DIGITALCOMMONS
@WAYNESTATE

Wayne State University

Wayne State University Dissertations

1-1-2011

Crystallographic, Molecular Dynamics, And Enzymatic Studies Of Multi-Drug Resistant Hiv-1 Protease And Implications For Structure Based Drug Design (project 1); Crystallographic Studies Of Human Myelin Protein Zero (project 2)

Zhigang Liu
Wayne State University

Follow this and additional works at: http://digitalcommons.wayne.edu/oa_dissertations

Part of the Biochemistry Commons

Recommended Citation
Liu, Zhigang, "Crystallographic, Molecular Dynamics, And Enzymatic Studies Of Multi-Drug Resistant Hiv-1 Protease And Implications For Structure Based Drug Design (project 1); Crystallographic Studies Of Human Myelin Protein Zero (project 2)" (2011). *Wayne State University Dissertations.* Paper 199.

This Open Access Dissertation is brought to you for free and open access by DigitalCommons@WayneState. It has been accepted for inclusion in Wayne State University Dissertations by an authorized administrator of DigitalCommons@WayneState.

www.manaraa.com

CRYSTALLOGRAPHIC, MOLECULAR DYNAMICS, AND ENZYMATIC STUDIES OF MULTI-DRUG RESISTANT HIV-1 PROTEASE AND IMPLICATIONS FOR STRUCTURE BASED DRUG DESIGN (PROJECT 1); CRYSTALLOGRAPHIC STUDIES OF HUMAN MYELIN PROTEIN ZERO (PROJECT 2)

by

ZHIGANG LIU

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2011

MAJOR: BIOCHEMISTRY &
MOLECULAR BIOLOGY

Approved by:

_____
Advisor Date

_____

_____

_____

_____

www.manaraa.com

DEDICATION

This dissertation is dedicated to my mother and my wife.

www.manaraa.com

ACKNOWLEDGMENTS

I want to express my sincere gratitude to my Ph.D. advisor, Dr. Ladislau C. Kovari, for the precious opportunity to work in his wonderful lab. He rendered me tremendous help not only in my academic development but also in my personal career development. The academic guidance and freedom I enjoy in his lab are the best gifts I have ever received.

I am also grateful to Dr. Iulia A. Kovari for her kindly help in my experiments, academic discussion, and manuscript preparation. In addition, I really appreciate her endeavors to help me to get used to American culture.

Also I want to extend my gratitude to my doctoral committee members, Dr. Brian F.P. Edwards, Dr. Robert M. Johnson, Dr. Patrick M. Woster and Dr. Barry P. Rosen (previous member) for their timely advice and encouragement towards a successful dissertation.

In addition, I want to express my appreciation to all the present and previous members in Dr. Kovari's lab, Yong Wang, Tamaria Dewdney, Maia Orabi, Paul Johnson, for your help in my daily lab life. Without your participation, my lab life would not be so pleasant.

Special thankfulness is given to Dr. Patrick M. Woster for his generous offer to let me work in his lab for one year, and Dr. John Kamholz for the wonderful collaboration on human myelin protein zero project.

Finally, I thank my mother and my wife for their unconditional love and support that has allowed me to finish this dissertation.

www.manaraa.com

PREFACE

This dissertation summarizes the work I have done during the past five years in Dr. Kovari's lab. I am currently working on two independent projects.

The first project addresses structure and function studies of multi-drug resistant HIV-1 protease. Protein crystallography, enzymatic assays, *in silico* modeling, and medicinal chemistry are the major tools employed to reveal the mechanism of HIV-1 protease drug resistance, to design, synthesize, and test the new drug candidates.

The second project focuses on structural studies of the extracellular domain of human myelin protein zero, which is responsible for Charcot-Marie-Tooth disease type 1B.

More Introductory information concerning the two projects, experimental procedures, results, and discussion are included in this dissertation. I hope that reading my dissertation is a pleasant scientific experience for the readers.

Thanks,

Zhigang Liu

www.manaraa.com

TABLE OF CONTENTS

www.manaraa.com

www.manaraa.com

www.manaraa.com

www.manaraa.com

www.manaraa.com

www.manaraa.com

www.manaraa.com

LIST OF TABLES

www.manaraa.com

LIST OF FIGURES

www.manaraa.com

www.manaraa.com

www.manaraa.com

CHAPTER 1: INTRODUCTION

1.1 Retroviruses

1.1.1 Characteristics of retroviruses

Retroviruses are a large family of enveloped RNA viruses that share some common taxonomic components such as structure, composition, and replicative properties(1). The diameter of these virions ranges from 80 to 100 nm with the viral glycoproteins embedded in the lipid envelope. The average size of the viron RNA is about 7–12 kb. The viral RNA is positive, linear, single-stranded, nonsegmented.

The hallmarks of retrovirus include the following key points. First, retroviruses replicate in the host through the enzyme reverse transcriptase which converts its RNA genome to linear double-stranded DNA. Second, the enzyme integrase incorporates the DNA into the host genome. Third, the viral genome then replicates itself as part of the host genome. The integrated viral DNA in the host genome is refered to as a provirus.

Based on the genome organization, the retroviruses comprise two categories: simple and complex retrovirus. Three major coding domains are essential for all retroviruses: *gag*, which encodes viral structural proteins including the matrix, the capsid, and the nucleoprotein protein; *pol*, which encodes viral enzyme consisting of the reverse transcriptase and integrase; and *env*, which encodes the surface and transmembrane segments of the

www.manaraa.com

envelope protein. Besides these essential major coding domains, the replication of retroviruses entails another smaller coding domain called *pro*, which encodes the enzyme protease. Retroviruses contain only these essential elements are referred to as simple retroviruses and complex retroviruses carry other regulatory nonvirion proteins arising from multiply spliced messages. Figure 1.1

1.1.2 Classification of Retroviruseses

Based on the relationship in evolution, retroviruses are classified into seven groups as shown in Table 1.1. Except for lentiviruses and the spumaviruses, retroviruses have oncogenic potential. As shown in Table 1.1, the human T-cell leukemia virus–bovine leukemia virus (HTLV-BLV) and the lentiviruses and spumaviruses are complex retroviruses and all other members are simple retroviruses.

1.2 Lentiviruses

1.2.1 Characteristics of lentiviruses

Lentivirus is a genus in retrovirus family and consists of several groups animal viruses. All these groups have some characteristics in common (2-3).

www.manaraa.com

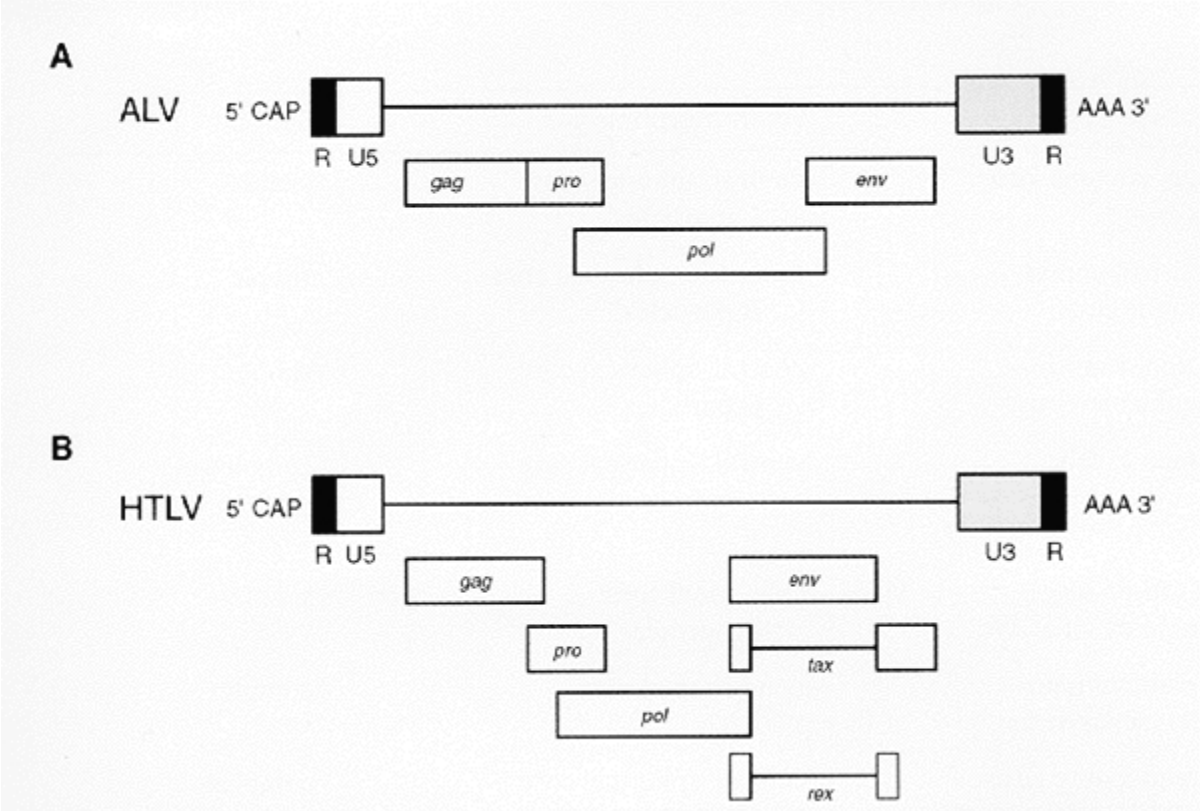


Figure 1.1 Genome of retrovirus (A) A simple retroviral genome. The genetic map of an ALV contains four major coding regions, *gag*, *pro*, *pol*, and *env*. Different reading frames are indicated by vertical displacement of the coding region. The *pro* gene is encoded in the gag reading frame. The terminal noncoding sequences include two direct repeats (R), a U5 (5′unique), and a U3 (3′unique) sequence. **(B) A complex retroviral genome.** The genetic map of HTLV contains, besides the major coding domains, information for two regulatory proteins, Tax and Rex, encoded in regions (boxes) joined by RNA splicing. In this case, *gag*, *pro*, and *pol* are all in different reading frames.

Source: Retroviruses 1997. John M. Coffin, Stephen H. Hugues and Harold E. Varmus. Cold Spring Harbor Laboratory Press.

www.manaraa.com

Table 1.1 Classification of Retroviruses

Genus	Example	Virion morphologya	Genome
Avian sarcoma and leukosis viral groupb	Rous sarcoma virus	central, spherical core “C particles”	simple
Mammalian B-type viral group	mouse mammary tumor virus	eccentric, spherical core “B particles”	simple
Murine leukemia-related viral group	Moloney murine leukemia virus	central, spherical core “C particles”	simple
Human T-cell leukemia bovine leukemia viral	human T-cell leukemia virus	central, spherical core	complex
D-type viral group	Mason-Pfizer monkey virus	cylindrical core “D particles”	simple
Lentiviruses	HIV	cone-shaped core	complex
Spumaviruses	human foamy virus	central, spherical core	complex

[a] Distinctive features seen in transmission electron micrographs.

Source: Retroviruses 1997. John M. Coffin, Stephen H. Hugues and Harold E. Varmus. Cold Spring Harbor Laboratory Press.

www.manaraa.com

Clinically, lentiviruses are associated with a disease with a long incubation period, immune deficiency, arthritis and autoimmunity. In addition, they are involved in diseases affecting the hematopoietic system and central nervous system.

Biologically, they are host species specific, exogenous and nononcogenic. The virus particles show a cone-shaped nucleoid under the electron microscope. The unintegrated circular and linear viral cDNA accumulates in the infected cell. In addition, the infection of the cells could be latent and persistent.

Molecularly, the genome is relatively large, more than 9kb long. The *gag* gene is truncated. The envelope gene is highly glycosylated and highly polymorphic. In addition, there is a novel central open reading frame in the viral genome that separates the *pol* and env regions.

1.2.2 Classification of lentiviruses

The classification of lentiviruses is shown in Table 1.2; the classification is based on the different hosts that are infected by the viruses.

www.manaraa.com

Table 1.2 Lentiviruses

Virus	Host	Primary cell	Clinical disorder
EIA virus[a]	Horse	Macrophages	Autoimmune hemolytic anemia, Encephalopathy
Visna/maedi virus	Sheep	Macrophages	Encephalopathy/pneumonitis
CAE virus[b]	Goat	Macrophages	Arthritis, encephalopathy
BIV[c]	Cow	Macrophages	Lymphocytosis, CNS disease
FIV[d]	Cat	T cell	Immune deficiency, encephalopathy
SIV[e]	Primate	T cell	Immune deficiency, encephalopathy
HIV[f]	Human	T cell	Immune deficiency, encephalopathy and enteropathy

[a]Equine infectious anemia virus, [b]Caprine arthritis-encephalitis virus, [c]Bovine immune deficiency virus ,[d]Feline immune deficiency virus, [e]Simian immune deficiency virus, [f]Human immune deficiency virus

Source: Levy, J. A. HIV and pathogenesis of AIDS -3rd Edition 2008. ASM press, American Society for Microbiology (Editors) & HIV Sequence Compendium 2008

www.manaraa.com

1.3 HIV

1.3.1 HIV epidemiology

HIV was discovered in 1981 (4), and its classification as a member of lentivirus was proposed one year later (5-8). Investigators first attributed AIDS to some known pathogens such at Epstein-Barr virus, cytomegalovirus and human T-cell leukemia virus (9-10). Later Montagnier and coworkers proved that a new human retrovirus, similar to human T-cell leukemia virus, which they called lymphadenopathy-associated virus (LAD), was associated with AIDS by destruction of $CD4^+$ cells(11). In 1984, HTLV-III was identified by Gallo et al to be the pathogen of AIDS (12-15). In addition, Levy and his colleagues reported that they had identified a retrovirus called AIDS-associated retrovirus (ARV) that caused AIDS(16). It was proven later that LAD and HTLV-III were the same virus. In 1986, a new name, HIV, was given to these viruses by the International Committee on Taxonomy of Viruses(17).

Another strain called HIV-2 was discovered in West Africa shortly after the discovery of HIV-1 (18). Both HIV-1 and HIV-2 lead to AIDS despite the more than 55% difference at the level of RNA sequence. The rest of this thesis will discuss HIV-1 only.

According to the AIDS epidemic update December 2009 by the Joint United Nations Programm on HIV/AIDS (UNAIDS) and the World Health

www.manaraa.com

Organization (WHO), 33.3 million people were living with HIV, 2.6 million were newly infected and 1.8 million died of HIV in 2009. The global distribution of HIV infection is summarized in Figure 1.2 (19).

