotting, proteins were transferred to a Hybond™ ECL™ nitrocellulose membrane (Amersham Pharmacia Biotech; Little Chalfont, Buckinghamshire, England) using the protocol established by Bio-Rad (Richmond, CA). The membrane was incubated in blocking buffer (20 mM Tris-HCl (pH 7.6), 137 mM NaCl, 0.1% (v/v) Tween and 5% (w/v) nonfat dry milk) for 1 hr at room temperature on an orbital shaker. The membrane was subsequently washed twice with TBS-T (20 mM Tris-HCl (pH 7.6), 137 mM NaCl, and 0.1% (v/v) Tween) for 20 minutes at room temperature on an orbital shaker. A mouse monoclonal anti-HA IgG antibody (Covance, Berkeley, CA) diluted 1:1000 in blocking buffer was added and incubated for 2 hr at 4°C on an orbital shaker. The membrane was subsequently washed twice with TBS-T for 20 minutes at room temperature on an orbital shaker. An anti-mouse sheep whole antibody coupled to horseradish peroxidase (NA931, Amersham, Arlington Heights, IL) was diluted 1:5000 in blocking buffer and incubated with the membrane at room temperature for 1 hr on an orbital shaker. The membrane was
subsequently washed twice with TBS-T for 20 minutes at room temperature. The protein was detected by chemiluminescence using ECL horseradish peroxidase substrate (Amersham Pharmacia Biotech) and autoradiography.

4.2.5 Protein Assays

Protein concentrations were estimated using the Coomassie blue dye binding assay as described by Bio-Rad Laboratories using bovine serum albumin (BSA) as the standard.

4.2.6 UDG and UGI Assays

UDG activity was determined as previously described (271, 288). Briefly, the reaction mixture contained the following in a total volume of 0.2 ml: 50 mM Tris-HCl (pH 7.5), 100 µg of bovine serum albumin per ml, 3-5 µg of calf thymus double-stranded DNA containing [3H]uracil residues (specific activity 99µCi/µmole, 30-60 pmol of dUMP/µg DNA), 10 mM EDTA and the enzyme sample. After incubation at 37°C, the reactions were terminated and the acid-soluble radioactivity was determined as described previously (271). A unit of UDG activity was defined as the amount of enzyme required to release 1 pmol of uracil as trichloroacetic acid soluble material per minute at 37°C.

UGI activity was measured as previously described (271,288) using the UDG assay described above. A unit of UGI was defined as the amount of inhibitor required to neutralize a unit of *E. coli* BL21 UDG under our assay conditions.
4.2.7  Stability Studies

Purified TaT-UGI was subjected to various buffer and/or temperature conditions to analyze its stability. Stability of the fusion protein was defined as the percentage of UGI activity recovered compared to the initial UGI activity added. UGI activity was measured using the method described above. The conditions included incubation in 10 mM Tris-HCl (pH 7.5) at –20°C, 40°C, 25°C, 37°C and incubation in growth media at –20°C and 37°C. Samples were incubated for 10 days upon which the sample was analyzed for UGI activity.

4.2.8  dUTPase Enzyme Assays

The standard assay procedure used has been previously described (365). The reaction mixture contained in a total volume of 0.1 ml: 50 mM Tris-HCl (pH 8.0), 0.1 mM [5-3H]dUTP (18 µCi/µmol), 2 mM 2-mercaptoethanol (BME), 1 mM MgCl₂, 0.1% (w/v) bovine serum albumin (BSA), 2 mM ρ-nitrophenyl phosphate (PNP), 2 mM ATP and the enzyme (0.02-52 µg protein). PNP and ATP were included in the reaction mixture to prevent the hydrolysis of dUTP by alkaline phosphatase and ATPase. The reaction mixtures were incubated at 37°C for 1 hr. Reaction mixtures were terminated by spotting 50 µl of the reaction mixture on a DE81 filter disc (Whatman) and immediately washing the disc twice in a solution containing 4 M formic acid and 1 mM ammonium formate for 5 min, and once in 95% ethanol for 5 min. Liquid scintillation spectrophotometry using a Beckman LS 6000IC quantitated the amount of radioactivity bound to the disc. A unit of dUTPase activity was defined as the amount of enzyme that hydrolyzed 1 nmol of dUTP/min/ml at 37°C.
4.2.9 Cell Culture and TaT-UGI Treatment

SW620 and HT29 cell lines derived from human colon adenocarcinomas and HeLa cell line derived from a human cervical epithelial carcinoma were grown and maintained in modified McCoy’s 5A medium supplemented with 10% fetal bovine serum and 1% penicillin (10 units)-streptomycin (10 µg) at 37°C in a 5% CO₂ atmosphere. The pro-monocytic cell line, U-937, was grown and maintained in RPMI 1640 with 2 mM L-glutamine containing 10 mM HEPES, 1 mM sodium pyruvate, 4.5 g/L glucose, 1.5 g/L sodium bicarbonate and supplemented with 10% fetal bovine serum and 1% penicillin (10 units)-streptomycin (10 µg) at 37°C in a CO₂ atmosphere. TaT-UGI was obtained from the purification scheme described in this chapter of this manuscript. TaT-UGI was dialyzed against two changes of 4 liters 10 mM Tris-HCl (pH 7.5) for 24 hr at 4°C then filter-sterilized. SW620, HT29, and HeLa cells were grown in 6-well plates and seeded at an initial density of 1.25 x 10⁵ cells/well. U-937 cells were grown in 25 cm² culture flasks and seeded at an initial density of 1.25 x 10⁵ cells/ml. Cells were allowed to recover for at least two doubling times before addition of TaT-UGI (one doubling time = 22-24 hr). Purified TaT-UGI was added to exponentially growing cells at specified concentrations and time periods.

Cells were collected by centrifugation at 4°C using a Sorvall RC-5B centrifuge (5000 x g for 10 min). The cell pellets were suspended in extraction buffer containing 10 mM Tris-HCl (pH 8.0), 20 % (v/v) glycerol, 2 mM 2-mercaptoethanol (BME), 1 mM MgCl₂ and a protease inhibitor cocktail containing antipain, benzamidine, chymostatin, leupeptin, peptstatin, TLCK, and TPCK (final concentration 1 µg/ml). Cells were lysed by sonication (Branson sonifier Model 350, microtip setting 4) for 1 min in a series of 15-
sec pulses separated by 30-sec cooling periods on ice. The resulting homogenate was clarified by centrifugation at 5000 x g for 10 min 4°C and the supernatant was used for further studies.

4.2.10 FITC Labeling

Purified TaT-UGI and GST-UGI were labeled with fluorescein isothiocyanate (FITC) using the protocol established by Pierce (Rockford, IL). Briefly, 60 µg of the purified proteins were equilibrated in 50 mM borate buffer (pH 8.5) using a dextran desalting column. A twenty-four molar excess of FITC was added to each protein and the resulting mixture was incubated in the dark at room temperature for 1 hr. The labeled sample was then applied to a dextran desalting column and the protein was eluted with 0.1 M phosphate, 0.15 M NaCl (pH 7.2). 100 µl of the purified fusion protein was added to ~1.0 – 1.5 x 10⁶ U-937 cells in McCoy’s 5A media, supplemented with 10% FBS, 1% penicillin (10 units)-streptomycin (10 µg) (growth media). Cells were incubated with the labeled protein for 1 hr. Cells were fixed with 4% paraformaldehyde for 20 min at room temperature in the dark. Cells were then centrifuged down for 5 min at 2000 x g, washed twice with PBS, and resuspended in 40 µl PBS. 10 ul of the protein sample was put onto slide and air-dried for 2 hr. Slides were then stained with Evan’s Blue for 45 sec and washed thoroughly with ddH₂O. Cells were checked for transduction using fluorescence microscopy.
4.2.11 Cell Proliferation/Chemosensitivity Assays

To measure the cell proliferation/chemosensitivity following treatment with TaT-UGI the CellTiter 96® AQueous non-radioactive cell proliferation assay (Promega, Madison, WI) was used. SW620, HT29, and HeLa cell lines were grown and maintained in previously described growth conditions (sec. 4.2.9). Cells were grown in 96-well plates and seeded at an initial density of 1.00 X 10^4 cells/well. Cells were allowed to recover for 24 hr before the addition of purified TaT-UGI. Cell lines were exposed to TaT-UGI for 72 hr at the indicated doses. At the end of TaT-UGI exposure, the medium containing TaT-UGI was aspirated and 100 µl of fresh growth media was added. A mixture of the tetrazolium compound 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) and the electron coupling reagent phenazine methosulfate (PMS) was added to the culture medium and incubated for 90 min. at 37°C in a humidified CO₂ atmosphere. Dehydrogenase enzymes found in metabolically active cells convert MTS into aqueous, soluble formazan. The quantity of formazan product, which is directly proportional to the number of actively replicating cells in culture, was determined by the amount of absorbance measured at 490 nm. The absorbance at 490nm was measured using an ELISA plate reader (BioTek FL600 Microplate Fluorescence Reader). Chemosensitivity of TaT-UGI on U-937 cells was determined by trypan blue staining.
4.3 Results

4.3.1 Purification of TaT-UGI

The TaT-UGI fusion protein was purified using a combination of affinity chromatography, heat denaturation, and centrifugal filtering. A summary of a typical purification is shown in Table 4.1. Briefly, TaT-UGI bound to Blue Sepharose matrix and eluted at a NaCl concentration of approximately 1.0 M. Fractions containing UGI activity were combined and subjected to heat denaturation. Following boiling, denatured proteins were clarified by centrifugation and the resulting supernatant was collected and assayed for UGI activity. The supernatant, which exhibited UGI activity, was subjected to concentration using a Centriprep-30 concentrator. UGI activity was detected in the retentate (data not shown). SDS-PAGE analysis (Fig. 4.1) demonstrated that the preparation following concentration contained one protein species with a molecular weight of 11 kDa, which corresponded to the predicted molecular weight of TaT-UGI.