1.3.2 HIV pathogenesis

In the initial stage of an acute HIV infection, HIV most likely attaches to host cells at the site of viral entry, such as activated lymphocytes, dendritic cells, or macrophages in the mucosae or the lymphoid tissues (20-21). The infection is established by the translocation of the HIV from the entry site to local lymph nodes (21-22). Within five to seven days of infection, as many as $>10^7$ viral RNA molecules per ml of plasma may be detected. After 10 to 14 days, up to 200 billion $CD4^+$ cell can be infected and destroyed (23). During acute infection, $CD8^+$ cell count increases. In addition, some innate antiviral cytokines are also produced. However, the cytokines are too late to prevent the establishment of infection (24-25), although they in turn cause acute retroviral syndrome. Due to the host immune response against HIV as well as the loss of target cells (26), the viremia is reduced dramatically within weeks in acute infection.

www.manaraa.com

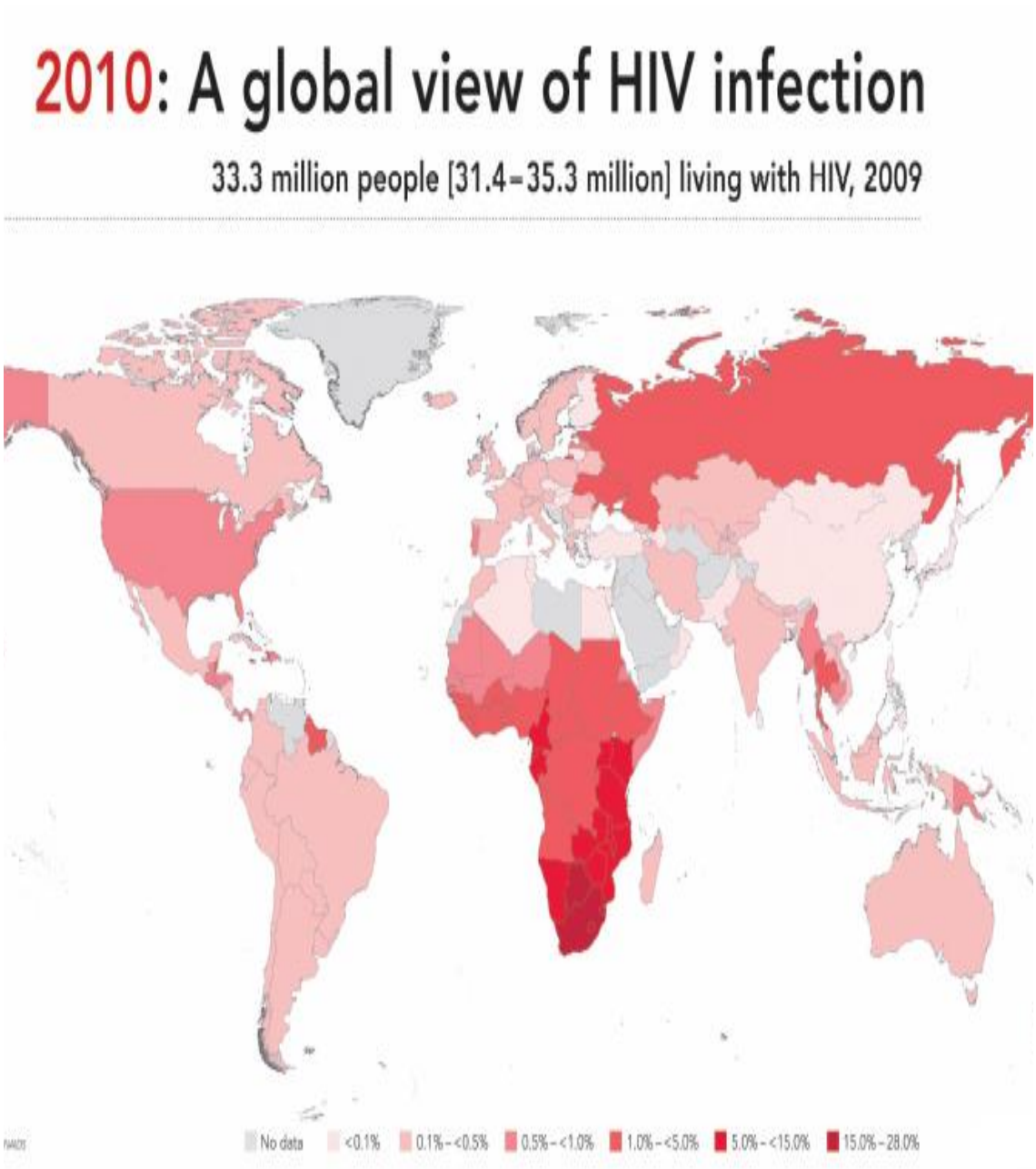


Figure 1.2 Global estimates for HIV infection as of December 2009

The $CD4^+$ count usually returns to values close to normal after three to six month of infection, and then it decreases steadily at a rate of 25 to 60 cells/µl per year due to unknown reasons (27). This asymptomatic period, when the

www.manaraa.com

HIV continues to replicate at a low rate, is called the persistent period and it leads to the heterogeneity of the HIV population under the selection pressure of the immune system (28).The anti-HIV activity of the immune system is thought to be meditated by $CD8^+$ cells (29-30).

The symptomatic period develops in many HIV infection patients within ten years when the $CD4^+$ cell count drops below 350 cell/µl; the viral load increases significantly, and the antiviral $CD8^+$ immune response decreases (31-33). In parallel there is a destruction of lymphoid tissue (34-35). In the end stage of AIDS, $CD4^+$ cells are still the major source of virus except in the GI tract (36). The overall progression of HIV infection is summarized in Figure 1.3 (37).

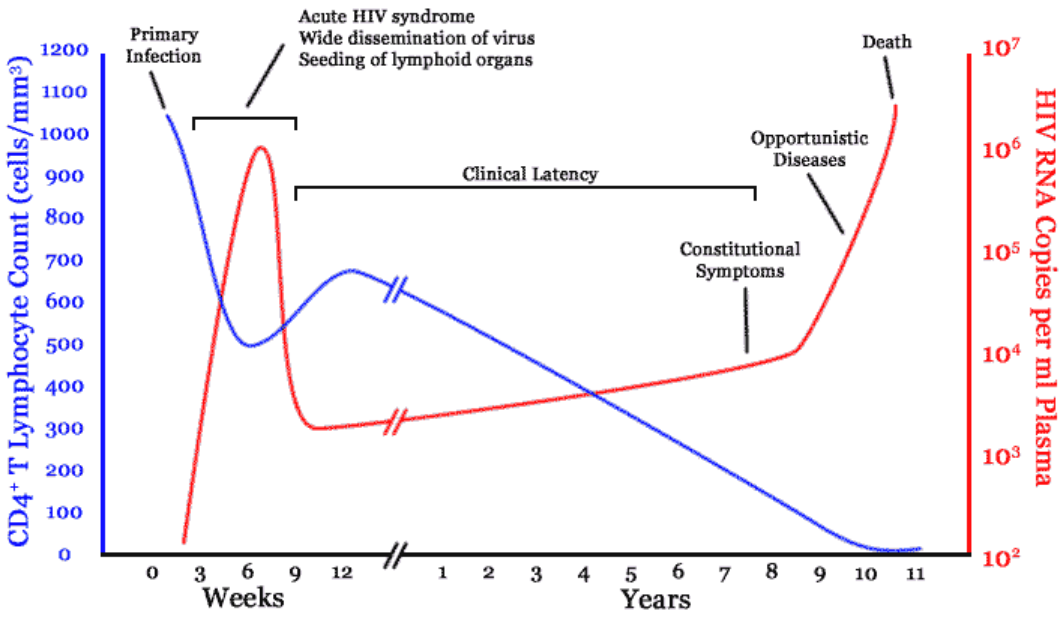


Figure 1.3 progression of HIV infection (37).

www.manaraa.com

1.3.3 HIV structure

Figure 1.4 shows that the HIV viral p24 Gag capsid protein forms a cone-shaped core, which is characteristic of lentiviruses. The viron is about 100 to 120 nm in diameter (38).

The HIV viron is a lipid bilayer enveloped virus, which is derived from the host cell membrane. Two proteins are present on the viral memberane, the external surface envelope protein gp160 and the transmembrane protein gp41, derived from a 160kDa precursor glycoprotein, gp160. They form spike-like structures and later are arranged as tripod-like structure on the viral membrane (39). Underneath the envelope is an inner shell composed of matrix protein (MA, p17), which is required for incorporation of the Env protein in the mature virions (40). Inside the inner shell is a conical core made of capsid protein (CA, p24), within which resides the viral genomic RNA and the viral structural protein nucleocapsid (NC. P7) interacting with viral RNA (41). Inside the capid core, there are three essential viral enzymes, the viral RNA-dependent DNA polymerase (also called reverse transcriptase, RT, p66, p51) (Figure 1.5). In addition, Vif and Nef proteins are closely associated with the core (42-44). Vpr protein is also found within the virion and most likely outside of the core (45-46). Finally, some cytoskeletal proteins are found within the virions (47)

www.manaraa.com

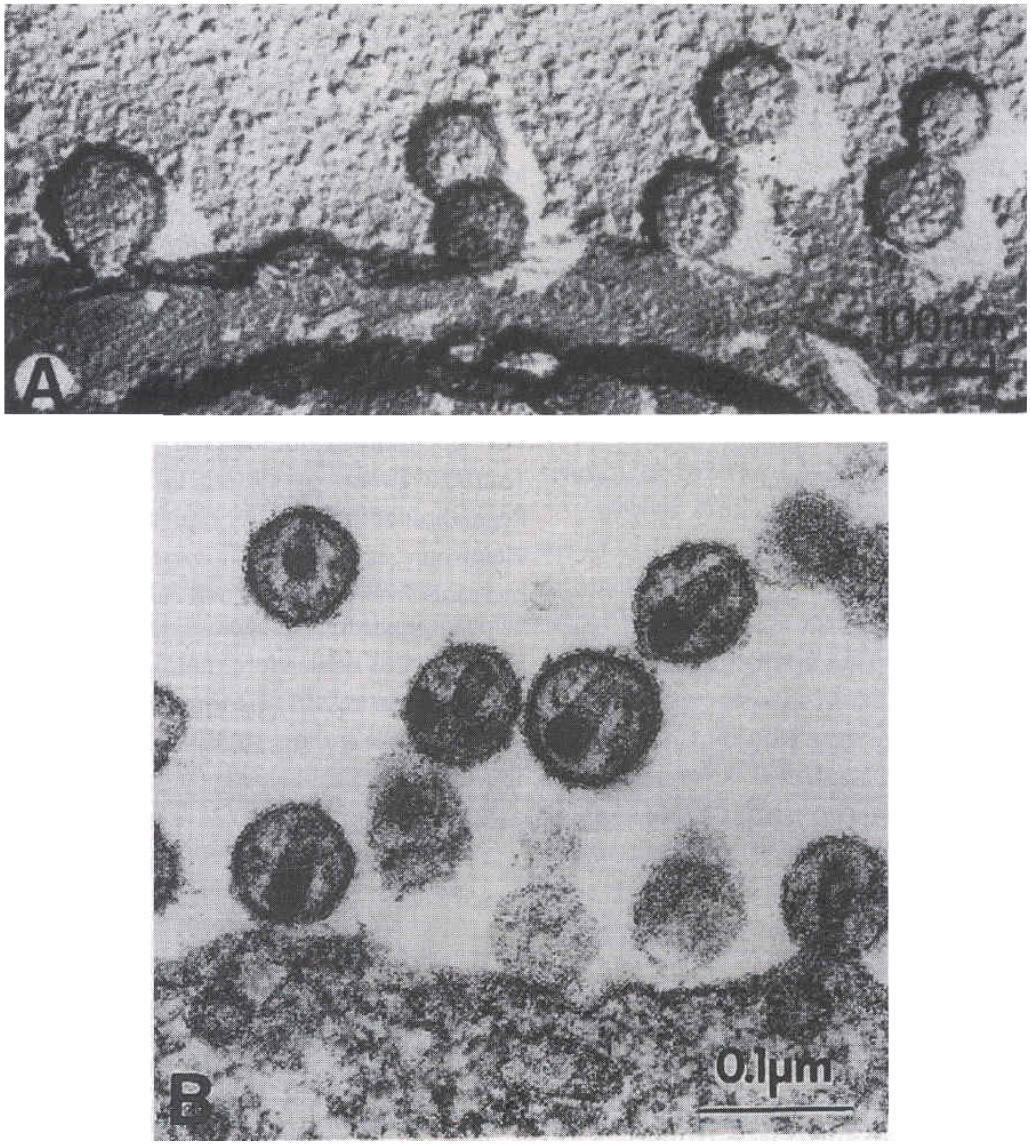


Figure1.4 Morphology of HIV. (A) Scanning electron micrograph of budding particles. (B) Transmission electron micrograph of HIV replication in a T cell.

Source: Levy, J. A. HIV and pathogenesis of AIDS -3rd Edition 2008. ASM press, American Society for Microbiology (Editors) & HIV Sequence Compendium 2008

www.manaraa.com

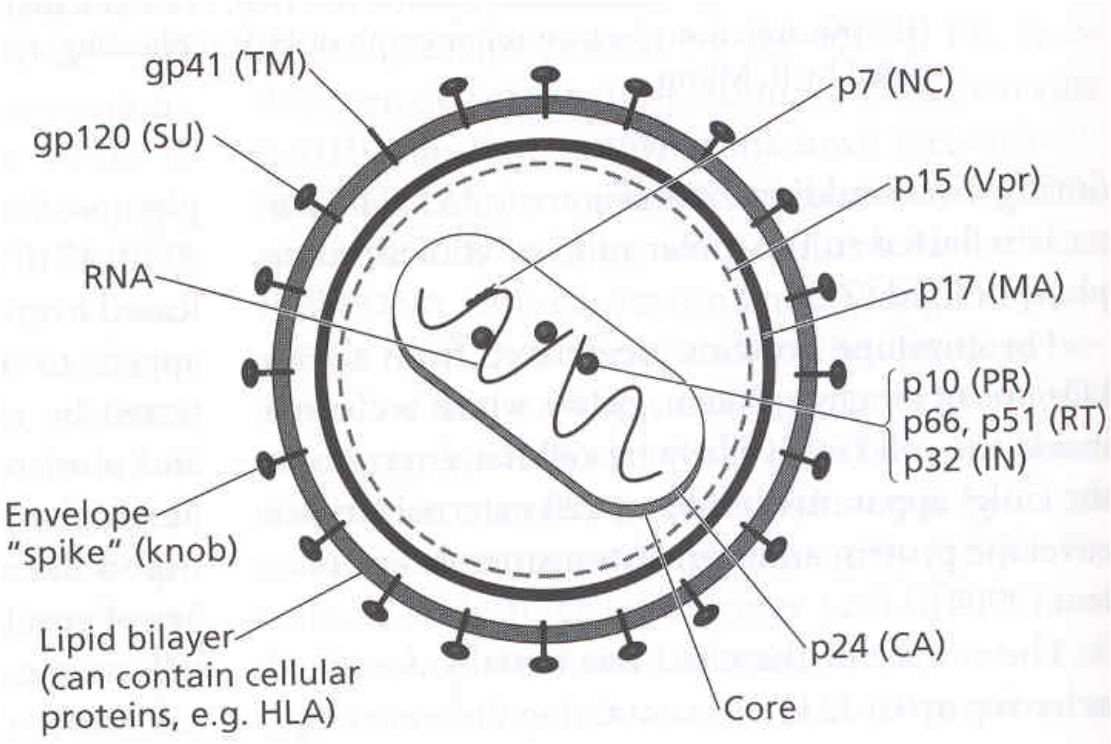


Figure 1.5 an HIV virion with the structural and other virion proteins identified

Source: Levy, J. A. HIV and pathogenesis of AIDS -3rd Edition 2008. ASM press, American Society for Microbiology (Editors) & HIV Sequence Compendium 2008

1.3.4 HIV genome

The HIV genome is about 10 kb long consisting of 46 open reading frames and two long terminal repeats as shown in Figure 1.6. 16 viral proteins are translated based on the open reading frames. The viral mRNA can be unspliced, singly spliced or multiply spliced. The key entity regulating the splicing is the *rev* gene, which is also the product of a multiply spliced mRNA (48).

www.manaraa.com

A full length viral mRNA, translated to Gag and Gag-Pol polyprotein later, is the primary transcription product of HIV. The ratio of Gag and Gag-Pol polyprotein is about 20:1 (49). The Gag-Pol polyprotein is 160kDa and processed by HIV protease into seven individual proteins: matrix protein (p17), capsid protein (p24), late domain (p7), nucleocapsid protein (p9), protease (p10), reverse transcriptase/RNase (p66/p51), and integrase (p32). The first four proteins are Gag protein and the last three are Pol protein.

Similarly, Env precursor (gp160), a translation product of a singly spliced message from the full-length viral mRNA, is proteolytically cleaved by the host endoprotease, furin, into the surface glycoprotein (gp120) and the transmembrane glycoprotein (gp41) (50).

A variety of viral regulatory and accessory proteins affecting the replication are products of the spliced mRNA, including Tat (p14), Tev (p20), Rev (p19), Nef (p27), Vif (p23), Vpr (p15), and Vpu (p16). Tat is a transactivating protein up-regulating HIV replication (51). Rev transports the full-length unspliced mRNA from the nucleus to cytoplasm (48,52). The Nef protein functions in signal transduction and cell activation (53-54). Vif, Vpr, and Vpu/Vpx proteins are important in virion assembly, cell cycling, budding, and infectivity (51,54).

www.manaraa.com

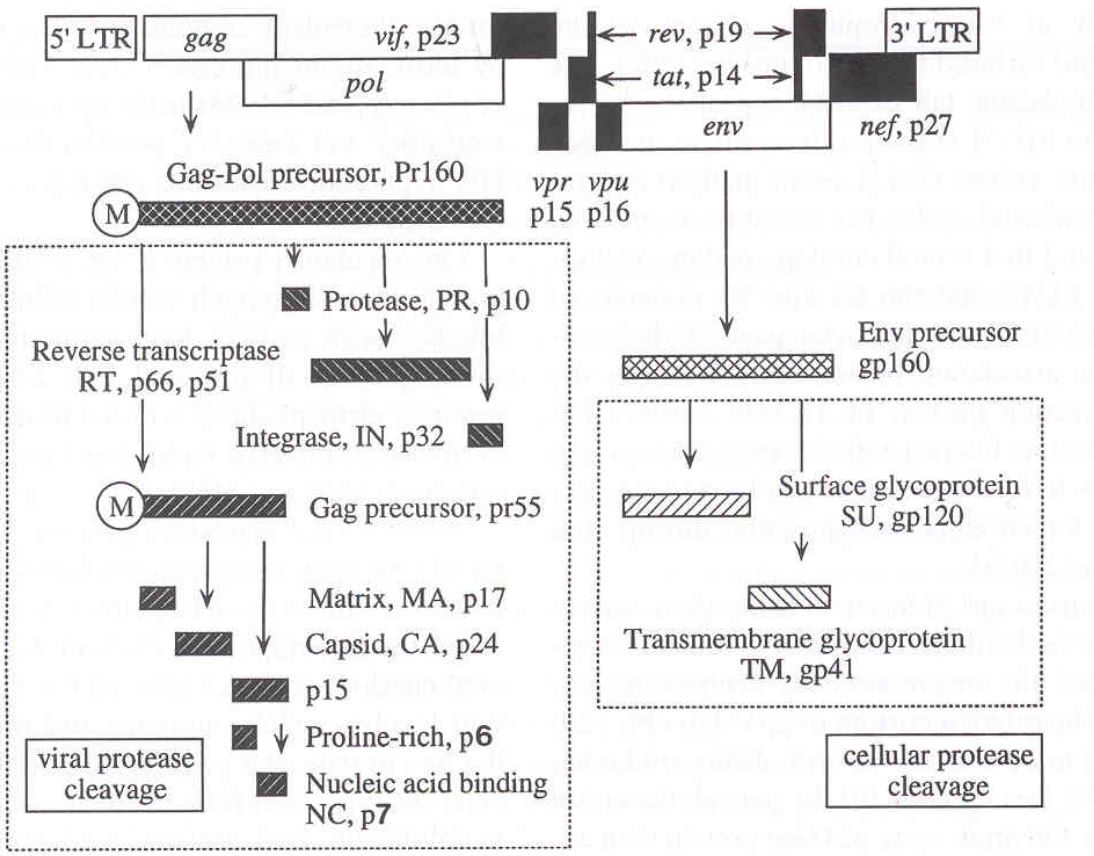


Figure 1.6 The HIV genome and the processing of viral proteins.

1.3.5 HIV replication

The overall replication of HIV is summarized in Figure 1.7

The HIV replication cycle starts with the binding of the HIV envelope protein gp120 to the CD4 molecule on the host T cells (55-57). Further studies revealed that the D1 region of the CD4 molecule participates in the virus binding directly (58-60). After the primary binding event, gp120 is displaced, to make gp41, which is necessary for virus and cell fusion, accessible to the host cell. The displacement of gp120 may be due to the dissociation of gp120 and

www.manaraa.com

gp41, or cleavage of gp120 (61-63). The HIV virion then enters the host cell as a ribonucleocapsid after the virus fuses with the host cell membrane.

In the presence of viral capsid protein and the host cyclophilin A, which is a peptidylprolyl isomerase binding to capsid protein, the HIV RNA exits from the viral capsid (64-66).

Reverse transcription of the viral RNA to double-strained cDNA is then carried out by the enzyme's RNA and DNA dependent DNA polymerase, and RNase H, with assistance from matrix protein and the nucleocapsid protein as a chaperon protein (67-68).

The viral cDNA, together with viral matrix protein, Vpr protein, and integrase as a preintegration complex, is transported to the nucleus as a nuclear pore complex due to a short nuclear localization signal sequence in integrase (69-71). Matrix protein and Vpr are not essential for the preintegration complex (72-73), although the latter may help nuclear importation (71,73).

Inside the nucleaus, linear viral cDNA is integrated into the host cell genome via the catalytic function of integrase and other host cell factors (65,74). The transcriptionally active regions of the genome are the regions where the integration primarily takes please (75-76). Among the host cell

www.manaraa.com

factors, emerin, an integral inner nuclear envelope protein, is necessary for the HIV cDNA integration (77).

The viral mRNA transcription is initiated with some doubly spliced transcripts of the major regulatory genes, such as *tat*, *rev*, and *nef* (51).

Viral assembly takes place in the host cell membrane. The viral RNA is incorporated into the capsid core. The capsid core, in turn, buds from the host cell membrane, where it obtains the lipid bilayer membrane. The gp120 and gp41 are incorporated on the viral membrane via the interaction between the cytoplasmic tail of gp41 and the Gag protein (78-79).

Two major models are proposed for the budding process(41). The lipid raft model, states that Gag and Gag-Pol precursors localize to the lipid rafts on the plasma membrane where sphingolipids and cholesterol are enriched (80-82). The second model, the Trojan exosome hypothesis, states that the assembly of HIV virions happens in multivesicular bodies which later fuse with the plasma membrane and releases the virus as an exosome (83).

www.manaraa.com

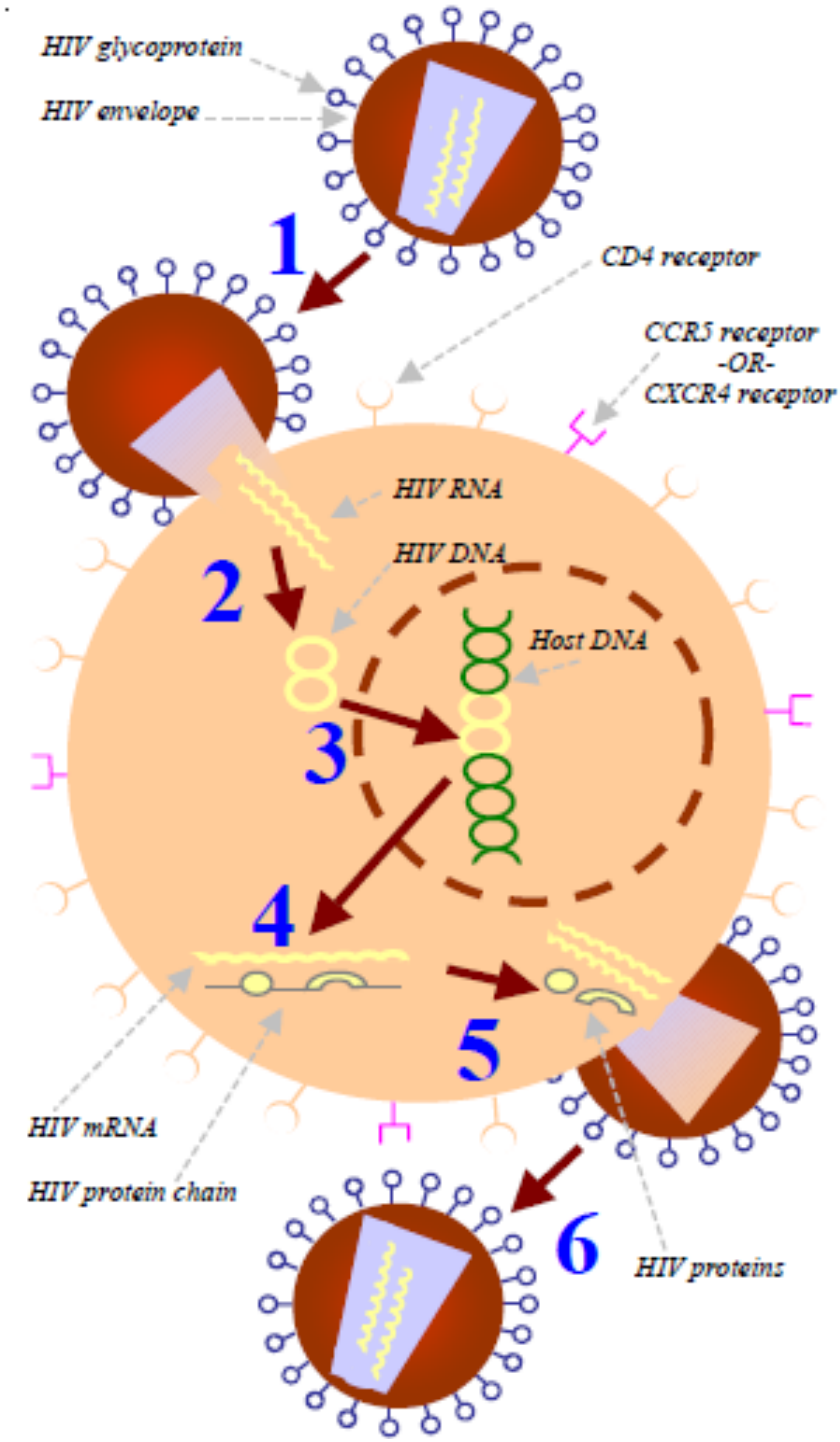


Figure 1.7 HIV replication cycle Adapted from http://aidsinfo.nih.gov/

1.Binding and fusion 2. Reverse transcription 3. Integration

4. Transcription 5. Assembly 6. Budding

www.manaraa.com

1.4 Antiretroviral therapy

1.4.1 Principles of antiretroviral therapy

Treatments for HIV infection have evolved at an extraordinary high rate compared to other infectious diseases. More than 20 different drugs belonging to several categories have been discovered and developed during the past 30 years as shown in Table 1.3. These drugs inhibit essential steps in HIV viral replication, such as virus cell entry (84-85), reverse transcription (86), viral cDNA intergration (87), viral polyprotein cleavage (88-89), and viral maturation (90-93). However, the emerging drug resistance, due to the lack of proofreading ability of the HIV-1 reverse transcriptase (94-96), requires the development of new and potent drugs to overcome the resistance problem (97-101).

Although these drug cannot eradicate HIV completely(102), careful use of of them can reach the following goals: reducing the HIV related morbidity, prolonging survival, improving the quality of life, restoring and preserving immunologic function, and preventing HIV transmission (103). However, a single drug is not potent enough to achieve all these goals due to the rapidly emerging drug resistance under the drug selection pressure (18,99-101).

As a result, a combined antiretroviral therapy, highly actively antiretroviral therapy (HAART) (104-107), is currently the mainstream treatment of HIV infection, which combines at least three drugs. The underlying rationale is: first,

www.manaraa.com

it is unlikely for the HIV virus to be resistant to three drugs in treatment naïve patients; second, the mutation rate, including the drug resistant mutations, is highly dependent on the viral replication rate (108). HAART, with at least three drugs, can suppress the viral RNA below 50 copies/ml for several years (109) and thus reduce the transmission rate (110).

However, if viral replication is not suppressed fully by the initial HAART, the drug combination must be changed to achieve maximal suppression of replication. However, if the full suppression of viral replication is not feasible, even after trials of all the different drug combinations, it is acceptable to have intervals of persisting viremia (111).

Table 1.3 Antiretroviral drugs used in the treatment of HIV infection

Adapted from FDA website.

Brand Name	Generic Name	Manufacturer Name	Approval Date	Time to Approval
Atripla	efavirenz, emtricitabine and tenofovir disoproxil fumarate	Bristol-Myers Squibb and Gilead Sciences	12-July-06	2.5 months
Nucleoside Reverse Transcriptase Inhibitors (NRTIs)				
Brand Name	**Generic Name**	**Manufacturer Name**	**Approval Date**	**Time to Approval**
Combivir	lamivudine and zidovudine	GlaxoSmithKline	27-Sep-97	3.9 months
Emtriva	emtricitabine, FTC	Gilead Sciences	02-Jul-03	10 months
Epivir	lamivudine, 3TC	GlaxoSmithKline	17-Nov-95	4.4 months
Epzicom	abacavir and lamivudine	GlaxoSmithKline	02-Aug-04	10 months

www.manaraa.com

Hivid	zalcitabine, dideoxycytidine, ddC (no longer marketed)	Hoffmann-La Roche	19-Jun-92	7.6 months
Retrovir	zidovudine, azidothymidine, AZT, ZDV	GlaxoSmithKline	19-Mar-87	3.5 months
Trizivir	abacavir, zidovudine, and lamivudine	GlaxoSmithKline	14-Nov-00	10.9 months
Truvada	tenofovir disoproxil fumarate and emtricitabine	Gilead Sciences, Inc.	02-Aug-04	5 months
Videx EC	enteric coated didanosine, ddI EC	Bristol Myers-Squibb	31-Oct-00	9 months
Videx	didanosine, dideoxyinosine, ddI	Bristol Myers-Squibb	9-Oct-91	6 months
Viread	tenofovir disoproxil fumarate, TDF	Gilead	26-Oct-01	5.9 months
Zerit	stavudine, d4T	Bristol Myers-Squibb	24-Jun-94	5.9 months
Ziagen	abacavir sulfate, ABC	GlaxoSmithKline	17-Dec-98	5.8 months

Nonnucleoside Reverse Transcriptase Inhibitors (NNRTIs)

Brand Name	Generic Name	Manufacturer Name	Approval Date	Time to Approval
Intelence	etravirine	Tibotec Therapeutics	18-Jan-08	6 months
Rescriptor	delavirdine, DLV	Pfizer	4-Apr-97	8.7 months
Sustiva	efavirenz, EFV	Bristol Myers-Squibb	17-Sep-98	3.2 months
Viramune	nevirapine, NVP	Boehringer Ingelheim	21-Jun-96	3.9 months

Protease Inhibitors (PIs)

Brand Name	Generic Name	Manufacturer Name	Approval Date	Time to Approval
Agenerase	amprenavir, APV	GlaxoSmithKline	15-Apr-99	6 months
Aptivus	tipranavir, TPV	Boehringer Ingelheim	22-Jun-05	6 months
Crixivan	indinavir, IDV,	Merck	13-Mar-96	1.4 months
Fortovase	saquinavir (no longer marketed)	Hoffmann-La Roche	7-Nov-97	5.9 months

www.manaraa.com

Invirase	saquinavir mesylate, SQV	Hoffmann-La Roche	6-Dec-95	3.2 months
Kaletra	lopinavir and ritonavir, LPV/RTV	Abbott Laboratories	15-Sep-00	3.5 months
Lexiva	Fosamprenavir Calcium, FOS-APV	GlaxoSmithKline	20-Oct-03	10 months
Norvir	ritonavir, RTV	Abbott Laboratories	1-Mar-96	2.3 months
Prezista	darunavir	Tibotec, Inc.	23-Jun-06	6 months
Reyataz	atazanavir sulfate, ATV	Bristol-Myers Squibb	20-Jun-03	6 months
Viracept	nelfinavir mesylate, NFV	Agouron Pharmaceuticals	14-Mar-97	2.6 months

Fusion Inhibitors

Brand Name	Generic Name	Manufacturer Name	Approval Date	Time to Approval
Fuzeon	enfuvirtide, T-20	Hoffmann-La Roche & Trimeris	13-Mar-03	6 months

Entry Inhibitors - CCR5 co-receptor antagonist

Brand Name	Generic Name	Manufacturer Name	Approval Date	Time to Approval
Selzentry	maraviroc	Pfizer	06-August-07	8 months

HIV integrase strand transfer inhibitors

Brand Name	Generic Name	Manufacturer Name	Approval Date	Time to Approval
Isentress	raltegravir	Merck & Co., Inc.	12--Oct-07	6 months

1.4.2 Mechanisms of action of antiretroviral drugs

Reverse transcriptase inhibitors: Nucleoside and nucleotide reverse transcriptase inhibitors (NRTIs) were the first drugs developed to inhibit HIV replication through the termination of viral DNA production and competition of nucleosides used by the viral polymerase. Among these, 3'-azido-3'-deoxythymidine (AZT) was the first drug in this category and

www.manaraa.com

reported to be effective in preventing the onset of HIV infection symptoms (112). It was initially synthesized as a potential drug against cancer (113). Subsequently, seven other nucleoside reverse transcriptase inhibitors have been used in various combinations. Nucleoside inhibitors must be phosphorylated *in vivo* to be active, Consequently, several nucleotide inhibitors of reverse transcriptase have also been developed because they are less dependent on cellular kinase activity for activation.