4.3.2 Biochemical Properties of TaT-UGI

Studies were conducted to determine the stability of the purified TaT-UGI fusion protein. The standard UGI assay was used to assess changes in TaT-UGI activity upon addition of 8 M urea or changes in temperature. A known amount of UGI activity (approximately 42,000 units/mg) was added to Tris-HCl (pH 7.5), Tris-HCl (pH 7.5) + 8M urea, and fetal bovine serum. The samples were incubated at −20, 4, 25, and 37°C for 10 days. After which they were assayed for UGI activity, stored at −20°C for 3 days, and then assayed for UGI activity once again. A summary of the results is shown in Table 4.2. TaT-UGI incubated at 4°C in Tris-HCl (pH 7.5) buffer with or without 8M urea
retained 91.1 ± 9.8 and 82.5 ± 18.3%, respectively of the input UGI activity. Whereas TaT-UGI incubated at –20, 25, or 37°C in Tris buffer retained 9.1 ± 0.1, 11.9 ± 0.2, and 12.0 ± 0.7%, respectively of its initial UGI activity. TaT-UGI incubated at 37°C in fetal bovine serum retained 94.5 ± 6.7% of its initial UGI activity. This amount increased to 101.1 ± 0.3% after storage at –20°C. Analysis of the other samples after storage at –20°C failed to detect any UGI activity.

4.3.3 Cellular UDG Activity

After extensive in vitro characterization of TaT-UGI I wanted to determine whether TaT-UGI could transduce cultured cells when added to the growth medium and ultimately inhibit UDG activity in intact cells. The ability of TaT-UGI to transduce cells was confirmed by FITC labeling TaT-UGI (data not shown). The ability to transduce cells was dependent upon TaT because FITC labeled GST-UGI did not transduce cells. To address the capability of TaT-UGI to inhibit cellular UDG, four cell lines were employed, the pro-monocytic cell line U-937, and the SW620, HT29, and HeLa cell lines. Initial characterization of U-937 treated with a single dose of TaT-UGI indicated that TaT-UGI inhibited cellular UDG in a concentration and time dependent manner (Fig. 4.2C). However, UDG levels recovered after 72 hours when compared to those of non-treated cells, except in those cell populations treated with 7000 and 14,000 units of TaT-UGI (Fig. 4.2C). Subsequent studies in U-937 cells in which TaT-UGI was added daily demonstrated that UDG levels were inhibited 94.41 ± 4.06% for up to 144 hours (length of experiment) for all concentrations of TaT-UGI used (Fig. 4.3D). The concentration and time dependent nature of TaT-UGI was also demonstrated in SW620 and HT29 cells.
The TaT-UGI fusion protein was purified as described in Materials and Methods from exponentially growing IPTG (1 mM) induced BL21 cells containing the pTaT-HA-ugi construct.

A unit of UDG activity was defined as the amount of enzyme required to release 1 pmol of uracil as trichloroacetic acid soluble material per minute at 37°C.

Specific activity represents units/mg.

### Table 4.1: Purification of TaT-UGI.

<table>
<thead>
<tr>
<th>Step</th>
<th>units&lt;sup&gt;b&lt;/sup&gt;/ml</th>
<th>mg/ml</th>
<th>Specific Activity&lt;sup&gt;c&lt;/sup&gt; (x10²)</th>
<th>Percent Recovery</th>
<th>Fold-Purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude Extract</td>
<td>39,506</td>
<td>4.19</td>
<td>94</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>Blue Sepharose Chromatography</td>
<td>21,300</td>
<td>0.25</td>
<td>852</td>
<td>159</td>
<td>9.06</td>
</tr>
<tr>
<td>Heat Denaturation</td>
<td>5238</td>
<td>0.032</td>
<td>1638</td>
<td>38</td>
<td>17.43</td>
</tr>
<tr>
<td>Centricon -30K</td>
<td>52,859</td>
<td>0.24</td>
<td>2202</td>
<td>48</td>
<td>23.43</td>
</tr>
</tbody>
</table>

<sup>a</sup> The TaT-UGI fusion protein was purified as described in Materials and Methods from exponentially growing IPTG (1 mM) induced BL21 cells containing the pTaT-HA-ugi construct.

<sup>b</sup> A unit of UDG activity was defined as the amount of enzyme required to release 1 pmol of uracil as trichloroacetic acid soluble material per minute at 37°C.

<sup>c</sup> Specific activity represents units/mg.
Figure 4.1: SDS-PAGE. SDS-PAGE was performed on 15% gels as described in Materials and Methods. *Lane 1.* Protein Standards: aprotinin (6.5 kDa), lactalbumin (14.2 kDa), soybean trypsin inhibitor (20 kDa), trypsinogen (24 kDa). *Lane 2.* TaT-UGI purified as described in Materials and Methods from IPTG induced BL21 cells containing the pTaT-HA-ugi construct. Approximately 5 µg of the protein from the UGI active fraction following concentration was analyzed. Proteins were visualized by silver staining as described in Materials and Methods.
Purified TaT-UGI was incubated in the following conditions for 10 days in the designated conditions then stored at –20°C for 3 days: 10 mM Tris-HCl (pH 7.5), 8M Urea in 10 mM Tris-HCl (pH 7.5), and dialyzed fetal bovine serum (serum). % Activity (units) was determined by comparing UGI activity of the treated TaT-UGI samples to the UGI activity associated with TaT-UGI prior to treatment. UGI activity associated with TaT-UGI prior to treatment was 420,000 units/mg. This value was standardized at 100% activity. Enzymatic assays were performed as described in Materials and Methods. A unit of UGI was defined as the amount of inhibitor required to neutralize a unit of *E. coli* BL21 UDG under our assay conditions. Values represent the average of two independent experiments.

<table>
<thead>
<tr>
<th>Temp (°C)</th>
<th>Buffer</th>
<th>% Activity Before Storage</th>
<th>% Activity After Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>-20</td>
<td>Tris-HCl</td>
<td>9.1 ± 0.1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Tris-HCl</td>
<td>82.5 ± 18.3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>8M Urea</td>
<td>91.1 ± 9.8</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>Tris-HCl</td>
<td>11.9 ± 0.2</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>Tris-HCl</td>
<td>12.0 ± 0.7</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>Serum</td>
<td>94.5 ± 6.7</td>
<td>101.1 ± 0.3</td>
</tr>
</tbody>
</table>

Table 4.2: Effect of Temperature and Buffer Conditions on the Stability of Purified TaT-UGI.
treated with TaT-UGI (Figs. 4.4A,B). The greatest inhibition of cellular UDG activity, approximately $78.89 \pm 1.00$ and $69.92 \pm 5.00\%$ in SW620 and HT29 cells, respectively, occurred when 8000 units of TaT-UGI was added for three consecutive days (Fig. 4.4B). A similar treatment protocol performed in HeLa cells inhibited cellular UDG activity $29.64 \pm 3.55\%$ (Fig. 4.4B).

### 4.3.4 Cellular Protein Concentration

In the previous experiments assessing the ability of TaT-UGI to inhibit cellular UDG activity it became apparent that the total cell protein concentration was also affected. In U-937 cells receiving one treatment of TaT-UGI there was a small initial decrease in the protein concentration during the first 48 hours after treatment with 7000 and 14,000 units, but the levels in these cell populations recovered to those of non-treated after 72 hours (Fig. 4.2B). In contrast to the early results (e.g. 24 hours) associated with a single treatment of 7000 and 14,000 units of TaT-UGI there was no measurable difference in the protein concentrations of U-937 cells treated daily with TaT-UGI and non-treated controls (Fig. 4.3B). Similar results demonstrating no significant difference in the total cell protein concentration between treated and non-treated cells were seen in HeLa cells treated with 8000 units of TaT-UGI for three consecutive days (Fig. 4.5A). The protein concentrations in SW620 and HT29 cells decreased, however, when treated with TaT-UGI (8000 units) (Fig. 4.5A). To determine whether TaT-UGI treatment affected other enzymatic activities, dUTPase activity was measured. In U-937 cells, treatment with TaT-UGI at a concentration below 7000 units had no affect on dUTPase
activity but treatment with 14,000 units resulted in a 48.13 ± 10.26% decrease. Conversely, treatment of SW620 and HeLa cells with 8000 units reduced dUTPase activity 83.99 ± 11.84 and 95.57 ± 5.50%, respectively (Figs. 4.3C, 4.5B).