Four nonnucleoside reverse transcriptase inhibitors (NNRTIs) are on the market: etravirine, delavirdine, efavirenz, and nevirapine. They function by blocking the hydrophobic pocket within the polymerase domain of the p66 reverse transcriptase subunit. Combined with NRTIs, these NNRTIs are potent inhibitors of HIV replication (114).

Protease inhibitors: The FDA has approved nine HIV protease inhibitors (PIs). HIV protease is an aspartic protease that cleaves newly synthesized HIV-1 polyproteins at nine major cleavage sites to generate the mature protein components of an infectious HIV-1 virion. The HIV-1 *gag* gene codes for three structural proteins: matrix protein (MA), capsid protein (CA), and nucleocapsid protein (NC), while the *gag-pol* gene encodes both the structural proteins (MA, CA, and NC) and three enzymes: protease (PR), reverse transcriptase (RT), and integrase (IN). Without effective HIV-1 PR, HIV-1 virions remain uninfectious, hence making HIV-1 protease inhibitors the

www.manaraa.com

most potent anti-AIDS drugs and essential therapeutic components of HAART (100,105). PIs are the most potent drugs currently against HIV replication and are used as one component of HAART. However, the bioavailability of certain PIs and the solubility are major concerns. For example, none of the PIs achieve a high concentration in the cerebrospinal fluid.

Integrase inhibitors: Raltegravir is an integrase inhibitor that disrupts the incorporation of viral cDNA into the host genome. It is proposed that raltegravir interacts with magnesium, a critical metal cofactor of integrase, in the integrase active site to block the catalytic function of integrase (115-117). HIV integrase is a 32 kDa protein of 288 amino acid residues. There are three functional domains: the N-terminal domain, the catalytic core domain and the C-terminal domain (118).

Fusion inhibitors: Enfuvirtide, which was approved by FDA in 2003, is a linear, 36 amino acid synthetic peptide with a sequence identical to part of the HR2 region of gp41 that competes for binding to HR1(119-120). The antiviral activity has been demonstrated in clinical trials, despite its variability in sensitivity against different viral strains (121-125). Enfuvirtide blocks the gp41-mediated membrane fusion of the virus to the host cell. In addition, there are other drug candidates that block the gp120-CD interaction and gp120-coreceptor interactions (126).

www.manaraa.com

1.5 Charcot-Marie-Tooth disease

This section is an introduction to the second project of my thesis, Charcot-Marie-Tooth disease.

Charcot-Marie-Tooth disease (CMT), a hereditary motor and sensory neuropathy (HMSN), is the most common genetic neuropathy with an incidence of 1 in 2500 (127). The disease is named for the three physicians who first identified it in 1886 - Jean-Martin Charcot and Pierre Marie in Paris, France, and Howard Henry Tooth in Cambridge, England. CMT is a heterogeneous group of disorders that affect peripheral nerves. The nervous system has two parts: the central nervous system (CNS) including the brain and the spinal cord; and the peripheral nervous system (PNS), which innervates the muscles and sensory organs. Disorders that affect the peripheral nerves are called peripheral neuropathies. Both motor and sensory functions are compromised in CMT. Atrophy of the foot and lower leg muscles results in foot drop and a high-stepped gait with frequent tripping or falls. Foot deformities are also commonly seen clinically due to weakness of the small muscles in the feet. In the advanced stage of CMT, muscles in the patient's hands also atrophy. Although the onset of symptoms is often in adolescence or early adulthood, the delayed mid-adulthood onset is not rare. Pain may occur, ranging from mild to severe. The mobility of the patients may be also limited due to the muscle atrophy and weakness. Usually CMT does not affect the life

www.manaraa.com

expectancy and thus is a non-fatal disease, although respiration failure may happen rarely if the respiratory muscles are severely affected.

The PNS contains two essential components: neurons and Schwann cells. These two components form extensive cell-cell interactions, which are necessary for the proper functioning of the PNS. Neuronal cells dominate the cell-cell interactions by regulating the survival, proliferation and differentiation of Schwann cells, which are important for the process of myelination (128-129). In turn, Schwann cells function structurally to determine the distribution of ion channels, the maintenance of axons, and neuron development and regeneration (130-132). If the network of the two components is disrupted, it is thought to lead to the development of CMT. The gene mutation may occur on either neurons or Schwann cells, or both of them, which directly or indirectly disrupt the cell-cell interactions (133). In addition, defects in axonal transport are commonly seen in CMT, because neurons in PNS are more sensitive to these defects considering the length of the axons (134). Table 1.4 and Figure 1.8 demonstrate the gene mutations and protein defects on the Schwann cell that lead to CMT (135).

www.manaraa.com

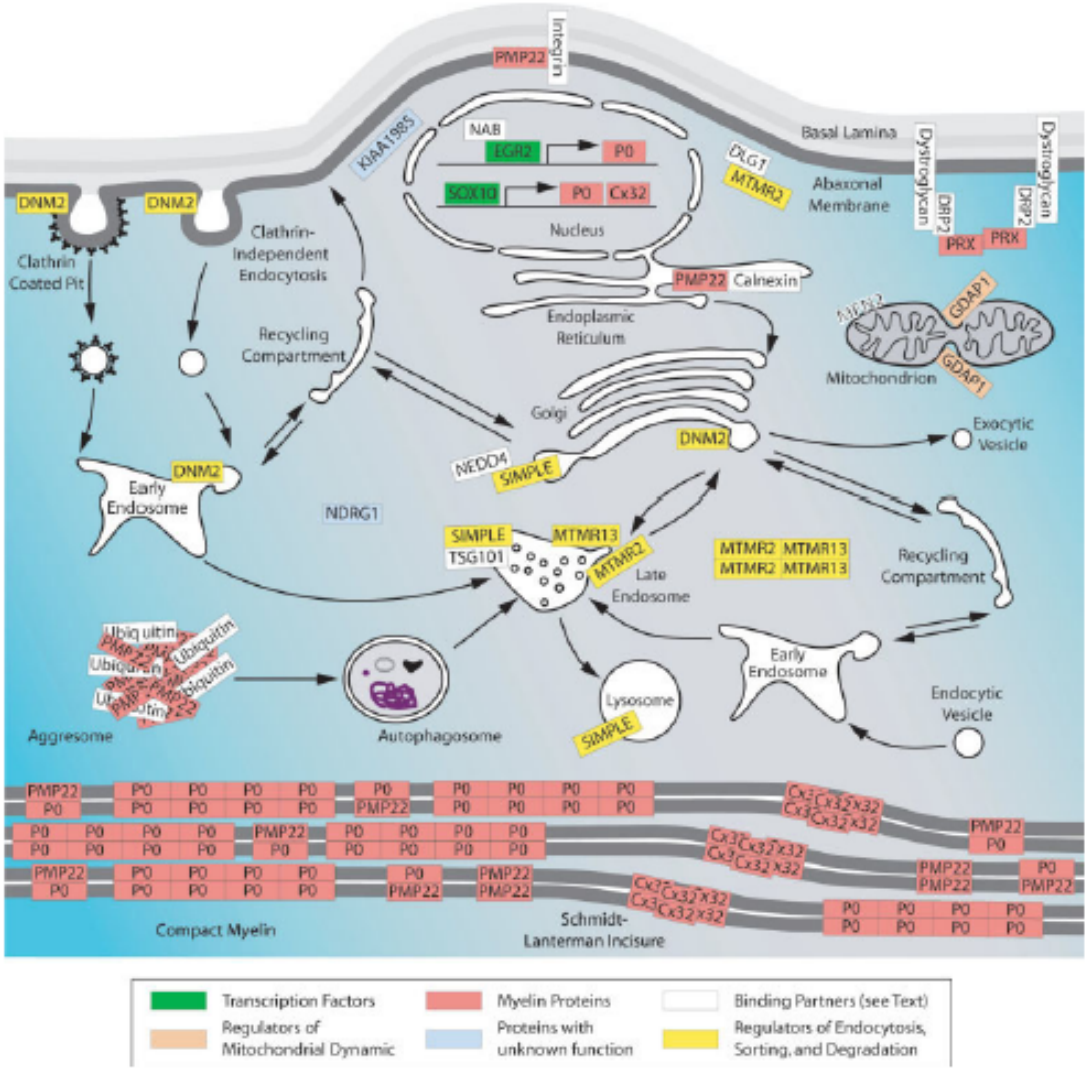


Figure 1.8. Schematic overview of a myelinating Schwann cell highlighting the subcellular localization of proteins mutated in demyelinationg peripheral neuropathies and their binding partners (135)

www.manaraa.com

Table 1.4 Classification of demyelinating neuropathies and their inheritance (135)

Protein	Synonyms	Disease	Inheritance
Myelin structure			
P0	MPZ,	CMT1B	Dominant
PMP22	PASII, SR13, Gas3	CMT1A	Dominant
Cx32	GJB1	X-linked CMT	X-linked
Periaxin	PRX	CMT4F	Recessive
Transcription factors			
EGR2	Krox20	CMT1D/CMT4E	Dominant/ recessive
SOX10		Shah-Waardenburg	Dominant
Vesicular transport			
MTMR2		CMT4B1	Recessive
MTMR13	SBF2	CMT4B2	Recessive
Dynamin2	DYN2	DI-CMTB	Dominant
SIMPLE	LITAF, p53-induced protein 7	CMT1C	Dominant
Mitochondrial morphology			
GDAP1		CMT4A	Recessive
Others			
YARS		DI-CMTC	Dominant
NDRG1		CMT4D/Lom	Recessive
KIAA1985		CMT4C	Recessive
NF-L		CMT1F	Dominant
ARFGEF10		Subclinical	Dominant

There are many forms of CMT disease, including CMT1, CMT2, CMT3, CMT4, and CMTX. The most common subtypes will be discussed here. CMT1 with associated peripheral nervous system (PNS) demyelination is the most frequent diagnosis. It demonstrates slowed nerve conduction velocities and segmental demyelination upon nerve biopsy(136).

It was first found the CMT1A was linked to chromosome 17p11.2 (137-138). Later in 1991, a large segmental duplication about 1.4 Mb of DNA was reported in two independent cases (139-140), within which the complete peripheral myelin protein22 gene (PMP22) was mapped (141). Consequently, defects in PMP 22 are confirmed to cause CMT1A. The duplication of

www.manaraa.com

chromosome 17p11.2 is the most common genetic mechanism of CMT1 as it is seen in about 70% of all cases of CMT1 (142-143). Later it was reported that deleting the same section of chromosome 17p11.2 causes hereditary neuropathy with liability to pressure palsies (HNPP), an autosomal-dominant condition characterized by episodic recurrent pressure palsies, events of paralysis or paresis, at points of nerve entrapment (144). Furthermore, point mutations, including missense, nonsense, splice site, and frameshift mutations, in chromosome 17p11.2 were also found to cause CMT1A (141,145-147). PMP22 is a hydrophobic 22-kDa glycoprotein of 160 amino acids with four transmembrane domains mainly expressed by myelinating Schwann cells in the PNS. It is only present in the compact myelin and constitutes about two to five percent of the total PNS myelin protein (148-149). The function of PMP22 is not clear yet, but it is proposed that PMP22 may function in the initiation of myelin spirals, regulation of growth differentiation of Schwann cells and control of thickness and stability of myelin sheaths (150).

Myelin protein zero (MPZ) is the causative gene for the CMT1B disease, which is much less common than CMT1A, accounting for less than five percent of all CMT1 (142,151-152). The MPZ gene is located on chromosome 1, region q22-q23 (151,153-154). To date, over 125 different mutations in the MPZ gene leading to peripheral neuropathy in patients have been reported worldwide (http://www.molgen.ua.ac.be/CMTMutations/default.cfm). The distribution of the mutations is illustrated in Figure 1.9 (155). It is found that all

www.manaraa.com

mutations resulting in change or deletion of amino acid residues in MPZ give rise to neuropathy with the only exception R215L, which instead causes a benign polymorphism (156). Furthermore, more detailed analysis classifies the MPZ mutations into two major groups (155). In the first group, the mutations disrupt the intracellular processing of MPZ and are mainly associated with early onset neuropathy(157). It is proposed that the mutated MPZ is trapped inside the cell rather than being transported to the plasma membrane. However, other evidence suggests that the mutated MPZ protein is expressed on the plasma membrane, but disrupts the structure of myelin by a dominant-negative mechanism (155,158-159). In the second group, the MPZ mutations are associated with a late onset neuropathy since the mutations cause only mild demyelination. The underlying mechanism is elusive but the current hypothesis is that the mutations cause minor abnormalities in the myelin sheath that over time may lead to aberrant Schwann cell-axon interactions and subsequently to axonal degeneration(157). Myelin protein zero (MPZ), associated with CMT1B, is a transmembrane protein of 219 amino acid residues. Human MPZ consists of three domains: 125 residues constitute the glycosylated immunoglobulin-like extracellular domain, 27 residues span the membrane, and 67 residues represent the highly basic intracellular domain (160-161). MPZ, a homophilic adhesion molecule(162-164), is a member of the immunoglobulin super-family (165) and is essential for normal myelin structure and function. It is proposed that MPZ forms homotetramers, a

www.manaraa.com

doughnut-like structure with a hole, and interacts with the opposite homotetramer with a head-to-head interaction (166). *In vivo* the MPZ knockout mice displayed abnormal myelin that severely affects the myelination pathway(167), and overexpression of MPZ causes congenital hypomyelination of peripheral nerves(168). The cytoplasmic domain of MPZ is also necessary for the proper adhesion function (169).

CMT1C is linked to chromosome 16p13.1 (170) and mutations in the lipopolysaccharide-induced TNF factor gene (LITAF gene) (171-173). Six mutations have been found to cause CMT1C. It is suggested that the LITAF sequence may be involved in the ubiquitin-mediated proteosome pathway and interactions with other proteins via the PPXY and the PSAP motifs (173).

CMT1D contains mutations in the early growth response 2 (EGR2) gene on chromosome 10q21-22 (174). ERG2, a zinc finger transcription factor, is critical in PNS development (175). Consequently, even before the onset of PNS myelination, the EGR2 gene is expressed in mice and rats (176). The regulatory functions of EGR2 include the expression of PMP22, MPZ, gap junction protein β1 and Periaxin, and the synthesis of lipids (177-179).

CMT2E is linked to mutations in the neurofilament protein, light polypeptide gene (NEFL), that cause reduced nerve conduction velocities, axonal degeneration, and the presence of small onion bulbs. NEFL is located within the cytoskeleton of myelinated axons. Neurofilaments (NF), structurally

www.manaraa.com

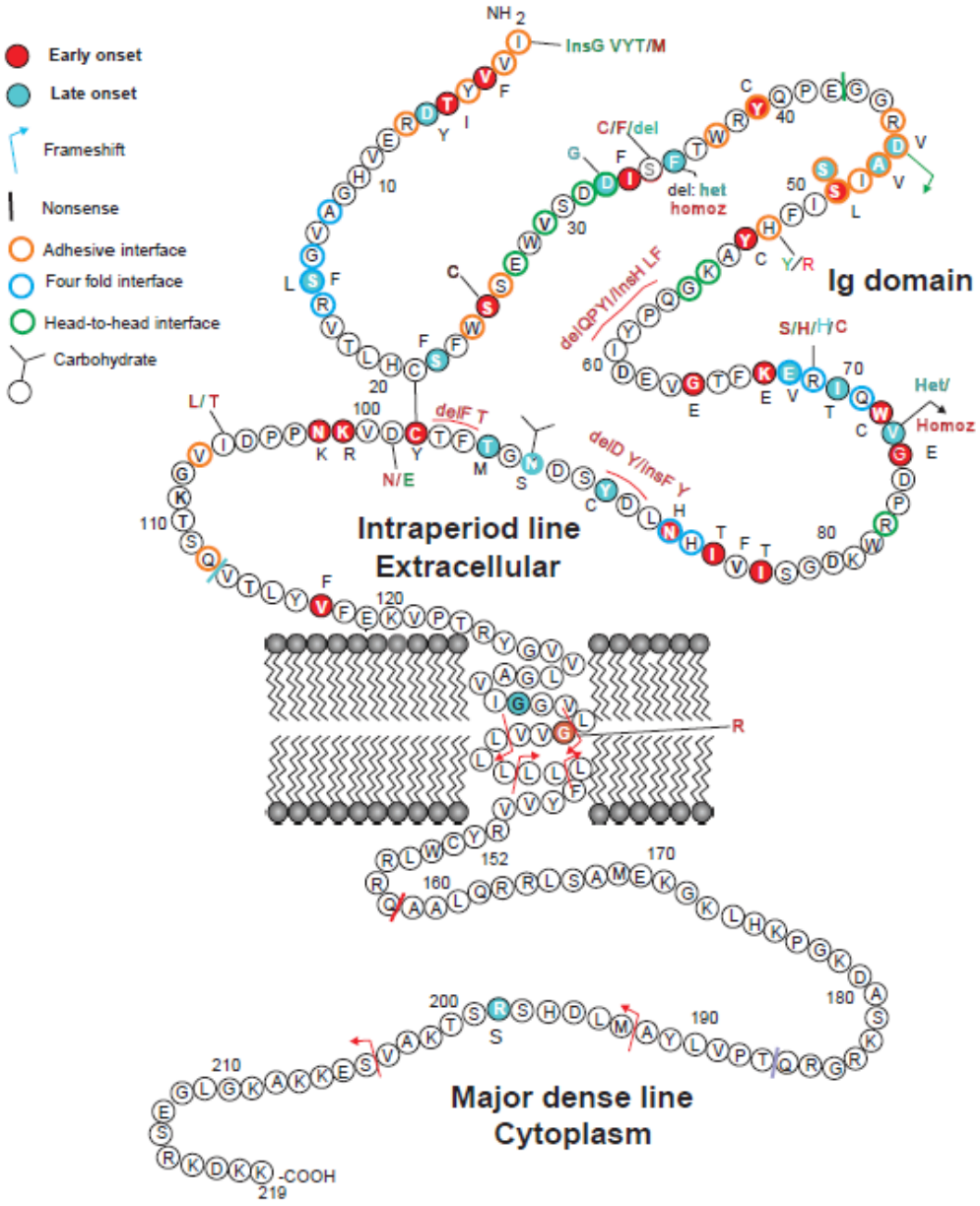


Figure 1.9 Mutations on the open reading frame of MPZ (155).

www.manaraa.com

and functionally important for axons, are composed of three neuron specific proteins: NEFL 68kD, NF medium 125 kD and NF heavy 200 kD on SDS polyacrylamide gels (180-181).

In addition, CMTX, an X-linked form of CMT, arises from a mutation in the connexin-32 gene(182).

www.manaraa.com

CHAPTER 2: MATERIAL AND METHODS

2.1 Plasmid purification and characterization

The plasmids were transformed into DH5α *E*. *c*oli and plated on LB plates in the presence of the antibiotic, ampicilin, as a selective agent. Then a single colony was selected and inoculated into liquid LB medium with the selective agent. The culture was keep at 37 ºC with shaking. The plasmids were purified with a MiniPrep purification kit (Invitrogen). The DNA concentration of the purified plasmids was determined as a function of absorbance at 260nm. The DNA purity was determined by the ratio of A_{260nm}/A_{280nm}. The plasmids were cut by restriction enzymes and the DNA was sequenced. Aliquots of the plasmids were prepared and stored at -80 degrees to avoid freeze and thaw cycles and to prevent potential degradation.

2.2 Agrarose gel electrophoresis

The agrarose gel electrophoresis was carried out with the HOEFER mini DNA electrophoresis apparatus and 0.8 to 1.0 grams of agarose that was dissolved in 100 ml TAE buffer and heated in a microwave for one to two minutes. After the melted agraose had cooled to 50 degrees, ethidium bromide was added for DNA staining. The melted agraose solution was poured into the electrophoresis apparatus. The sample loading wells were made by using a comb with tooth dimensions of 1.5 mm thickness and 5 mm width.

www.manaraa.com

DNA samples were mixed with BlueJuice loading buffer (65% sucrose, 10 mM Tris-HCl, pH 7.58, 10 mM EDTA, 0.3% Bromophenol blue).

The gel was run at 10V/cm for one to two hours until the dye reached about one centimeter from the bottom of the gel. The electrophoresis pattern was viewed on a short wave ultraviolet transilluminator to visualize the DNA stained with ethidium bromide.

2.3 *E. coli.* transformation

The plasmid (50ng) and competent cells (50 µl) were mixed and incubated on ice for ten minutes to achieve maximal contact between the plasmids and competent cells. The *E.c*oli–DNA mixture was heat shocked at 42ºC for 20 seconds. The *E.c*oli–DNA mixture was kept on ice for two minutes to reduce potential damage to the competent cells. LB liquid medium (950 µl) without the antibiotic was added to each tube and incubated at 37 ºC for one hour. Then the liquid culture was spread on two LB plates containing the antibiotic using 200 µl and 800 µl culture respectively. The Petri dishes were grown overnight at 37 ºC to obtain isolated single colonies.

2.4 Protein expression, purification, and dialysis

The conditions for protease expression, purification and dialysis were highly similar between the inactive and active MDR769 HIV-1 protease. The only difference is the pH value of the buffer at the ionic exchange

www.manaraa.com

chromatography step. The buffer pH of the inactive protease preparation is 8.6, and the pH of the active protease preparation is 7.8. Sections 2.4.1 and 2.4.2 describe the preparation of the inactive and active HIV-1 MDR protease.

2.4.1 Inactive MDR769 HIV-1 protease

The inactive MDR769 HIV-1 protease A82T was over expressed by using a T7 promoter expression vector in conjunction with the *E*. *c*oli host, BL21 (DE3). In brief, a fresh transformant of BL21 (DE3) with the MDR769 plasmid was cultured in 5 ml LB medium containing 100 mg/ml ampicillin for 7 hours, which was used to inoculate 50 ml LB medium containing 100 mg/ml ampicillin overnight. Later, the culture was inoculated into two liters of LB medium with 100 mg/ml ampicillin to an to an absorbance at 600nm of 0.1. After 4 to 6 hour of incubation with shaking at 37ºC, the A600 was increased to 0.5 and these cells were harvested at 10,000xg for 10 min at 4ºC by using a Sorvall RC5B Plus centrifuge (183). The inactive MDR769 HIV-1 protease was isolated from inclusion bodies by using a series of buffered washes, followed by denaturing in 6 M urea. For purification of the unfolded protease, an anion exchange resin with pH 8.6 (Q Sepharose, Amersham Biosciences) was used that allowed the protease to pass through and the contaminants to bind. The protease was determined to be purified greater than 95% through Coomassie blue-stained SDS-PAGE (184).The 6M urea was removed to refold the protease by using a series of dialysis exchanges that were carried out at 4ºC. The first buffer

www.manaraa.com

consisted of 0.2M sodium phosphate monobasic, 0.2M sodium phosphate dibasic, 1.0M urea, 0.2% βME, and 10% glycerol. The next three buffer changes consisted of 0.2M sodium phosphate monobasic, 0.2M sodium phosphate dibasic, 0.2% βME, and 10% glycerol. The last two buffer changes were 10 mM sodium acetate, 1 mM DTT and 10% glycerol (pH 5). The protease was concentrated by Amicon to between 2 and 3 mg/ml for storage at -80ºC in aliquots. Figure 2.1 shows the SDS gel of MDR769 HIV-1 protease.

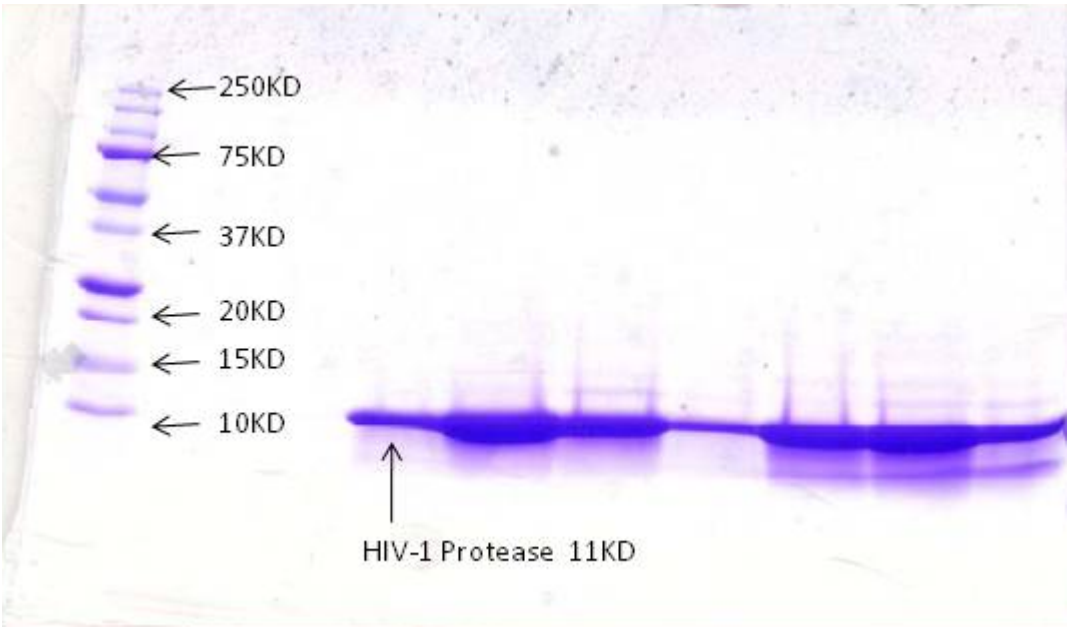


Figure 2.1. SDS PAGE gel of MDR769 HIV-1 protease after ion exchance column at pH 8.6. The left column was the size market. The bands on the right were the denatured inactive MDR769 HIV-1 protease around marker level of 10K Da. The fractions with pure protein were collected, pooled together and subjected to dialysis in order to remove the urea and refolded the protein.

www.manaraa.com

2.4.2 Active MDR796 HIV-1 protease

The active MDR769 HIV-1 protease was over expressed by using a T7 promoter expression vector in conjunction with the *E*. *c*oli host, BL21 (DE3). In brief, a fresh transformant of BL21 (DE3) with the active MDR769 HIV-1 protease plasmid was cultured in 5 ml LB medium containing 100 mg/ml ampicillin for 7 hours, which, in turn, was used to inoculate 50 ml LB medium containing 100 mg/ml ampicillin overnight. Later the culture was inoculated to two liters of LB medium with 100 mg/ml ampicillin to an A600 of 0.1. After 4 to 6 hour of incubation with shaking at 37ºC, the A600 had increased to 0.5 and these cells were harvested at 10,000xg for 10 min at 4ºC by using a Sorvall RC5B Plus centrifuge (183). The active MDR769 HIV-1 protease was isolated from inclusion bodies by using a series of buffered washes, followed by denaturing in 6 M urea. For purification of the unfolded protease, an anion exchange resin with pH 7.8 (Q Sepharose, Amersham Biosciences) was used that allowed the protease to pass through and the contaminants to bind. The protease was determined to be purified greater than 95% through Coomassie blue-stained SDS-PAGE (184).The 6M urea was removed to refold the protease by using a series of dialysis exchanges that were carried out at 4ºC. The first buffer consisted of 0.2M sodium phosphate monobasic, 0.2M sodium phosphate dibasic, 1.0M urea, 0.2% βME, and 10% glycerol. The next three buffer changes consisted of 0.2M sodium phosphate monobasic, 0.2M sodium phosphate dibasic, 0.2% βME, and 10% glycerol. The last two buffer changes

www.manaraa.com

were 10 mM sodium acetate, 1 mM DTT and 10% glycerol (pH 5). The protease was concentrated by Amicon to between 2 and 3 mg/ml for storage at -80ºC in aliquots.

2.4.3 Extracellular domain of human wild type myelin protein zero

The extracellular domain of the wild type human MPZ (hP0ex) was cloned into pMAL™-p2X vector (New England Biolabs), at the *XmnI/XbaI sites*. The insertion of hP0ex downstream from the *malE* gene, which encoded the maltose binding protein (MBP), resulted in a translational fusion of the *E.c*oli MBP to the N-terminus of hP0ex. To facilitate protein crystallization, a flexible triple alanine linker was inserted between the MBP and hP0ex. This polyalanine linker allowed the MBP and hP0ex domains to fold independently. Also, in recent reported crystal structures, the presence of MBP did not disturb the 3D structure of the target protein in crystals (185-187). In addition, the structures of both apo-MBP (188) and maltose-bound forms (189-190) have been determined and could be employed as search models to solve the phase problem.

E. *c*oli BL21 (DE3) cells were transformed with the MBP-hP0ex fusion plasmid, and the cells were grown in LB medium supplemented with ampicillin (100µg/ml) at 37^{o}C to a cell density of OD600 = 0.5. Fusion protein expression was induced with isopropylthiogalactoside (IPTG) for two hours. Cell lysates were prepared, and the fusion protein purified twice by amylose-sepharose

www.manaraa.com

affinity chromatography (191). The final protein was concentrated to 5.0 mg/ml in buffer of 20 mM Tris, pH 8.0, 200 mM NaCl, and 1 mM EDTA. The gel pattern is shown in Figure 2.2

2.5 Crystallization

2.5.1 In-house crystallization screening

The inactive MDR769 HIV-1 protease variants with concentration 1.5 to 2.5 mg/ml were subjected to centrifugation to remove insoluble impurities. The hanging drop vapor diffusion method was used to form the bipyrimidal crystals of the inactive MDR769 HIV-1 protease. Using a matrix screen consisting of pH values (5.5 to 8.1) versus sodium chloride (0.7 to 1.4 M), the HIV-1 protease-substrate complex crystals were formed overnight at 22 ºC. Routinely, 0.2 mm crystals in the longest dimension were obtained after 14 days of incubation. In each well, there were two droplets, containing 1µl of protein, 1µ of reservoir solution and 2µl of protein, 1µ of reservoir solution respectively. The later combination was adopted to increase protein sample concentration to facilitate the crystallization.

www.manaraa.com

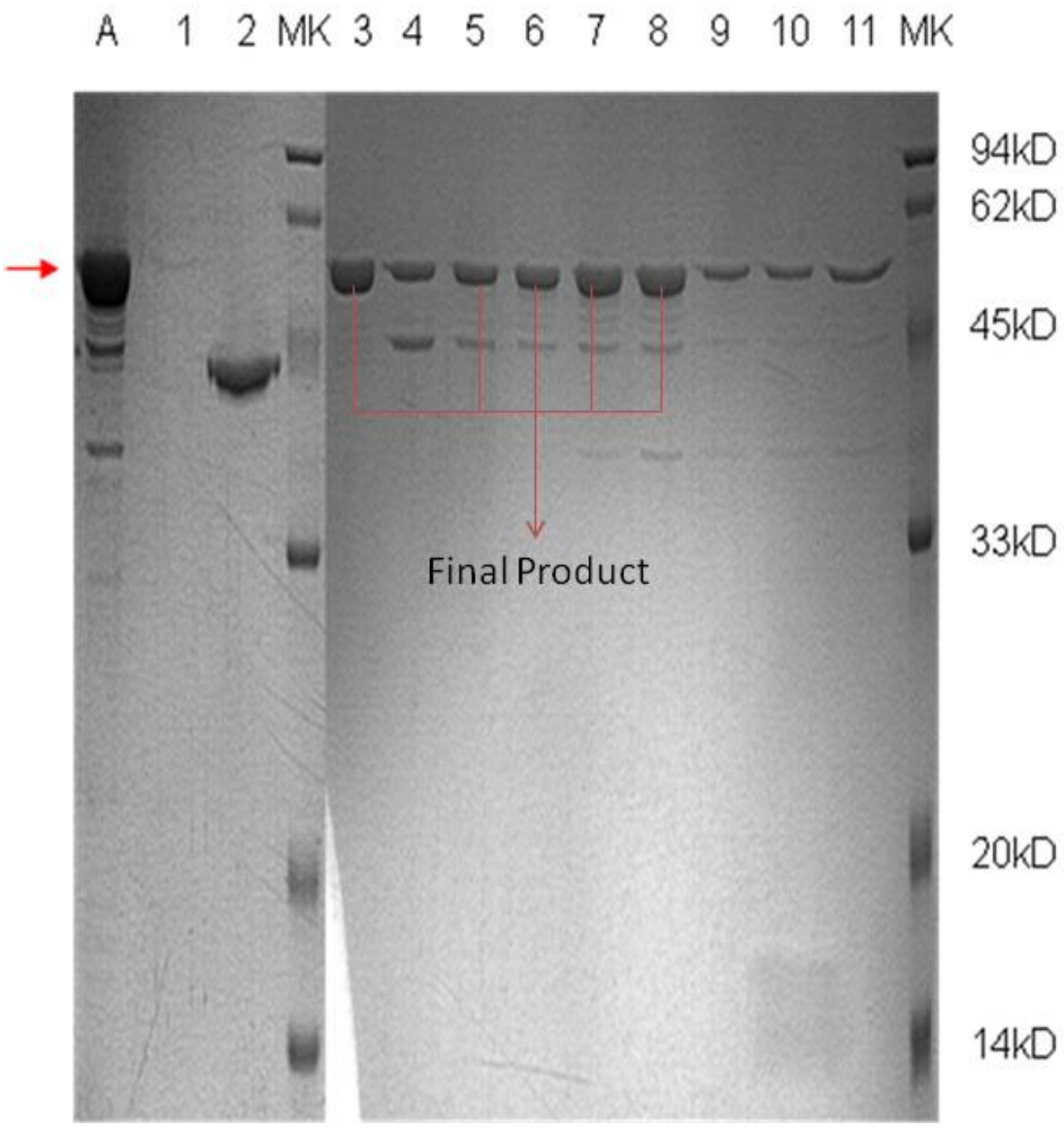


Figure 2.2. The SDS PAGE gel of MBP-hP0ex fusion protein. The red arrow indicates the band of MBP-hP0ex fusion protein. Column A was the initial pool of MBP-hP0ex fusion protein after the first amylose-sepharose affinity chromatography and it was not pure enough. The second amylose-sepharose affinity chromatography gave columns 3, and 5 to 8, which were pooled together to get pure MBP-hP0ex fusion protein.