4.3.5 Chemosensitivity Associated with TaT-UGI

The discovery that certain concentrations and number of treatments of TaT-UGI decreased total cell protein concentration in certain cell types lead me to investigate whether TaT-UGI had any affect on cell proliferation. Cell lines were treated with varying concentrations of TaT-UGI for three consecutive days and then compared to the cell numbers of non-treated controls. Calculation of viable U-937 cells using trypan blue staining indicated no difference in the treated versus non-treated populations (Fig. 4.3A). Likewise, using a colormetric assay that allowed me to approximate the number of viable cells in a sample population, I determined that TaT-UGI had no significant effect on HeLa cell proliferation when compared to non-treated controls (Fig. 4.6, Table 4.3). Conversely, treatment of SW620 and HT29 cells with TaT-UGI resulted in a dose-dependent inhibition of cellular proliferation, with maximum inhibition (approximately 30%) following multiple treatments with 10,000 units (Fig. 4.6, Table 4.3). The results of these studies are summarized in Table 4.4.
Figure 4.2: Effects of a Single Dose of TaT-UGI on U-937 Cells. Purified TaT-UGI was added to exponentially growing cells at the indicated concentrations. Cells were collected at 24 hr intervals after the initial administration and cell extracts were prepared as described in Materials and Methods. (A) Cell number was quantitated by trypan blue staining; (B) protein concentration, and (C) UDG activity was determined as described in Materials and Methods. The values represent the average of three independent experiments.
Figure 4.3: Effects of a Multiple Dose of TaT-UGI on U-937 Cells. Purified TaT-UGI was added to exponentially growing cells (5.00 x 10^5) at the indicated concentrations at 24 hr intervals. Cells were collected at 24 hr intervals after the initial administration and cell extracts were prepared as described in Materials and Methods. (A) Cell number was quantitated by trypan blue staining; (B) protein concentration, (C) dUTPase activity, and (D) UDG activity was determined as described in Materials and Methods. The values represent the average of three independent experiments.
Figure 4.3
Figure 4.4: Effects of TaT-UGI Dosage on UDG Activity in SW620, HT29, and HeLa Cells. (A) Purified TaT-UGI was added to exponentially growing cells (5.00 x 10^5) at the indicated concentrations. Cell extracts were prepared 24 hr after TaT-UGI administration and UDG activity was determined as described in Materials and Methods. (B) 8000 units of purified TaT-UGI were added to exponentially growing cells (5.00 x 10^5). Treatment groups either received one dosage (single dose), two dosages (double dose) or three dosages (triple dose). The dosages occurred at 24 hr intervals. Cell extracts were prepared for each treatment group 24 hr after the final dosage and UDG activity was determined. 1 unit of TaT-UGI represents the amount of purified protein needed to inhibit 1 unit of *E. coli* BL21 UDG. The percent UDG activity remaining was determined by comparing UDG activity in the treated population to non-treated controls. UDG activity in the control non-treated population was 1464, 1180, and 810 units/mg for SW620, HT29, and HeLa cells, respectively. These values were standardized at 100% activity. The values represent the average of three independent experiments.
Figure 4.4
Figure 4.5: Cellular Effects of TaT-UGI on SW620, HT29, and HeLa Cells. (A) Cellular protein concentration and was determined after 8000 units of purified TaT-UGI was added to exponentially growing cells (5.00 x 10^5). Treatment groups either received one dosage (single dose), two dosages (double dose) or three dosages (triple dose). The dosages occurred at 24 hr intervals. Cell extracts were prepared for each treatment group 24 hr after the final dosage and UDG activity was determined. Protein concentration was determined as described in Materials and Methods. (B) Cellular dUTPase activity was determined after exponentially growing cells were treated with three consecutive dosages occurring every 24 hr of 8000 units of purified TaT-UGI. Cell extracts were prepared for each treatment group 24 hr after the final dosage and dUTPase activity was determined as described in Materials and Methods. The values represent the average of three independent experiments.
A

Protein (mg/ml)

Control

Single Dose

Double Dose

Triple Dose

SW620 HT29 HeLa

B

dUTPase Activity (units/mg)

Control

TaT-UGI

SW620 HeLa

Figure 4.5
Figure 4.6: Effect of TaT-UGI on Cell Proliferation. Purified TaT-UGI was added to exponentially growing cells (1.00 x 10⁴) at the indicated concentrations at 24 hr intervals for a total of three dosages. Twenty-four hr after the final TaT-UGI administration a solution containing (MTS/PMS) was added to the cell culture. Cultures were incubated 90 min in a 37°C, CO₂ humidified incubator after which the absorbance at 490nm was recorded using an ELISA plate reader. Cell number values were extrapolated from a standard curve established from the absorbances of known cell concentrations. The values represent the average of four independent experiments.
<table>
<thead>
<tr>
<th>TaT-UGI (units)</th>
<th>HeLa</th>
<th>SW620</th>
<th>HT29</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100 ± 4.8</td>
<td>100 ± 3.0</td>
<td>100 ± 2.1</td>
</tr>
<tr>
<td>2000</td>
<td>109.5 ± 11.6</td>
<td>99.2 ± 3.6</td>
<td>96.8 ± 10.9</td>
</tr>
<tr>
<td>4000</td>
<td>104.7 ± 11.9</td>
<td>93.3 ± 8.7</td>
<td>91.6 ± 11.8</td>
</tr>
<tr>
<td>6000</td>
<td>106.8 ± 15.9</td>
<td>85.5 ± 12.8</td>
<td>84.6 ± 10.3</td>
</tr>
<tr>
<td>8000</td>
<td>96.9 ± 3.9</td>
<td>76.4 ± 11.7</td>
<td>80.0 ± 1.9</td>
</tr>
<tr>
<td>10,000</td>
<td>83.4 ± 9.5</td>
<td>69.6 ± 9.5</td>
<td>63.5 ± 4.9</td>
</tr>
</tbody>
</table>

* Purified TaT-UGI was added to exponentially growing cells (1.00 x 10⁴) at the indicated concentrations at 24 hr intervals for a total of three dosages. 24 hr after the final TaT-UGI administration a solution containing (MTS/PMS) was added to the cell culture. Cultures were incubated 90 min in a 37°C, CO₂ humidified incubator after which the absorbance at 490nm was recorded using an ELISA plate reader.

* The percent control was determined by comparing the cell number in the treated population to non-treated controls. Cell number in the control non-treated population was 48,262 ± 1452, 37,788 ± 809, and 21,317 ± 1031 for SW620, HT29, and HeLa cells, respectively. These values were standardized at 100% activity. The values were extrapolated from a standard curve established from the absorbances of known cell concentrations. The values represent the average of four independent experiments.

**Table 4.3: Effect of TaT-UGI on Cell Proliferation.**
### Table 4.4: Summary of the Cellular Effects Associated with TaT-UGI Administration.

<table>
<thead>
<tr>
<th>Cell Type</th>
<th>Protein (mg)</th>
<th>UDG Activity</th>
<th>dUTPase Activity</th>
<th>Proliferation</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-937&lt;sup&gt;b&lt;/sup&gt;</td>
<td>100.36 ± 2.53</td>
<td>5.59 ± 4.06</td>
<td>51.87 ± 10.26</td>
<td>93.03 ± 17.54</td>
</tr>
<tr>
<td>SW620&lt;sup&gt;c&lt;/sup&gt;</td>
<td>31.81 ± 1.87</td>
<td>21.11 ± 1.00</td>
<td>16.07 ± 11.84</td>
<td>76.38 ± 11.74</td>
</tr>
<tr>
<td>HeLa&lt;sup&gt;d&lt;/sup&gt;</td>
<td>80.22 ± 2.52</td>
<td>70.36 ± 3.55</td>
<td>4.53 ± 5.50</td>
<td>96.88 ± 3.9</td>
</tr>
<tr>
<td>HT29&lt;sup&gt;e&lt;/sup&gt;</td>
<td>24.97 ± 3.22</td>
<td>30.08 ± 5.00</td>
<td>N.D.</td>
<td>80.01 ± 1.9</td>
</tr>
</tbody>
</table>

<sup>a</sup> The percent control was determined by comparing the values generated in the treated population to non-treated controls. Cells were treated daily with approximately 8000 units/ml TaT-UGI for 72 hr. The non-treated values were normalized at 100%.

<sup>b</sup> The following are the values obtained from non-treated controls: protein concentration (0.187 ± 0.013 mg/ml), UDG activity (365.98 ± 60.62 units/mg protein), dUTPase activity (6.499 ± 0.845 units/mg protein), cell number (3.04 x 10<sup>6</sup> ± 3.31 x 10<sup>5</sup>).

<sup>c</sup> The following are the values obtained from non-treated controls: protein concentration (0.83 ± 0.06), UDG activity (1018 ± 187.5 units/mg protein), dUTPase activity (8.06 ± 1.02 units/mg protein), cell number (48,262 ± 152).

<sup>d</sup> The following are the values obtained from non-treated controls: protein concentration (0.51 ± 0.04 mg/ml), UDG activity (810.64 ± 98.01 units/mg protein), dUTPase activity (16.23 ± 2.55 units/mg protein), cell number (21,317 ± 1031).