www.manaraa.com

2.5.2 High-throughput crystallization screen (HTS)

Protein samples were sent to the Hauptman-Woodward Institute (HWI) for high throughput crystallization screening experiments using the batch method (1536-well micro assay plates)(192). We optimized the initial crystallization hits with the hanging-drop vapor diffusion method in the lab. The best crystals (hP0ex) were grown from 40% PEG 8000, 0.1 M KSCN, 0.1 M Tris, pH 11.1. Needle shaped crystals of MBP-hP0ex fusion protein appeared within one week at 22 ºC. The crystals were soaked for two minutes in 30% glycerol and flash frozen for diffraction experiments.

Figure 2.3 shows the original crystallization hit from HTS and Figure 2.4 represents the crystals after condition optimization.

2.5.3 Cryoprotection of crystal

To protect the crystal from cracking when fresh frozen, two different cryoprotectents were used. 30% glucose (W/V) and 30% glycerol (v/v) were used respectively for HIV-1 protease and myelin protein zero crystals. Fresh cryoprotectent was made before flash freezing the crystals. One drop of the cryoprotectent was placed on a clean slide into which a harvested crystal was immensed for about 30 seconds. The crystals were harvested with CyrstalCap HT(Hampton Research), a complete crystal mounting system for manual and

www.manaraa.com

automated cryocrystallography, and were stored with aluminum CryoCanes (Hampton Research) in liquid nitrogen.

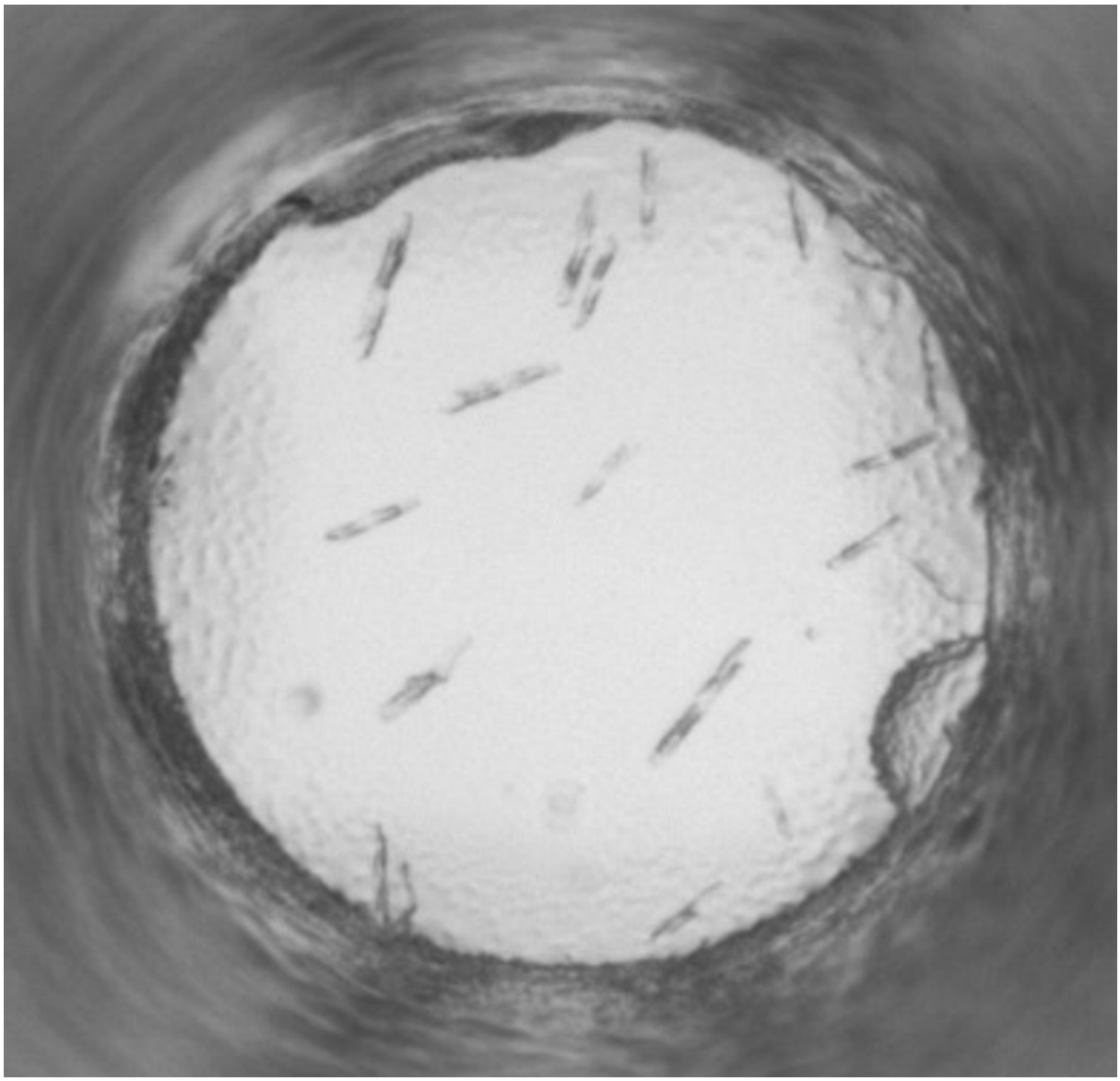


Figure 2.3 Hit from high-throughput crystallization screen showing crystals of the the extra-cellular domain of wild type human myelin protein zero. The crystals of the extra-cellular domain of wild type human myelin protein zero in the leading crystallization condition hit from the high-throughput crystallization screen at Hauptman-Woodward Institute (HWI) using the batch method. The leading hit condition was 20% PEG 8000, 0.2 M KSCN, 0.1 M Tris, pH 9.

www.manaraa.com

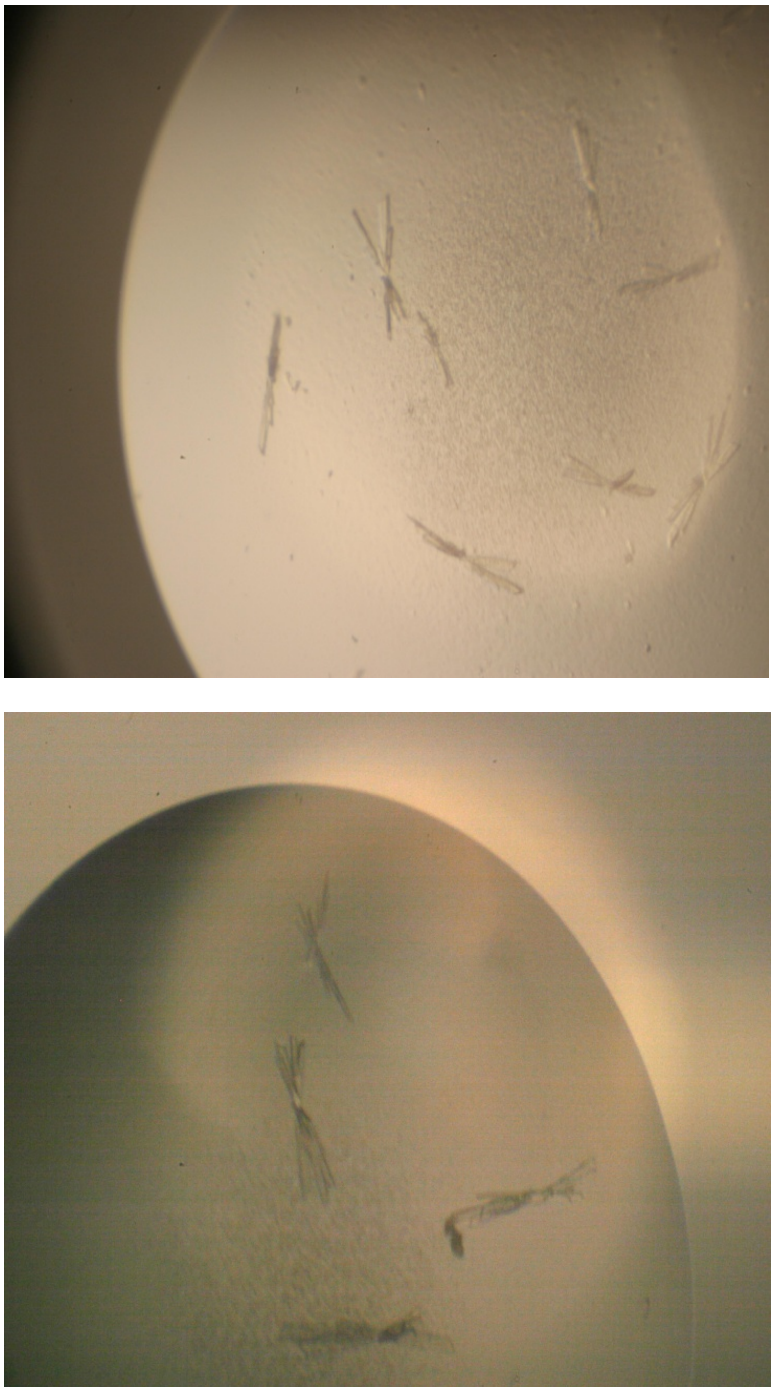


Figure 2.4 The crystals after in-house condition optimization with the hanging drop method based the leading HTC crystallization condition hits. Individual needle shape crystals were harvested and frozen for data collection. The crystallization temperature was at 22 ^{o}C and the condition was 40% PEG 8000, 0.1 M KSCN, 0.1 M Tris, pH 11.1.

www.manaraa.com

2.6 Crystal diffraction data collection

Data were collected from a single crystal at the Advanced Photon Source (APS) (LS-CAT 21), Argonne National Laboratory (Argonne, IL). The beamline used was 21-ID-D with MD-2 diffractometer hardware and software from EMBL for sample mounting, centering, beamline diagnostics, and data collection. The detector was a Mar 300 CCD. Raw data were reduced with program HKL2000.

2.7 Diffraction data analysis

2.7.1 CrystalClear

CrystalClear is a program provided by Rigaku Americas Corporation. It incorporats the concept that data acquisition and data processing should be combined under one common, modern interface called d*TREK, a powerful compute engine designed to process data from all Rigaku detectors including those at synchrotrons. As a result, it provides a complete graphical user interface for data collection and processing, which includes instrument control, data collection, data collection strategy, data processing, image viewing, manipulation tools, and, continuous instrument and processing analysis modules. In the data collection module, there were functions such as integration, empirical and numerical absorption corrections, scaling and

www.manaraa.com

calculation of ESDs, Laue group and space group determination and reciprocal lattice viewer.

2.7.2 HKL2000

HKL2000 was also used to reduce the diffraction data. The HKL-2000 program package was based on the extended versions of Denzo, Xdisplayf, and Scalepack. Similar to CrystalClear, it also employs a graphical user Interface. HKL2000 is a new, extended and significantly improved version of the HKL-package that includes all the features of HKL plus major new options, such as 3-D processing, that improve the integration of data from crystals with very high mosaicity. It also processes data collected with any spindle axis orientation and refines the direction of the spindle axis. The absorption correction dramatically improved the anomalous signal. ????

2.7.3 CCP4

The Collaborative Computational Project Number 4 in Protein Crystallography was set up in 1979 to develop a comprehensive collection of software for crystallography.

For the HIV protease drug resistance project, molecular replacement was performed with Molrep-autoMR in CCP4 (193) with the model previously solved

www.manaraa.com

in our lab (194). Initial refinements were performed by Refmac5 (195-196). The models were further built according to the electron density maps in program COOT (197) as the refinement processed. Crystallographic solvent water was added in program ARP/wARP(198). The structures were refined to the final resolution in Refmac5. The final stereochemical parameters were checked using PROCHECK(199). B factor analysis was done by program Temperature Factor Analysis in CCP4. Structure superpositions and their corresponding RMSD values were calculated with the CCP4 program Superpose Molecules (200). Protease superpositions were based upon the Cα atoms of residues 1 through 99..

In the MPZ structural project, CCP4 suite(193) was used to convert .sca file to .mtz file. Molecular replacement was performed by Molrep-autoMR to get phase information(195) in CCP4 suite of programs with the maltose binding protein model 1ANF and rP0ex model 1NEU(166,190). The maltose molecule was built into the difference electron density maps in the COOT program(197). Solvent water was added using ARP/wARP program(198). The structures were refined to resolution 2.1 Å in Refmac5. The final stereochemical parameters were checked using PROCHECK(199),

www.manaraa.com

2.7.4 COOT

The Crystallographic Object-Oriented Toolkit (COOT) is a convenient tool for macromolecular model building, model completion and validation. It also displays maps and models and allows model manipulations such as idealization, real space refinement, manual rotation/translation, rigid-body fitting, ligand search, solvation, mutations, rotamers, Ramachandran plots, skeletonization, non-crystallographic symmetry and more.

2.7.5 PyMol

PyMol is a molecular visualization system created by Warren Lyford DeLano and commercialized by DeLano Scientific LLC. It is a powerful tool for producing high quality three=dimensional images of both small molecules and biological macromolecules.

2.7.6 PISA Server

PISA is an acronym for protein interfaces, surfaces, and assemblies. It is an interactive tool that analyzes macromolecular interfaces, predicts probable quaternary structures, and searches the PDB and other databases for structurally similar interfaces. Pisa Server was used to calculate the dimer interface of MDR769 HIV-1 protease variants (201).

www.manaraa.com

2.8 Protease enzyme assays

The active MDR769 HIV-1 protease produced in our lab and the WT HIV-1 protease (NL4-3), purchased from AnaSpec, were cryo-protected by 50% glycerol and stored at -20 ºC. The fluorescent substrate (FRET substrate I) which mimicked the natural substrate peptide MA/CA with excitation/emission wavelength 340/490 nm respectively was dissolved in DMSO at a concentration of 500 µM (202-203). The composition of the fresh enzyme buffer was 0.1 M sodium acetate, 1.0 M sodium chloride, 1.0 mM ethylenediaminetetraacetic acid (EDTA), 1.0 mM dithiothreitol (DTT), 10% dimethylsulfoxide (DMSO), 1 mg/ml bovine serum albumin (BSA), and the pH was adjusted to 4.7. The ligand was dissolved in DMSO at a concentration of 20 mM and stored at -20 ºC. It was diluted on a two-fold series (the first concentration is two fold higher than the next one) before enzyme assay.

A 96 well flat bottom assay plate was used for the assay. A SpectraMax M5 from Molecular Devices was used to read the fluorescent signal and SoftMax Pro was used to plot and analyze the data.

In each well of the assay plate, 97 µl of enzyme buffer was added first. Then, 1 µl of one of the proteases was added to the buffer to make the final concentration 20 µM. 1 µl of ligand at different concentrations was then added to the system. The whole system was mixed gently and incubated at 37 ºC for

www.manaraa.com

ten minutes. 1 µl of fluorescent substrate was added to make the final concentration of 5 µM. The fluorescence signal was read at one minute intervals for twenty minutes at 37 ºC. The IC50 value were calculated with the program SoftMax Pro.

The fluorogenic substrate used in the enzyme assay was purchased from AnaSpec. It is a MA/CA substrate peptide analog with the sequence DABCYL-GABA-Ser-Gln-Asn-Tyr-Pro-Ile-Val-Gln-EDANS; this peptide is also referred to as HIV protease substrate I. It was widely used for the continuous assay for HIV protease activity. Upon the cleavage of the substrate, the EDANS molecule fluroresces at 490nm wavelength after excitation at 340 nm. This emission is quenched by nearby DABCYL in the uncleaved substrate. As a result, the experiment demonstrated a time-dependent increase in fluorescence intensity that was linearly related to the extent of substrate hydrolysis.

2.9 Chemical synthesis

The Abbott protocol was adopted to synthesize the library of lopinavir analogs as shown in Figure 2.5 and Figure 2.6 (204). A hydroxylethylene core consisting of the P1 and P1' residue was synthesized first. Next, the P2 and P2' residues were added to the core to obtain the final desired molecule by a regular peptide bond coupling. The amine and carboxylic groups of L-Phenylalanine, the starting material of the hydroxylethylene core, were

www.manaraa.com

protected with benzyl groups in step1 to obtain a benzyl ester. In step 2, the carboxyl group was deprotected to synthesize the nitrile from the ester. In step 3 the Grignard reaction was carried out where the P1 residue was attached to the core. At this point the core was a keto-enone and needed to be further reduced to obtain the final hydroxyethylene core. In step 4, a reduction reaction was used to obtain the hydroxyl group between the P1 and P1' residues. Enantiomers were produced in the reaction and then separated on a silica column to get the desired pure isomer (done by our collaborators). In step 5, the free amine group was protected and the remaining benzyl groups were removed. In step 6, the P2' residue (acid1) was coupled to the free amine group of P1' residue. In step 7, the Boc protection was removed and the P2 residue (acid3) was coupled to the free amine group of the P1 residue. In my experiment, P1 site is the side chain of Phenylalanine and P1's site is the side chain of Homo- Phenylalanine. The supporting documents (MS analysis results of the reaction intermediate) for each reaction are shown in the appendix.

www.manaraa.com

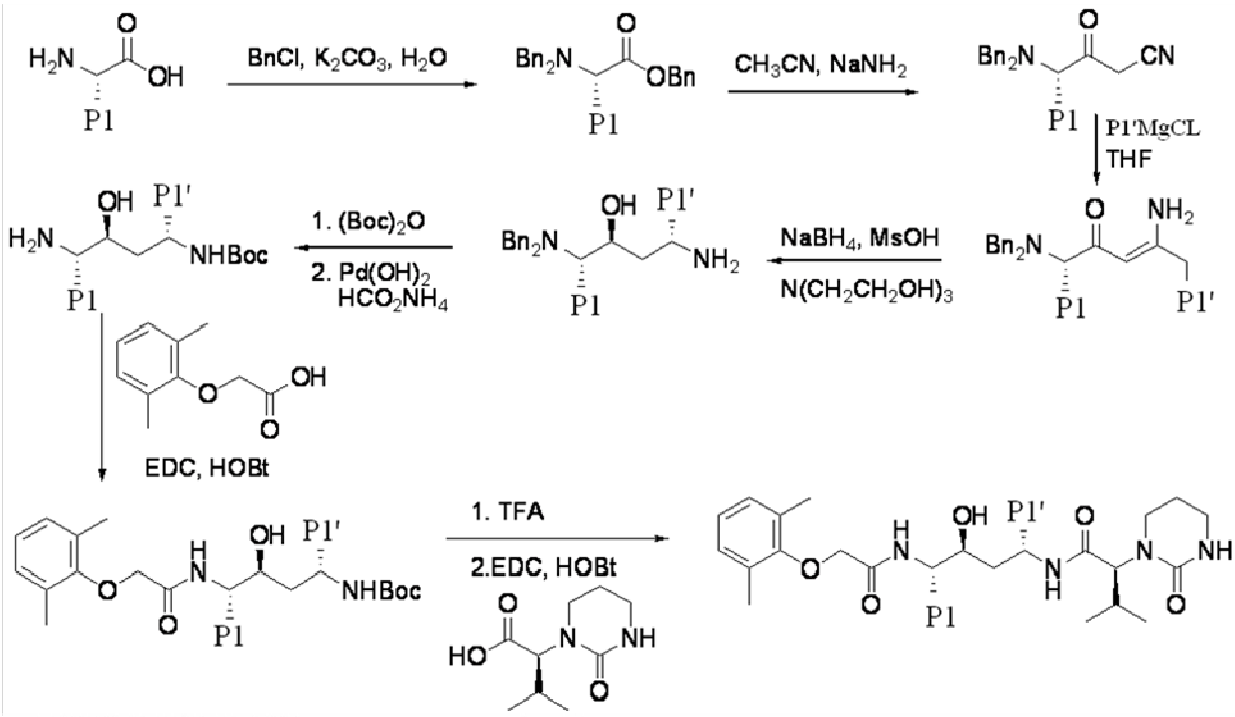


Figure 2.5 The synthetic scheme of lopinavir analog. This synthetic scheme was adapted from the Abbott protocol. The hydroxylethylene core was synthesized in the lab with the desired modifications on the P1 and P1' sites of lopinavir. Four stereoisomers were produced after the step 4, which require further purification. The core then was coupled with prepared acid 1 and acid 3 to yield the final product of a lopinavir analog.

www.manaraa.com

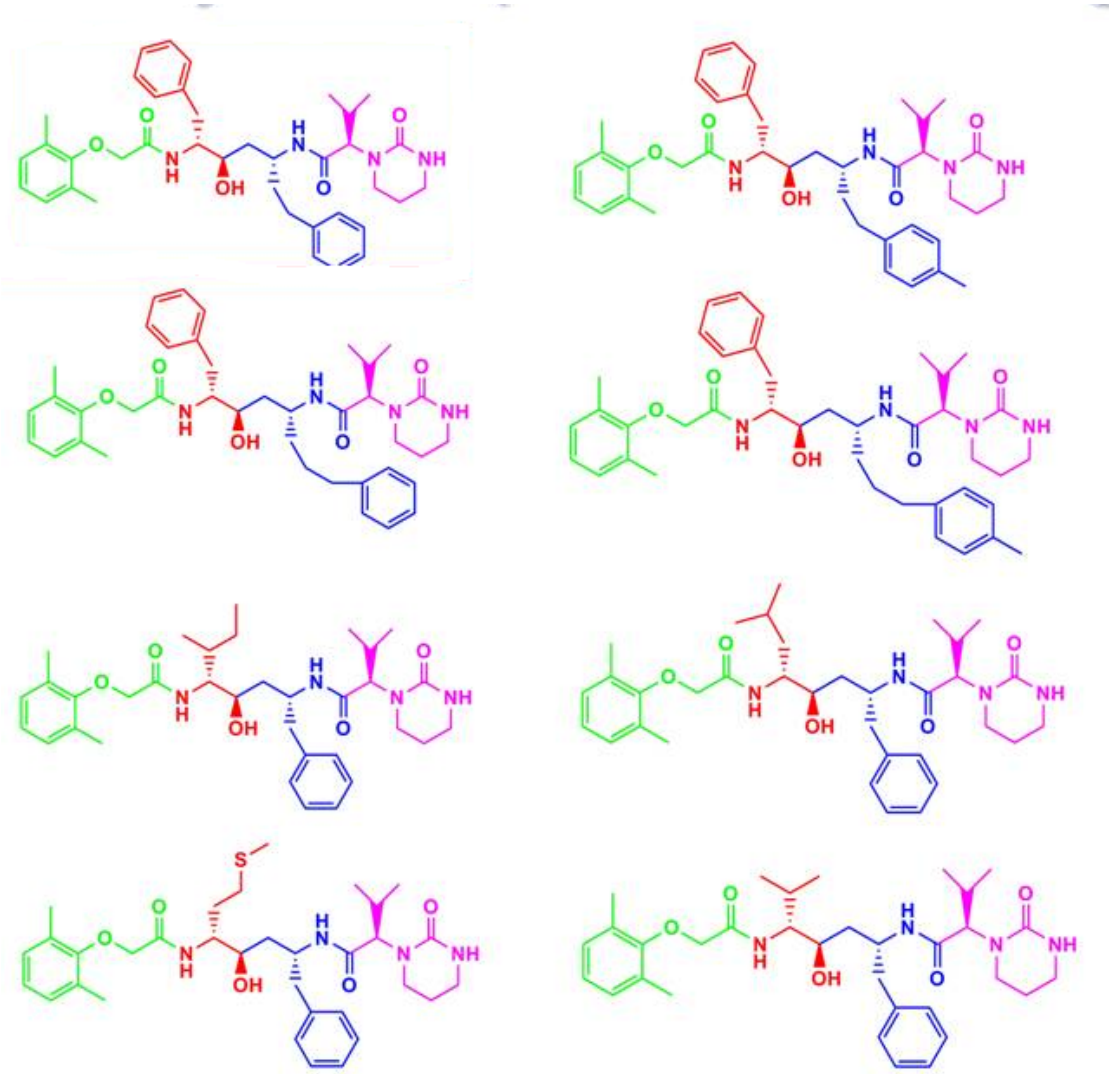


Figure 2.6 Proposed library of lopinavir analogs. The green color, red color, blue color and purple color represents the P2, P1, P1' and P2' of lopinavir respectively. P2 and P2' were kept the same as the original lopinavir. P1 and P1' underwent different modifications to restore the binding affinity of lopinavir to MDR HIV-1 protease. This figure listed part of the proposed library considering the tremendous combination of different modifications. The first four lopinavir analogs show the modifications at P1's site, while the last four lopinavir analogs contain modification in P1 site.

www.manaraa.com

2.10 Construction of protease – substrate models

2.10.1 Construction of wild type HIV-1 protease – substrate (Gag-Capsid) model

For the substrate, we used the NMR (nuclear magnetic resonance spectroscopy) based solution structure of a large protein, the MA4CA sub domain of Gag that includes the C-terminus last four residues (Ser-Gln-Asn-Tyr) of Matrix and full length CA totaling 150 residues. We superposed the Cα atoms of P2'-P1'-P1-P2 residues from MA4CA onto the equivalent peptide substrate residues of the crystal structure, 1KJ4 (205), with a root mean square (RMS) delta of 0.4 Å using the LSQKAB from the CCP4 (193) suite of programs. We then replaced the peptide coordinates in 1KJ4 with the substrate, MA4CA as shown in Figure 2.7 to construct the wild type protease with substrate docked into the active site cavity.

www.manaraa.com

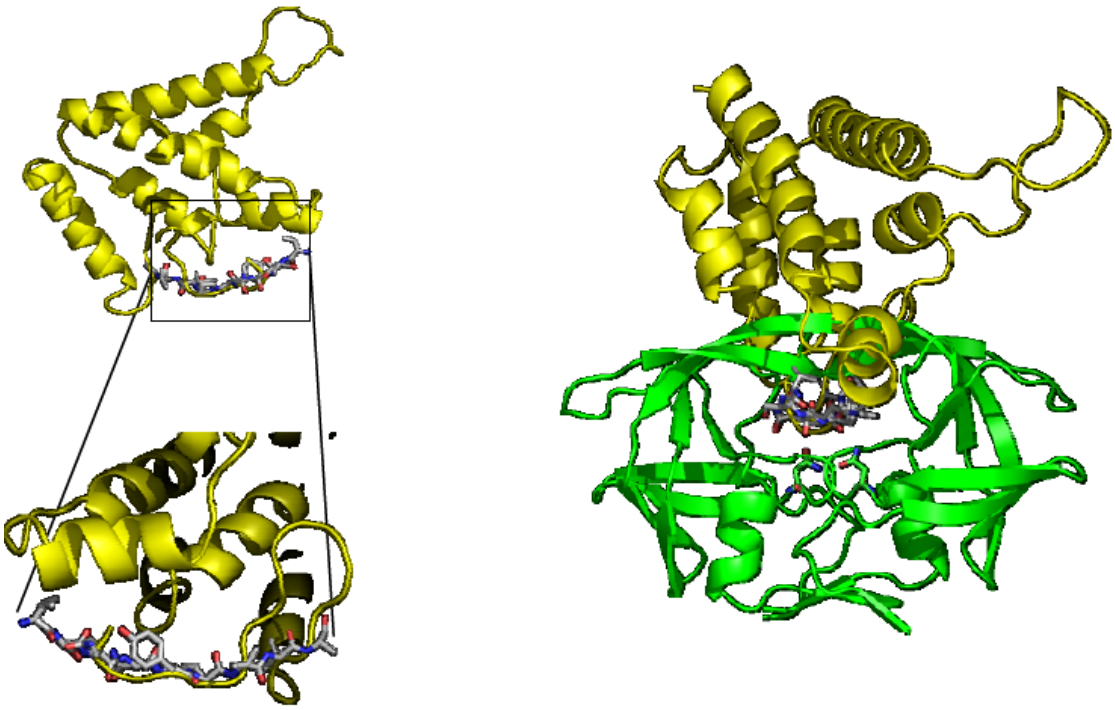


Figure 2.7 Modeling the Gag-Capsid substrate into the active site cavity of HIV-1 protease. The N-terminus of the NMR structure of Capsid (shown in yellow on the left) with MA/CA cleavage site sequence (SQNY/PIV) was superposed onto the corresponding residues in the heptapeptide taken from the crystal structure 1KJ4 (wild type protease in complex with the MA/CA cleavage site heptapeptide). The Capsid was then docked into the active site cavity of the wild type protease by replacing the peptide coordinates in 1KJ4 to obtain the final model shown on the right.

www.manaraa.com

2.10.2 Construction of MDR769HIV-1 protease – substrate (Gag-Capsid) model

Mutation N25D was introduced in the 1TW7 (206) crystal structure in both chains of the MDR protease dimer prior to the substrate docking. Both the catalytic aspartates were unprotonated to represent physiological pH in wild type as well as MDR models. The substrate was then docked into the expanded active site cavity of the MDR protease by superposing the MDR protease onto the wild type protease-substrate model. Residues 1 to 7, 24 to 29 and 95 to 99 were chosen for superposing the MDR protease onto the corresponding Cα atoms of the wild type protease in the model. The coordinates for the substrate from the wild type HIV-1 protease-substrate model were then copied into the MDR769 coordinate file. Both the initial models showed multiple steric clashes of the substrate with the protease flaps (Figure 2.8) indicating the requirement of major conformational changes in the protease flaps before substrate docking.