<sup>e</sup> The following are the values obtained from non-treated controls: protein concentration (1.00 ± 0.14 mg/ml), UDG activity (467 ± 114.69 units/mg protein), dUTPase activity (Not Determined (N.D.)), cell number (37,788 ± 809).
4.4 Discussion

Numerous obstacles exist when attempting to express a foreign protein in a cell. Some of which, include: (i) efficient transfection of the target cells with a construct carrying the gene that encodes for the protein, (ii) retention in the cell of this construct, and most importantly (iii) proper expression of your protein, and (iv) length of time of expression. An alternative approach is to deliver the protein directly to the cell. It was initially demonstrated by Fawell et al (366), that the Tat protein encoded for by HIV-1 could be used to deliver heterologous proteins into cells. Dr. Steven Dowdy (Washington University School of Medicine) developed a vector (pTaT-HA), containing the protein transduction domain of TaT that could be used for delivering heterologous proteins to cells in a process termed “protein transduction” (367-369). This construct has been used to deliver a variety of heterologous proteins to cultured cells, tissues and mice (367-369).

I hypothesized that such a system could be used to elucidate the role of uracil DNA glycosylase (UDG) in various biological environments, including HIV-1 and colon cancer. The heterologous protein created to study the role and/or importance of UDG consisted of the uracil DNA glycosylase inhibitor (UGI) fused in frame to pTaT-HA (pTaT-HA-UGI). In order to achieve efficient transduction of cells the Tat fusion protein must undergo a series of denaturing and renaturing steps during its purification. Because of the time involved and propensity to lose the majority of the fusion protein an alternative purification approach was undertaken. Using a three-step procedure, I have successfully purified a functionally active TaT-HA-UGI fusion protein. Generally, 0.5-1.0 mg of TaT-UGI is recovered with an average specific activity of $2 \times 10^5$
The purified protein is highly stable, with no loss of activity following storage for 15 days (last day tested) at room temperature, incubation in 8M urea and boiling for 30 min.

To determine whether TaT-UGI could inhibit UDG activity in intact cells the promonocytic cell line U-937, the cell lines SW620 and HT29, which were derived from adenocarcinoma of the colon, and the HeLa cell line, which is derived from a human cervical epitheloid carcinoma were employed. Table 4.4 summarizes the cellular effects associated with TaT-UGI treatment. Initial studies in U-937 cells demonstrated that TaT-UGI inhibited UDG in intact cells in a dose and time dependent manner (Figs. 4.2C and 4.3D). In U-937 cells treated with a single exposure of TaT-UGI, as little as 1400 units of TaT-UGI reduced UDG activity 47.4% 24 hr following treatment, while 5600 units of TaT-UGI completely inhibited UDG 24 hrs following treatment when compared to non-treated cells. UDG activity was restored in cells by 72 hr after treatment with 2800 units or less of TaT-UGI. Moreover, while cells treated with at least 5600 units of TaT-UGI did recover some UDG activity, it was still significantly decreased (approximately 70%) in treated cells 144 hr following treatment (Fig. 4.2C). Conversely, daily addition of TaT-HA-UGI to U-937 cells resulted in the depletion of cellular UDG activity for 6 days (length of experiment) without having any affect on growth rate, protein concentration, and dUTPase levels (Figs. 4.3A-D).

Subsequent studies in SW620, HT29 and HeLa cells demonstrated that TaT-UGI also inhibited UDG in a dose and time dependent manner (Fig. 4.4). Maximum inhibition occurred when cells were treated for three consecutive days with 8000 units of TaT-UGI. Using this treatment protocol 78.9, 69.9 and 29.6% of UDG activity was
inhibited in SW620, HT29 and HeLa cells, respectively (Fig. 4.4B). In contrast to the results observed in U-937 cells there was a significant decrease in the protein concentration in SW620 and HT29 cells and a significant decrease in the dUTPase levels in SW620 and HeLa cells (Fig. 4.5). Furthermore, there was a slight decrease in the proliferation of SW620 (23.62 ± 11.74) and HT29 (19.99 ± 1.9) cells (Fig. 4.6, Table 4.3). The differences that were observed in the various cell lines examined can be explained by differences in the transduction efficiency of TaT-UGI, the replication status of the cells, levels of UDG protein, and, also, the rate of UDG expression.

A somewhat surprising result was the global effect that treatment of cells with high amounts (>7000 units) of Tat-UGI had on protein synthesis, dUTPase activity and cell proliferation. Daily treatment all cells with TaT-Ugi (<7000 units) for 72 hours resulted in a statistically significant decrease in UDG and dUTPase activities in all cell lines tested, a statistically significant decrease in the total protein concentration in SW620, HeLa and HT29 cells and a statistically significant decrease in the proliferation of SW620 and HT29 cells. There have been many studies using the TaT domain to transduce protein into cells and while these studies have been designed towards demonstrating a gain of function, there have been no reports that the TaT domain has a negative affect on cells. Likewise, there have been numerous studies on UGI demonstrating that it is a specific inhibitor of UDG. There are at least two mechanisms that could possibly explain my results. (i) Since UDG and dUTPase function in tandem to exclude uracil residues from DNA, depletion of UDG activity with UGI may negatively regulate dUTPase, resulting in an increased misincorporation of dUTP into DNA. While dUTPase and UDG operate in tandem, there is no evidence to support this
premise. Furthermore, this would not explain the decline in protein levels in the treated cells. (ii) Depletion of UDG by UGI results in cell cycle arrest/decreased cell proliferation.

Recently, Tinkelenberg et al (370) reported that in *Saccharomyces cerevisiae*, UDG deletion mutants and isogenic strains that were depleted of UDG by the expression of UGI arrested in G2-M phase of the cell cycle when the cells were treated with aminopterin and sulfanilamide. The investigators suggested that this might represent a novel mechanism by which arrest at the G2-M checkpoint is signaled by the presence of uracil residues in DNA. Since the nuclear isoform of dUTPase is cell cycle dependent (294), the novel mechanism described above is attractive because it could explain the decrease in the protein levels and also the decrease in dUTPase activity that I observed following exposure of cells to high amounts of TaT-UGI. However, there are some inconsistencies that cannot be easily explained. My studies were performed in the absence of anti-folate or thymidylate synthase inhibitors. Thus, under the conditions of my experiments there should not be an increased incorporation of dUTP into DNA in the TAT-UGI treated cells. While the inhibition of UDG by TaT-UGI would increase the retention of uracil residues in DNA, which could signal DNA checkpoint arrest, it seems unlikely that this occurred based upon what is presently known concerning the rate of dUTP misincorporation and/or dCMP deamination. Furthermore, my results are not consistent with what has been reported when UGI was expressed in *S. cerevisiae* (370) and in a glioma cell line (289). Although the level of expression of UGI in *S. cerevisiae* was not reported, the level of expression in four established glioma cell lines expressing UGI ranged from 2400 to 19,800 units (based upon my definition of a unit of UGI)
activity). Moreover, neither study presented any data suggesting that the expression of UGI resulted in a decrease in cellular proliferation. In fact, it was reported that the growth rate and cell cycle distribution of the UGI expressing glioma cells did not differ from the parental cell (289). Furthermore, there has not been any data presented to indicate that the lack of expression of UDG in cells established from UDG knock-out mice exhibit altered cellular proliferation even though the steady state level of uracil increased to approximately 2000 residues per genome.

It is possible that these inconsistencies reflect differences in the cells or organisms used in these studies and/or technical problems associated with such studies. The study in *S. cerevisiae* demonstrated that UGI expressing cells were more resistant to aminopterin and sulfanilamide than the parental strain, which suggest that UGI is inhibiting UDG activity in intact cells (370). However, none of the studies over-expressing UGI have demonstrated, conclusively, that UGI inhibited UDG activity in intact cells, or did they demonstrate cellular localization of the expressed UGI. It is possible that the UGI expressed in these cells was retained in the cytoplasm and only inhibited UDG activity following extraction of the cells. Likewise, further analysis has demonstrated that subcellular sorting and expression of mouse UDG differs from human UDG and it has been suggested that such differences may contribute to the phenotype of UDG knock-out mice (371). While it is obvious from this discussion that additional studies are required, my results are consistent with the feature that the depletion of UDG activity with UGI results in a decrease in cellular proliferation, but not cell death at least during the time frame of my studies. While the role of UDG2 in pre- and post-replicative DNA repair is fairly well established, the interaction of UDG2 with other proteins
involved in these processes remains to be established as well as the mechanism(s) by which depletion of UDG2 results in decreased cellular proliferation and cell senescence. However the results of my study support the premise that UDG can be used as a target for the development of specific antiviral and cancer chemotherapeutic agents and that TaT-UGI can be used to elucidate the role of UDG in cellular metabolism.
CHAPTER 5

Down-Regulation of Human Deoxyuridine Triphosphate Nucleotidohydrolase (dUTPase) using Small Interfering RNA (siRNA)

5.1 Introduction

The physiological role of deoxyuridine triphosphate nucleotidohydrolase (dUTPase, EC 3.6.1.23), which is well established in prokaryotic and eukaryotic organisms, is to keep the intracellular dUTP pool level low by hydrolyzing dUTP to dUMP and pyrophosphate. Since DNA polymerases do not exhibit a significant preference for dTTP over dUTP (243) a low intracellular pool of dUTP reduces the likelihood that DNA polymerases will incorporate dUTP instead of dTTP. Expression of dUTPase is essential for the replication of Escherichia coli (296) and Saccharomyces cerevisiae (297). Furthermore, recent studies in cultured human cells and in S. cerevisiae have demonstrated that increased expression of dUTPase results in the increased resistance to certain antifolates (370,372,373). It has been proposed that dUTPase may be a potential target for the development of specific agents that could be useful for the treatment of infections caused by several microorganisms and also in cancer chemotherapy. However, few inhibitors have been described that target dUTPase (374-379).
While there is considerable information known concerning the biochemical and structural properties of dUTPases from eukaryotic cells, studies on dUTPase in human cells have been hampered because cells deficient in dUTPase activity have not been isolated or constructed. Human dUTPase is a member of the homotrimeric family (Fig. 5.1) of dUTPases that include dUTPases from most eukaryotes, prokaryotes and RNA viruses (reviewed in 307). The homotrimeric, as well as the monomeric, dUTPases contain five conserved domains that contribute to the formation of the catalytic site of the enzyme. Recent studies, using a process termed RNA interference (Fig. 5.2), have demonstrated that specific gene expression can be down regulated by introducing small double stranded RNA (siRNA) molecules (20-23 bp) into cells that target a homologous sequence in mRNA (380-385). The purpose of this study was to determine whether dUTPase activity could be down regulated using siRNA technology.

Transfection of a small interfering double stranded RNA molecule (siRNA, 21 bp) synthesized to a portion (nucleotides 337 to 357) of domain 3 of the human dUTPase (siRNAdUT3) into HeLa and HT29 cells decreased dUTPase activity approximately 50%, while transfection of SW620 cells decreased dUTPase activity approximately 27% when compared to non-treated controls. Subsequent northern blot analysis confirmed that dUTPase specific mRNA levels were decreased compared to non-treated controls. There was no effect on UDG activities in transfected cells when compared to non-transfected controls, thus demonstrating that the decrease in enzymatic and mRNA expression was specific for dUTPase. siRNAdUT3 decreased cellular proliferation and this inhibition could be enhanced when siRNAdUT3 was used in combination with TaT-UGI, a specific inhibitor of uracil-DNA glycosylase.
Figure 5.1: Homotrimeric dUTPase.
RNAi is initiated by the Dicer enzyme, which cleaves dsRNA into 21- to 23-nt siRNAs. siRNAs are complexed with a large multiprotein complex, the RISC. RISC is thought to unwind siRNA, which helps target the appropriate mRNA. The siRNA/mRNA hybrid signals mRNA degradation by endo and exonucleases.

Figure 5.2: Model for the RNAi pathway. RNAi is initiated by the Dicer enzyme, which cleaves dsRNA into 21- to 23-nt siRNAs. siRNAs are complexed with a large multiprotein complex, the RISC. RISC is thought to unwind siRNA, which helps target the appropriate mRNA. The siRNA/mRNA hybrid signals mRNA degradation by endo and exonucleases.
5.2 Materials and Methods

5.2.1 Materials

Double stranded RNA (21 bp) containing the sequences (GAUUAUAGAGGAAAUGUUG)D(TT) and r(CAACAUUUCCCUAUAAUC)d(TT) (siRNA)dUT3 was synthesized by Qiagen-Xeragon (Germantown, MD). [5-3H]dUTP was purchased from Moravek Biochemicals Inc. (Brea, CA). DE81 Whatman ion exchange filters were purchased from Fisher. Cell culture media and medium supplements were obtained from Invitrogen-GIBCO. Filter sterilized TaT-UGI was obtained from the purification scheme described in chapter 4 of this manuscript.

5.2.2 UDG and UGI Assays

UDG activity was determined as previously described (271, 288). Briefly, the reaction mixture contained in a total volume of 0.2 ml: 50 mM Tris-HCl (pH 7.5), 100 µg of bovine serum albumin per ml, 3-5 µg of calf thymus double-stranded DNA containing [3H]uracil residues (specific activity 99µCi/µmole, 30-60 pmol of dUMP/µg DNA), 10 mM EDTA and the enzyme sample. After incubation at 37°C, the reactions were terminated and the acid-soluble radioactivity was determined as described previously (271). A unit of UDG activity was defined as the amount of enzyme required to release 1 pmol of uracil as trichloroacetic acid soluble material per minute at 37°C.

UGI activity was measured as previously described (271,288) using the UDG assay described above. A unit of UGI was defined as the amount of inhibitor required to neutralize 1 unit of E. coli BL21 UDG under our assay conditions.
5.2.3 Protein Assays

Protein was estimated using the Coomassie blue dye binding assay as described by Bio-Rad Laboratories using bovine serum albumin (BSA) as the standard.

5.2.4 dUTPase Enzyme Assays

The standard assay procedure used has been previously described (365). The reaction mixture contained in a total volume of 0.1 ml: 50 mM Tris-HCl (pH 8.0), 0.1 mM [5-^3^H]dUTP (18 $\mu$Ci/$\mu$mol), 2 mM 2-mercaptoethanol (BME), 1 mM MgCl$_2$, 0.1% (w/v) bovine serum albumin (BSA), 2 mM $\rho$-nitrophenyl phosphate (PNP), 5 mM ATP, and the enzyme (0.02-52 $\mu$g protein). PNP and ATP were included in the reaction mixture to prevent the hydrolysis of dUTP by alkaline phosphatase and ATPase. The reaction mixtures were incubated at 37°C for 1 hr. Reaction mixtures were terminated by spotting 50 $\mu$l of the reaction mixture on a DE81 filter disc and immediately washing the disc in a solution containing 4 M formic acid and 1 mM ammonium formate for 5 min, twice. Liquid scintillation spectrophotometry using a Beckman LS 6000IC quantitated the amount of radioactivity bound to the disc. A unit of dUTPase activity was defined as the amount of enzyme that hydrolyzed 1 nmol of dUTP/min/ml at 37°C.

5.2.5 Preparation of siRNA

Double-stranded siRNA_dUT3 was resuspended in buffer containing: 100 mM potassium-acetate, 30 mM Hepes-KOH, and 2 mM Mg-acetate (pH 7.4). The solution was heated to 90°C for 1 min then incubated at 37°C for 60 min. The solution was stored at –20°C until use.
5.2.6 Transfection of Cells with siRNA-dUT3

SW620, HT29, and HeLa cells were grown and maintained in modified McCoy’s 5A medium supplemented with 10% fetal bovine serum and 1% penicillin (10 units)-streptomycin (10 µg) at 37°C in a CO₂ atmosphere. Transfection of these cells with siRNA-dUT3 followed the protocol described for the TransMessenger™ Transfection Reagent (Qiagen). Briefly, siRNA-dUT3 was added to a reaction mixture containing Enhancer R diluted in Buffer EC-R. A RNA (µg) to Enhancer R (µl) ratio of 1:2 was held constant. The reaction mixture was incubated at room temperature for 5 min. Either TransMessenger Transfection reagent (Qiagen) or Metafectene (Biontex Laboratories) was added to the reaction mixture at a ratio of 4 µl of reagent: 1 µl RNA and incubated at room temperature for 15 min. Cells that had been plated at a density of 4.0 X 10⁵ cells/well (6-well plates) 24 hr prior to siRNA transfection were washed twice with sterile phosphate-buffered saline (PBS). Growth medium (without antibiotics and serum; 900 µl) was then added to the transfection reaction mixtures. Transfection reaction mixtures were added to cells (70-80% confluent, 0.8 to 1.0 x 10⁶) for 4 hr at 37°C in a CO₂ atmosphere. The transfection complexes were removed from the cells, washed once with PBS, and then 2 ml of normal growth medium containing serum and antibiotics was added. Forty-eight hours following transfection, cells were harvested, resuspended in general extraction buffer [10 mM Tris-HCl (pH 7.5), 2 mM MgCl₂, 1 mM 2-mercaptoethanol and 20% v/v glycerol, lysed by sonication (Branson sonifier Model 350, microtip setting 4) for 1 min in a series of 15-sec pulses separated by 30-sec cooling
periods on ice, and centrifuged at 4°C for 5 min. at 14,000 x g. The supernatants were used immediately for the determination of dUTPase and uracil DNA glycosylase activities.

5.2.7 Northern Blot Hybridization

RNA was isolated from HeLa and SW620 cells by the acid guanidinium isothiocyanate phenol chloroform extraction method previously described by Chomczynski and Sacchi (386). RNA (15 µg/lane) was size fractionated in 1% formaldehyde agarose gels and RNA transferred by capillary blotting onto Hybond-N+ membranes (Amersham, Arlington Heights, IL). Membranes were hybridized overnight at 42°C in 50% formamide, 5X SSPE, 5X Denhart’s solution, 0.50% SDS, 10% dextran sulfate, 100 µg/ml herring sperm DNA, and 1 x 10⁶ cpm/ml of radiolabeled probe. The membranes were washed twice at 42°C in 2X SSPE, 0.50% SDS followed by a high stringency wash at 65°C in 0.5X SSPE, 0.50% SDS. The β-actin probe was isolated from a macrophage cDNA library by screening with a mouse β-actin oligonucleotide probe (Oncogene Science, Manhasset, N.Y). The human dUTPase probe was a 500 bp BamH1/EcoR1 fragment obtained from plasmid pGEX2ThdUT, which was kindly provided by E. M. McIntosh. Gel purified insert DNA as well as β -actin were radiolabeled with ³²P-dCTP by the High Prime solution (Roche, Manheim, Germany). Membranes were exposed for 90 min. to Kodak BioMax MR film and autoradiographs were scanned with an Epson scanner and quantified using SigmaScan Pro 4 (SPSS, Chicago, IL). To account for differences in loading, dUTPase specific signals were normalized with the β -actin signal.
5.2.8 Cell Proliferation/Cytotoxicity Assays

To measure the cytotoxicity associated with siRNA\textit{dUT3} and siRNA\textit{dUT3/TaT-UGI} the CellTiter 96® AQueous non-radioactive cell proliferation assay (Promega, Madison, WI) was used. SW620, HT29 and HeLa cells were grown in 96-well plates in previously described growth conditions (sec. 5.2.6). Cells were seeded at an initial density of 1.00 X 10^4 cells/well and allowed to recover for 24 hr before the addition of siRNA (0.25 \( \mu \)g) and/or purified TaT-UGI (4000 units/ml). After an additional 72 hr the medium was aspirated and 100 \( \mu \)l of fresh growth media was added. A mixture of the tetrazolium compound 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) and the electron coupling reagent phenazine methosulfate (PMS) was added to the culture medium and incubated for 90 min. at 37°C in a humidified CO₂ atmosphere. Dehydrogenase enzymes found in metabolically active cells convert MTS into aqueous, soluble formazan. The quantity of formazan product, which is directly proportional to the number of viable cells in culture, was determined by the amount of absorbance measured at 490 nm. The absorbance at 490nm was measured using an ELISA plate reader (BioTek FL600 Microplate Fluorescence Reader).

5.3 Results

5.3.1 siRNA\textit{dUT3}-Mediated Inhibition of dUTPase

dUTPase expression has been associated with decreased responsiveness of certain solid tumors to chemotherapeutic agents (372,373). Furthermore, dUTPase expression had been shown to be critical for virulence and infectivity of viruses (127-130,300). Because dUTPase maybe an essential gene in mammalian cells it has not been possible to
critically analyze the effects of a dUTPase deficient cell. While several compounds have been developed that inhibit purified dUTPases (374-379), none of these compounds have been shown to selectively inhibit dUTPase in cultured cells or in vivo. Earlier studies have demonstrated that introduction of small double stranded RNAs have the ability to decrease gene expression. To evaluate if such an approach could specifically block dUTPase expression a small double stranded RNA corresponding to a portion of the conserved domain 3 of the human dUTPase (siRNA$d_{UT3}$) was synthesized. As shown in Fig 5.3, siRNA$d_{UT3}$ utilizing the Qiagen transfection reagent decreased cellular dUTPase activity in a dose dependent manner. The greatest decrease (56 ± 3.6%) in dUTPase activity occurred in HeLa cells transfected with 3 µg of siRNA$d_{UT3}$ and in HT29 cells (56 ± 1.4%) transfected with 4 µg siRNA$d_{UT3}$. Conversely, transfection of SW620 cells with 4 µg of siRNA$d_{UT3}$ resulted in only a 27 ± 6% reduction in activity. A greater decrease in dUTPase activity was observed in SW620 (95 ± 4%) and HeLa (94 ± 10%) cells when 2 µg of siRNA$d_{UT3}$ was transfected into cells utilizing the Metafectene transfection reagent (Fig. 5.3).

To analyze this inhibition of dUTPase expression at an RNA level, RNA was purified from transfected SW620 and HeLa cells and subjected to Northern hybridization to dUT and β-Actin cDNA probes (Fig. 5.4). Quantitation of the relative intensities showed a decrease in the levels of the $d_{ut}$ mRNA obtained from SW620 cells transfected 3 µg siRNA$d_{UT3}$ and HeLa cells transfected with either 2 or 3 µg siRNA$d_{UT3}$, compared with non-treated cells or cells treated with only the transfection reagent (Fig. 5.5). The percentage decrease was approximately 23% for SW620 cells transfected with 3 µg siRNA$d_{UT3}$ in the presence of the Metafectene transfection reagent, when
compared to cells treated with only the Metafectene reagent. *dut* specific mRNA levels decreased approximately 19% in HeLa cells transfected with 3 µg siRNA*dUT3* in the presence of the Qiagen TransMessenger transfection reagent. There was a greater decrease in HeLa *dut* mRNA levels, 55 and 65%, when Metafectene was used to transfect either 2 µg or 3 µg siRNA*dUT3*, respectively.

### 5.3.2 siRNA*dUT3* Influence on Cell Proliferation

To assess the fate of cells transfected with siRNA*dUT3*, SW620, HT29 and HeLa cells were assayed for cell proliferation using MTS. Cells were transfected with 2 µg siRNA*dUT3* with either Qiagen TransMessenger or Metafectene transfection reagent. Cells were assayed for cell survival 48 hr post siRNA*dUT3* addition. As shown in Fig. 5.6 cell proliferation was decreased in all cell populations compared to non-treated and transfection reagent treated cell populations. The largest decrease in cell proliferation (57.9 ± 2.6%) was observed in HeLa cells. Conversely, cell proliferation was decreased 31.1 ± 2.5 and 32.1 ± 12.21% in HT29 and SW620 cells, respectively. The results shown in Fig. 5.6 were generated from transfection with Metafectene. Similar results were observed when Qiagen TransMessenger Transfection Reagent was used (data not shown).

The function of dUTPase is to maintain low intracellular dUTP pools, thus reducing the likelihood of dUTP being incorporated into DNA. On the rare occasion when dUTP is incorporated there are several cellular enzymes that have the capability to remove it. One of the most well characterized enzymes is uracil DNA glycosylase (UDG). In fact, it has been reported that UDG is responsible for >90% of the cell’s uracil excising capability (239). To investigate the fate of cells when they are transfected with
Figure 5.3: Effect of siRNA<sub>dUT3</sub> on dUTPase Activity. Cells were transfected as described in Materials and Methods and examined for residual dUTPase activity. Results are expressed as the percent of dUTPase activity in treated cells when compared to non-treated controls. (A) Cells transfected with siRNA<sub>dUT3</sub> using the Qiagen TransMessenger transfection reagent. dUTPase activity (units/mg protein) in non-treated controls: SW620, 23.22 ± 1.65; HT29, 5.98 ± 0.78; HeLa, 18.67 ± 3.38. These values were standardized at 100% activity. The values represent the average of three independent experiments. (B) SW620 and HeLa cells transfected with 2 µg siRNA<sub>dUT3</sub> using the Metafectene transfection reagent. dUTPase activity (units/mg protein) in non-treated controls: SW620, 7.93 ± 0.924; HeLa, 12.30 ± 1.324. These values were standardized at 100% activity. The values represent the average of three independent experiments.
Figure 5.3

A

![Bar chart showing dUTPase Activity (% Control) for SW620, HT29, and HeLa cells with different µg siRNA-dUT3 treatments.](image)

B

![Bar chart showing dUTPase Activity (% Control) for SW620 and HeLa cells.](image)
Figure 5.4: Effect of siRNA<sub>dUT</sub> on dUT mRNA Expression. (A) HeLa and (B) SW620 cells were transfected with either 2 or 3 µg siRNA<sub>dUT</sub> for 48 hr using either Qiagen TransMessenger or Metafectene transfection reagent. Northern blots were hybridized with DUT and β-actin. Lanes 1 & 7: non-treated cells; Lane 2: Qiagen transfection reagent (Q); Lane 8: Metafectene transfection reagent (M); Lanes 3 & 4: 2 µg siRNA<sub>dUT</sub> (Q); Lanes 5 & 6: 3 µg siRNA<sub>dUT</sub> (Q) in HeLa blot and 2 µg siRNA<sub>dUT</sub> (M) in SW620 blot; Lanes 9 & 10: 2 µg siRNA<sub>dUT</sub> (M) in HeLa blot and 3 µg siRNA<sub>dUT</sub> (Q) in SW620 blot; Lanes 11 & 12: 3 µg siRNA<sub>dUT</sub> (M).
Figure 5.5: Effect of siRNA_dUT3 on DUT mRNA Expression. (A) HeLa and (B) SW620 cells were transfected with either 2 or 3 µg siRNA_dUT3 for 48 hr using either Qiagen TransMessenger or Metafectene transfection reagent. Northern blots were hybridized with DUT and β-actin. Densitometric scanning analysis of DUT expression. Results were normalized to β-actin expression and expressed as relative intensity.
siRNAdUT3 and treated with a protein inhibitor of UDG (TaT-UGI), SW620, HT29, and Hela cells were assayed for cell proliferation using MTS. Cells were treated with 2 µg siRNAdUT3 as previously described. In addition, 4000 units/ml of TaT-UGI was added daily for a total of three treatments. As shown in Fig 5.6, the decreased cellular proliferation observed with siRNAdUT3 transfection was enhanced when used in combination with TaT-UGI.

5.4 Discussion

While several compounds have been developed that inhibit purified dUTPases (374-379), none of them have been shown to selectively inhibit dUTPase in cultured cells or in vivo. While structural data has been obtained for the human dUTPase that could be used to model compounds capable of inhibiting dUTPase, we do not believe, at this time, it is possible to develop nucleotide analogs that would effectively inhibit the activities of this enzyme in vivo. This is because these nucleoside analogs would have to be delivered to cells as nucleosides and be metabolically activated. The intermediates formed during metabolic activation may function as substrates for several enzymes, including dUTPase. The complexities of such reactions make accurate analysis and interpretation of the data difficult if not impossible. Therefore, alternative approaches are required to determine whether the dUTPase can be used as a potential chemotherapeutic target.

We chose to determine whether dUTPase activity could be decreased in cells by using RNA interference (RNAi). Transfection of SW620, HT29 and HeLa with siRNA that targeted motif 3 of the human dUTPase (siRNAdUT3) resulted in a significant decrease in dUTPase activity in the transfected cells when compared to non-treated
Figure 5.6: siRNAdUT3 Influence on Cell Proliferation. 2 µg siRNAdUT3 and 4000 units/ml TaT-UGI were added to exponentially growing cells (1 x 10^4) as described in Materials and Methods. Cells were harvested 72 hr after treatment and assayed for cell proliferation. The percent control was determined by comparing the cell number in the treated population to non-treated controls. Cell number in the control non-treated population was 39,141 ± 2786, 23,491 ± 2482, and 26,208 ± 2895 for SW620, HT29, and HeLa cells, respectively. These values were standardized at 100% activity. The values were extrapolated from a standard curve established from the absorbances of known cell concentrations. The values represent the average of four independent experiments.
controls (Fig. 3). The amount of inhibition was dependent upon the amount of siRNA-dUT3 used, the cell line used, and the transfection reagent (Fig 3). When the Qiagen transfection was used the greatest decrease (56 ± 3.6%) in dUTPase activity was observed in HeLa cells transfected with 3 µg of siRNA-dUT3 and in HT29 cells (56 ± 1.4%) transfected with 4µg siRNA-dUT3 (Fig. 5.3A). Conversely, transfection of SW620 cells with 4 µg of siRNA-dUT3 resulted in only a 27 ± 6% reduction in activity. When the Metafectene transfection reagent was used, the ability of siRNA-dUT3 to inhibit dUTPase expression was increased, approximately 95% in SW620 and HeLa cells, compared to cells transfected using the Qiagen transfection reagent (Fig. 5.3B). The differences in siRNA-dUT3 inhibition of dUTPase expression probably reflect differences in the expression of dUTPase specific mRNA, differences in the stability of dUTPase and differences in transfection efficiencies of the cells or a combination of the above.

The specificity of siRNA-dUT3 on dUTPase expression in transfected cells was confirmed indirectly and directly. Analysis of total cell protein concentration and UDG activity indicated there was no decrease in transfected cells when compared to non-transfected controls (data not show), thus indicating that the decrease in dUTPase activity was caused neither by a global decrease in protein expression nor an effect on the expression of an unrelated gene. Northern blot analysis confirmed that the decreased dUTPase activity observed was specific for DUT mRNA (Figs. 5.5 and 5.6). These results demonstrate that a siRNA approach can be used to specifically decrease expression of dUTPase in human cells.
A long-range goal of these studies is to determine whether the two key enzymes responsible for controlling the incorporation and/or retention of uracil can be used as potential targets for the development of specific anti-viral or cancer chemotherapeutic agents. The data I presented in Chapter 4 demonstrated that the transduction of cells with TaT-UGI resulted in the inhibition of UDG activity and a decrease in cell proliferation. In this chapter I presented data demonstrating that siRNA\textsubscript{dUT3} specifically down-regulated dUTPase gene expression and this resulted in a decrease in dUTPase activity in the treated cells. The results also demonstrated that decreasing dUTPase activity resulted in a decrease in cell proliferation. Co-treatment of cells with TaT-UGI and siRNA\textsubscript{dUT3} had an additive anti-proliferative effect when compared to cells treated with either agent. This increased anti-proliferative effect was statistically significant in HT29 and SW620 cells. At this time I do not know what mechanism(s) is responsible for this anti-proliferative effect. Since the MTS assay is actually based on metabolism, it is possible that the overall metabolism in treated cells is reduced when compared to non-treated controls. The cells treated with either TaT-UGI and/or siRNA\textsubscript{dUT3} still retain their membrane integrity based upon their ability to exclude trypan blue. Thus it appears the cells have become senescent. It is possible that the increased incorporation and/or retention of uracil into DNA is causing cell cycle arrest. Additional studies are necessary to determine if this is the case. Regardless of the mechanism, I believe my data supports the hypothesis that dUTPase and/or UDG can be used as targets of the development of specific anti-viral or cancer chemotherapeutic agents.
CHAPTER 6

GENERAL DISCUSSION

6.1 Introduction

Organisms are constantly challenged with the potentially deleterious consequences of uracil incorporation/retention in their DNA. Uracil in the DNA can alter the structure of DNA (353-357), interfere with protein-DNA interactions (358-360), interfere with RNA processing (351-353), and is a premutagenic lesion (242,351,352). However, organisms have evolved very efficient mechanisms that prevent uracil incorporation/retention in DNA. Two of the most well characterized mechanisms involve dUTPase and UDG. These enzymes have been demonstrated to be important in maintaining homeostasis in various organisms including viruses. While the role(s) of these enzymes in the viral life-cycle have not been elucidated, data is accumulating which suggests that these enzymes are important for virus replication and pathogenesis. Expression of dUTPase has been demonstrated to be beneficial to non-primate lentiviruses (125-130). Non-primate lentiviruses encode, as a structural component, a dUTPase. Studies using dUTPase deficient non-primate lentiviruses, demonstrated that they exhibited a higher spontaneous mutation rate and lower replication kinetics than the parental virus (129,130). Furthermore, the preclusion of uracil from lentivirus DNA has
been shown to be important in primate lentiviruses. Primate lentiviruses incorporate a cellular UDG (174,193,194,203,321-323) and subsequent studies have demonstrated that uracil in their DNA inhibits steps in their life-cycle (339,340). Although the precise roles of the virus encoded dUTPase and the cellular UDG is unknown, it is apparent that the lentiviruses have evolved mechanisms to exclude the incorporation and/or retention of uracil in their DNA.

The precise role(s) of UDG and dUTPase in DNA replication and DNA repair is unclear. While an UDG deficient mouse has been constructed (239), investigators who have constructed (289) or attempted to construct UDG deficient cell lines (288) have presented data that conflicts with the results and/or interpretation of the results generated from the UDG knockout mouse. As is the case with viruses, there appears to be some controversy regarding the precise role of UDG in mammalian cells.

The inability to isolate and/or construct dUTPase mutants in mammalian cells suggests that it performs a function essential for cell survival. The function of dUTPase is to maintain low intracellular levels of dUTP, so that this nucleotide is not incorporated into DNA. As mentioned before, there are several potentially deleterious outcomes when uracil is introduced into the DNA. One of the most commonly targeted enzymes for drug therapy is thymidylate synthase (TS). TS is the rate-limiting step in \textit{de novo} thymidine triphosphate (dTTP) synthesis. Inhibition of TS prevents the reductive methylation of dUMP to dTMP and, ultimately, a build-up of dUMP and dUTP. Thus, inhibition of TS causes a decrease in the dTTP and an increase in the dUTP pool. This increases the likelihood of dUTP incorporation in DNA during replication and repair. Increased dUTPase expression has been shown to induce resistance to thymidylate synthase
inhibitors in cancer cells (370,372,373). An increased expression of dUTPase would decrease the amount of dUTP available for DNA incorporation by DNA polymerases. Thus, it appears that certain mutations and/or physiological alterations that increase dUTPase expression would give tumor cells a selective advantage in the presence of certain anti-neoplastic agents. The role of UDG in determining the responsiveness of cancer cells to chemotherapeutic agents has not been clearly elucidated.

Several reports have suggested either dUTPase or UDG could be targets for the development of specific chemotherapeutic agents (374,375,387-391). However, few inhibitors have been described that target either dUTPase or UDG, and most of these are targeted against the monomeric dUTPases and UDGs encoded by the members of the Herpesviridae family (374-379). Although the crystal structures have been determined for the human dUTPase (reviewed in 306) and UDG (reviewed in 392), I do not believe, at this time, it is possible to develop nucleotide analogs that would effectively inhibit the activities of the enzymes in vivo. Because these nucleotide analogs would have to be delivered to cells as nucleosides, they would require the cell to metabolically activate them in order to either inhibit dUTPase or be incorporated into DNA to inhibit UDG. The intermediates formed during metabolic activation may function as substrates or irreversibly inhibit several other enzymes. The complexities associated with such reactions make accurate analysis and interpretation of the results difficult. Therefore, alternative approaches are required to determine the roles of dUTPase and UDG in the normal metabolism of an organism, as well as to target these enzymes for anti-viral and cancer chemotherapy.
6.2 Review of Results

The purpose for conducting the preceding studies was to determine the function of UDG in primate lentivirus replication and to determine whether dUTPase and/or UDG could be used as potential targets for the development of specific anti-viral or cancer chemotherapeutic agents. The focus of the studies in chapters 2 centered on the incorporation of UDG into HIV-1, the role of UDG in the virion, and whether UDG could be a viable target of chemotherapeutic agents. The data I presented in chapter 2 demonstrates that Vpr is able to associate with and incorporate into virions the \textit{Escherichia coli} UDG (UDGcoli). The interaction was dependent upon the same residues shown to be critical for the human UDG (193,200). Furthermore, I developed a novel system for the purification of Vpr complexed with the UDGcoli. Using the Vpr-UDGcoli complex, it was possible to demonstrate that the interaction of Vpr with UDG neither altered the biochemical or functional properties of UDG nor did it impart a new function upon UDGcoli (e.g. dUTPase or DNase). Additionally, the binding of Vpr to UDGcoli did not affect the ability of UGI to inhibit UDGcoli activity. The ability of Vpr to interact directly with \textit{E. coli} UDG was confirmed using a yeast two-hybrid assay and it was also demonstrated that UDGcoli was incorporated into virions.

The studies conducted in Chapter 3 attempted to determine a mechanism(s) responsible for uracil formation in HIV DNA during reverse transcription and to explain why UDG, an enzyme responsible for excising uracil from DNA, is incorporated into HIV-1/SIV. The data I presented demonstrated that while several reverse transcriptases (RTs) were capable of incorporating dUTP just as efficiently as dTTP when a polyA template was used, dUTP was only stably incorporated into extracts from SIV virions
when UGI was present, indicating that UDG is functional. My results, also, demonstrated using purified replicative polymerase that dUTP was not incorporated, to any significant degree, across from a guanine. Measurement of the dNTP pool levels in three different cell lines demonstrated that dTTP pools were larger than dUTP and dCTP pools. dCTP pools were larger than dUTP pools. In fact, dUTP pools were not detected using my assay.

The data I presented in Chapter 4 demonstrated that by constructing a recombinant protein containing the protein transduction domain of the HIV-1 encoded TaT protein fused to UGI, a natural protein inhibitor of UDG, I could inhibit the UDG activity in various mammalian cell types. The benefit of making such a protein is that is capable of transducing cells in culture when added directly to the media, thus, eliminating the difficulties associated with DNA transfection and expression. In addition to the UDG inhibition, treatment of cells with TaT-UGI resulted in decreased protein levels, dUTPase activity and cell proliferation when compared to non-treated controls.

In chapter 5 I demonstrated that a siRNA molecule synthesized to recognize motif 3 of the human dUTPase (siRNA\textsubscript{dUT3}) specifically down-regulated dUTPase gene expression at the post-transcriptional level and this resulted in a decrease in dUTPase activity in the treated cells. The results also demonstrated that decreasing dUTPase activity resulted in a decrease in cell proliferation. Co-treatment of cells with TaT-UGI and siRNA\textsubscript{dUT3} had an additive anti-proliferative effect when compared to cells treated with either agent. This increased anti-proliferative effect was statistically significant in HT29 and SW620 cells.
6.3 Overall Conclusions

The purpose of these studies was to investigate the role of UDG and dUTPase in viral replication and cancer cell metabolism, and to determine whether these enzymes were viable targets for the development of novel chemotherapeutic agents. While the initial studies examined UDG and the later studies dUTPase, I believe that both enzymes play a major role in maintaining viral and cellular homeostasis. Although my studies only focused on uracil metabolism in HIV-1/SIV, numerous other viruses encode for UDG and/or dUTPase, thus uracil incorporation/retention maybe a more global issue. There are reports demonstrating the importance of UDG in replication of cytomegalovirus (317) and poxviruses (319).

The data I presented did not define a precise role for UDG in HIV-1/SIV replication, nor did I describe a mechanism that could explain the G-to-A hypermutagenesis associated with lentiviruses. However, my studies demonstrated that RT was not able to incorporate uracil across from guanine. This type of mispair is required to form a G-to-A transition mutation. In addition, my studies conducted on the dNTPs pool levels in three different cell types demonstrated that was no measurable amount of dUTP in any of the three cell types. Although, these studies were not conducted in HIV-1/SIV infected cells nor were they conducted in macrophages or T-cells, natural sites of HIV-1/SIV infection, previous studies measuring dUTP/dTTP and/or dCTP/dTTP ratios in macrophages indicate this phenomenon is very unlikely to account for hypermutation.
Thus, if uracil is not being introduced by RT, and the possibility of spontaneous deamination of dCMP is very low, why does the virus incorporate a cellular enzyme that removes uracil? Recently, studies have demonstrated an association between the cellular protein CEM 15, also known as APOBEC3G, and the HIV-1 encoded Vif protein (393). This protein is a member of a family of cytidine deaminases. This same study demonstrated that in the absence of Vif, expression of CEM 15 rendered the progeny virions non-infectious. Thus, they concluded that CEM 15 may represent a form of innate anti-viral resistance and that Vif expression may prevent this ant-viral pathway. Subsequent studies observed a G-to-A hypermutagenesis in the progeny virions produced from HIV-1 containing a mutation in the vif gene when compared to wild-type virus (394). However, the hypermutagenesis was only detected in the mutant virus when cells expressing CEM 15 were infected.

After the initial study that demonstrated UDG incorporation into HIV-1 (193), a subsequent study reported that the inability to package cellular UDG into HIV-1 resulted in an approximately four-fold increase in the mutation rate, with G-to-A transition mutations accounting for the majority of mutations observed (174). This led to the premise that UDG was acting as an anti-mutator by specifically removing uracil residues during reverse transcription or by some other mechanism(s) such as preventing the access of deoxynucleotide triphosphates to RT or by interacting with RT to influence its enzymatic activity (174). However, I believe this premise is neither supported by existing data nor by the data I presented in this study. Rather, I believe UDG is functioning to remove uracil residues from the viral DNA that potentially could inhibit an essential step in HIV-1 replication (e.g. plus strand DNA synthesis and integration). The
recent studies reporting the association of a cellular cytidine deaminase protein, CEM 15, with Vif, and G-to-A hypermutation found in HIV-1 when Vive is absent provides a source of uracil in the viral DNA. While, I do not believe that uracil introduction into the DNA by RT or spontaneous deamination is the dominant mechanism of uracil introduction, I cannot rule out the possibility that some uracil can arise in the DNA by these mechanisms. Additional studies need to be performed to accurately address the contribution of uracil by these mechanisms. Furthermore, additional studies need to be performed to address the influence of HIV-1/SIV infection on the metabolism of deoxynucleotide triphosphates. I have presented two models (Figs. 6.1 and 6.2) that illustrate what I believe are the functions of the viral encoded dUTPase in non-primate lentiviruses and the cellular UDG in primate lentiviruses.

The data I presented demonstrates that I have a mechanism by which I can decrease dUTPase expression. I believe this a valuable reagent not only to understand the biological role and significance of dUTPase but, also, to determine whether dUTPase is a viable target for cancer chemotherapy. Targeting dUTPase as a means for cancer chemotherapy is not unfounded. Increased dUTPase expression has been shown to induce resistance to thymidylate synthase inhibitors in cancer cells (370,372,373). Thus, I believe that a way to improve the responsiveness to current cancer therapies would be to decrease dUTPase expression in cancer cells. Decreasing the expression of dUTPase, would lead to an increase in the incorporation of uracil into the DNA, resulting in DNA fragmentation and ultimately cell death. Figure 6.3 illustrates the potential outcomes of inhibiting dUTPase and/or UDG activity in a cell.
While the studies performed on HIV-1/SIV focused on UDG and the studies performed in the cancer cell lines focused on dUTPase, I believe that both UDG and dUTPase could potentially be targets of anti-viral and anti-cancer chemotherapy. Since the two enzymes are related in that one, dUTPase, prevents uracil incorporation and the other, UDG, removes uracil if it is incorporated, targeting one or both enzymes could potentially inhibit viral replication and cancer cell growth and/or improve the responsiveness of current chemotherapeutic agents.
Figure 6.1: Hypothetical Model for the Role of dUTPase in Non-Primate Lentivirus Replication. The non-primate lentivirus encoded dUTPase hydrolyzes dUTP as the deoxynucleotides enter into the viral replicating core, thus minimizing the incorporation of uracil residues into DNA.
Figure 6.2: Hypothetical Model for the Role of UDG in Primate Lentivirus Replication. Uracil is formed in viral DNA either by incorporation of dUTP by RT during reverse transcription or by the cytidine deaminase, CEM 15. UDG clenses the DNA of uracil so that subsequent steps in the replication of the virus are not inhibited (e.g. plus-strand DNA synthesis and/or integration).
**Figure 6.3: Model Depicting the Potential Effects of Inhibiting dUTPase and/or UDG.** Decreasing the expression of dUTPase, would lead to an increase in the incorporation of uracil into the DNA. In the presence of UDG uracil residues would be excised. However, the excess amount of uracil being removed would result in DNA fragmentation and ultimately cell death. The introduction of UGI would decrease UDG expression. Thus, in the presence of UGI uracil residues would be maintained in DNA. The excess amount of uracil DNA leads to cell death.
AT  CGCAGGCCG  TACAA  GGC
TAAGCGTCCGCAATGTACCG

AT UCGCAGGCCGU TACAA TGGC
TAAGCGTCCGCAATGTACCG

UDG
UGI

dUTP
siRNA
siRNA
siRNA

dUMP
dTMP
dTTP

Fig. 6.3

173
LIST OF REFERENCES


298. McGeoch, D. J. 1990. Protein sequence comparisons show that the “pseudoproteases” encoded by poxviruses and certain retroviruses belong to the deoxyuridine triphosphatase family. *Nucleic Acids Res.* **18:**4105-4110.


358. Verri, A., Mazzarello, P., Biamonti, G., Spadari, S., and Focher, F. 1990. The specific binding of nuclear protein(s) to the camp responsive element (CRE) sequence (TGACGTCA) is reduced by the misincorporation of U and increased by the deamination of C. *Nucleic Acids Res*. **18**:5775-5780.