www.manaraa.com

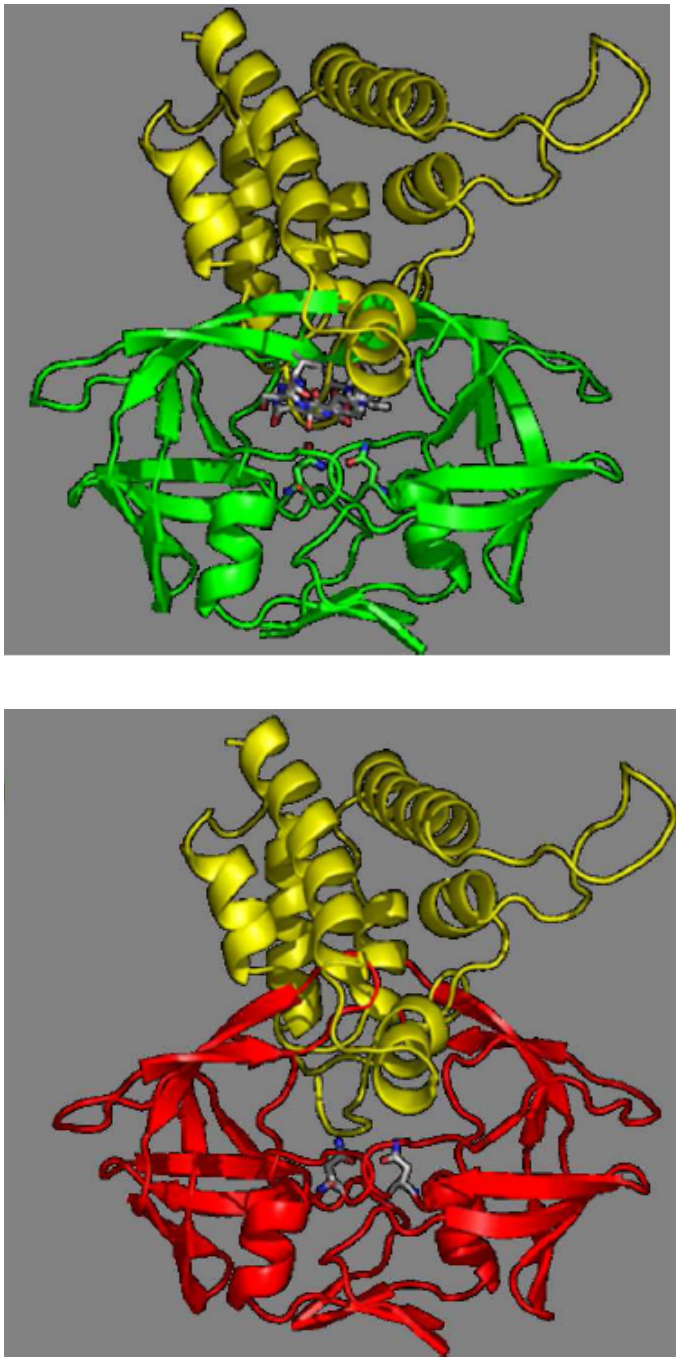


Figure 2.8 Initial models of the protease-substrate complexes. Shown in green is the wild type HIV-1 protease taken from 1KJ4 and in red is the MDR769HIV-1 protease taken from 1TW7, both docked with the substrate (Gag-Capsid shown in yellow) into their active site cavities. Both models show multiple steric clashes with the protease flaps

www.manaraa.com

2.11 Molecular dynamics simulations

The apo-protease structures taken from 1TW7 and 1KJ4 (peptide coordinates were deleted) were initially subjected to a 5 nano second MD simulations. All MD simulations were run using the GROMACS (207) package with the GROMOS96 force field at a temperature of 300 K and a pressure of 1 bar. The temperature and pressure were loosely coupled using Berendsen's procedure. Crystallographic waters were deleted from all the starting structures and the final models were solvated with the SPC (single point charge) (208) water model. Long range electrostatics was incorporated using the PME (particle mesh Ewald) method with a 9 Å cutoff. After 5 nano seconds of MD, the substrate was docked into the active site cavity and both the models (wild type protease-substrate complex model and MDR protease-substrate complex model) were then further subjected to a 5 nano second MD simulation using the protocol described earlier.

www.manaraa.com

CHAPTER 3: CRYSTAL STRUCTURE OF FOUR MDR769 HIV -1 PROTEASE VARIANTS

3.1 Introduction

3.1.1 Wild type HIV-1 protease

HIV-1 protease is a homodimeric, aspartic protease that is critical for the viral maturation (209) and infectivity (210). Each monomer of the protease consists of 99 amino acids. Because disabling the protease would significantly slow down the viral infection, HIV-1 protease has become one of the most important drug targets to design inhibitors to combat the HIV/AIDS problem. Error prone viral replication randomly incorporates multiple mutations resulting in the emergence of multidrugresistant strains under the selection pressure of various treatment regimens. HIV-1 protease is one of the viral proteins that have been studied in detail to understand the affect of various mutations.

3.1.2 Crystal structures of WT and MDR HIV-1 protease

It is very important to study the structures of apo-protease to understand the overall stability of the protease dimer as well as the conformational state of various critical domains of the protease such as flaps, 80s loops (Pro79 – Ile84) and active site cavity. As shown in Table 3.1, there is a plethora of structures of HIV-1 protease variants (194,206,211-218), but only two crystal structures of MDR apo-protease strains (showing altered flap conformation) from a total of twelve structures of apo-protease are available to date. The first medium and high resolution crystal structures of a clinical isolate MDR769

www.manaraa.com

HIV-1 protease with a wide-open conformation of the flaps was reported by our group and the second structure of an MDR strain with such open conformation of flaps was reported by another group.

Table 3.1 List of currently available HIV-1 protease apo-protein structures

PDB entry	Resolution (Å)	Space group	No. of mutations	Remarks (reference)
2HVP	3.0	$P4_12_12$	0	Wild type (4)
3HVP	2.8	$P4_12_12$	6	2X mutant (5)
3PHV	2.7	$P4_12_12$	1	Wild type (3)
1HHP	2.7	$P4_12_12$	1	Wild type (6)
1GL6	1.9	$P6_1$	3	Wild type (7)
1LV1	2.1	$P6_1$	3	Wild type (8)
1Q9P	N/A	NMR	5	Wild type (9)
1RPI	1.86	$P4_1$	10	9X mutant (10)
1TW7	1.3	$P4_1$	11	10X mutant (11)
2G69	1.35	$P4_12_12$	6	1X mutant (12)
2HB2	2.3	$P4_12_12$	7	6X mutant (13)
2HB4	2.15	$P4_12_12$	1	Wild type (13)

N/A - Not available

3.1.3 Mutations in HIV-1 protease gene

Among the mutations seen in the wild type HIV-1 protease gene, 32.32% are constituted by the natural polymorphisms (NP) (shown in Figure 3.1a). Such NPs are observed in the strains isolated from the treatment naïve patients. Once the patients are started on specific treatment regimens during the HAART (Highly active anti-retroviral therapy), the drug resistant strains of protease are selected clinically. Depending on the combination cocktail used in the HAART, some patients might eventually get selected for the MDR strains

www.manaraa.com

of the protease. Analysis of the drug resistance mutations (DM) from the Stanford HIV database, revealed that, based on more than 4000 clinical isolates, the DMs constitute 60.60% of the protease gene. Frequency distribution of such NPs and DMs analyzed from the Stanford HIV database is shown in Figure 3.1b. With more than 50% of the gene prone to DMs, designing inhibitors against such a drug target can be hypothetically compared to shooting a moving target.

3.1.4 Multi-drug resistant 769 HIV-1 protease

The MDR769 HIV-1 protease is one such challenging strains of the HIV-1 protease that resists almost all the protease inhibitors available to date in the HAART. The MDR769 HIV-1 protease consists of a set of ten mutations in the protease gene out of which all ten are known DMs and four out of the ten are NPs (Figure 1a). Mutations, such as V82A and I84V due to a change from longer side chain to short side chain, cause an overall expansion of 3 Å in the active site cavity resulting in the loss of contacts with the inhibitors. In addition to the active site expansion, due to the conformational rigidity acquired by the MDR HIV-1 protease from other mutations, the flaps show a wide-open conformation with an inter-flap distance (measured between the Cα atoms of Il50 on either flaps) of 12.25 Å (twice that of the wild type protease). In combination, the expanded active site cavity and the wide-open flaps cause a highly unstable binding of the inhibitor. Thus, the MDR strain is propagated further even in the presence of most potent inhibitors.

www.manaraa.com

```
WT -PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI
NP -PQITLWQRP* V*VRVEEQ*K EALLDTGADD TVVEDI*LSG KWRPKMIGGI
769-PQITLWQRPI VTIKIGGQLK EALLDTGADD TVLEEVNLPG RWKPKLIGGI
DM -PQITLWQRP* I*VRV*GH** EVIIDTGADN TI*QD**LSG KW*PR*VVGV

WT -GGFIKVRQYD QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
NP -GGFIKVKQYE EV**DICG*R **GTVLVGPT PINIIGRNML TQLGSTLNF
769-GGFVKVRQYD QVPIEICGHK VIGTVLVGPT PANVIGRNLM TQIGCTLNF
DM -GGL*RVKEYE *V**D*FG*R ****VVIGAT P*NVVGR**M S*LGFTLNF
```

WT Wild type
NP Natural polymorphism
769 MDR769 Clinical isolate
DM Clinically selected drug resistant mutations
* Indicates more than one possible substitution

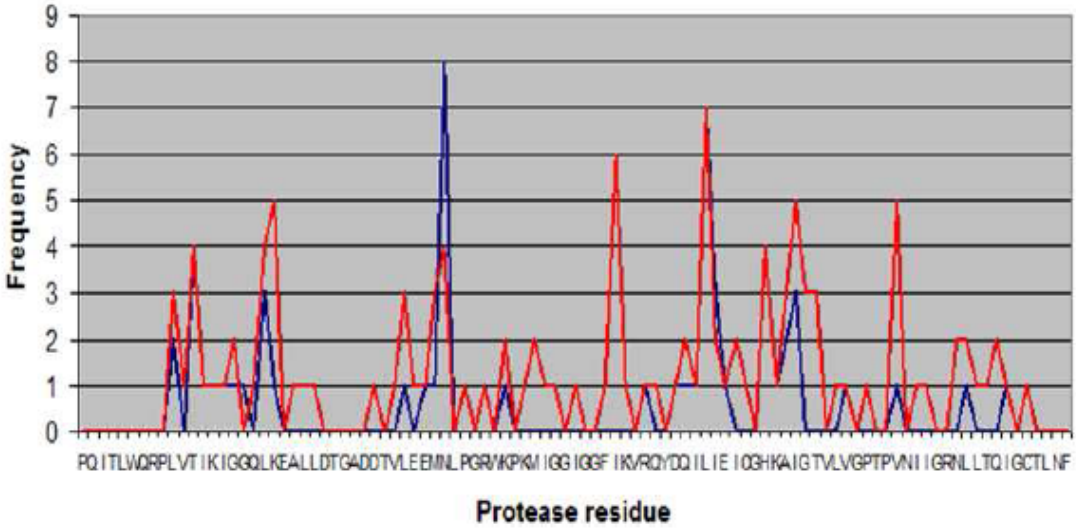


Figure 3.1 Mutations in the HIV-1 protease gene. (a) Amino acid sequence of wild type (NL43) HIV-1 protease aligned with isolate from treatment naïve patients showing the natural polymorphisms (green), the MDR769 clinical isolate (red) and isolate from patients that received treatment showing drug resistance mutations (brown). (b) Frequency distribution of various natural polymorphisms (blue plot) and drug esistance mutations (red plot) in the HIV-1 protease gene. Mutation frequency on y-axis is plotted against the amino acid sequence of the HIV-1 protease.

3.1.5 Variants of multi-drug resistant 769 HIV-1 protease

www.manaraa.com

The current study focuses on understanding the affects of additional point mutations on the MDR protease structure with MDR769 background (consisting of the ten characteristic mutations). Figure 3.2 shows the mapping of the characteristic MDR769 mutations and the additional point mutations further analyzed in this study. There are three reasons to choose these point mutations. First, these mutations represent the side chain changes in either volume or polarity. Second, these mutations show high possibility in DNA random mutations. Third, all these mutations are seen in clinical isolates and documented by the Stanford HIV database.

The active site cavity, flaps, 80s loop and dimerization domains (Figure 3.2) were mainly analyzed. Nine clinically relevant additional point mutations, as summarized in Table 3.2, were selected for further analysis with an MDR769 HIV-1 protease background. Out of the nine point mutations, four mutants (I10V, A82F, A82S and A82T) yielded diffraction quality crystals. Crystallographic analysis of the four structures mainly geared towards understanding: (a) the capability of additional point mutations on causing further changes in the three-dimensional structure of the MDR769 HIV-1 protease making the multidrug-resistance problem worse and (b) if there is a common conserved trend among the four structures that can be targeted to design future inhibitors against an ensemble of the MDR strains.

Before we obtained the crystals and analyzed the crystal structures of the MDR769 HIV-1 protease variants, we proposed that MDR769 HIV-1 protease

www.manaraa.com

represents the end stage form of the HIV-1 protease. As we already know, two independent phenomena contribute the most to the drug resistance of MDR769 HIV-1 protease. One is the wide open flaps, and the other is the expanded active site cavity of the MDR769 HIV-1 protease. Both of the two structural changes make the MDR769 HIV-1 protease incapable of holding inhibitors stably and tightly in its active site cavity. The extent of the structural changes determines the severity of the drug resistance. Larger flap distance and larger active site cavity give rise to more drug resistance. As a result, if additional mutations introduced to the MDR769 HIV-1 protease fail to further open the flaps or to further expand the active site cavity, it indicates that the conformation of MDR769 HIV-1 protease is the most resistant form of HIV-1 protease against FDA approved HIV-1 protease inhibitors and thus the end stage form of the HIV-1 protease. On the other hand, if the additional mutations further open the flaps and enlarge the active site cavity of MDR769 HIV-1 proteaes, then MDR769 HIV-1 protease is an intermediate toward the the most resistant form in the evolution progress. Only the end stage form of HIV-1 protease is worth our further investigation considering the extraordinary mutation rate of HIV.

www.manaraa.com

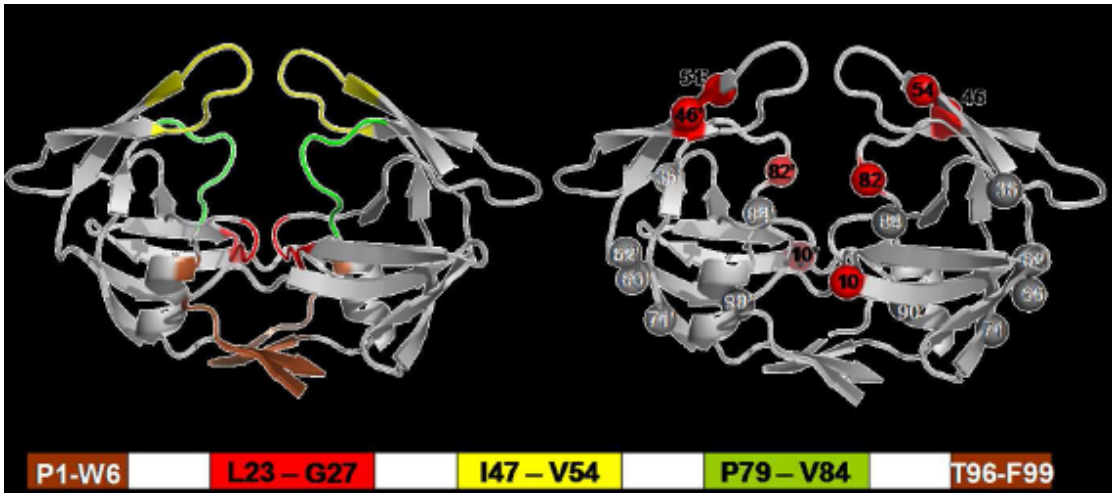


Figure 3.2 Mapping the MDR769 HIV-1 protease. (a) Crystal structure (PDB code: 1TW7) of MDR769 HIV-1 protease showing color coded domains – dimerization domain shown in brown, active site area shown in red, 80s loops shown in green and the wide-open flaps shown in yellow. (b) The characteristic ten mutations of the MDR769 clinical isolate are highlighted in spheres and red spheres are the focus of current study. (c) Domain diagram of the MDR769 HIV-1 protease showing the amino acid residues that constitute different domains. Color codes are same as in (a).

Table 3.2 Mutations chosen in the current study

Wild type residue	MDR769 residue	Mutations introduced
L10	I10	R, V, F
M46	L46	I
I54	V54	M, L
V82	A82	L, S, T

3.2 Results

www.manaraa.com

3.2.1 Crystallization results, electron density maps, and structure analysis

From the ten mutants with MDR769 background, only four mutants (I10V, A82F, A82S and A82T) yielded diffraction quality crystals (Table 3.2). Mutants I10R and I10F did not express adequately and mutants L46I, V54M and V54L yielded crystals that diffracted only to low resolution. Crystal structures of MDR769 HIV-1 protease variants I10V, A82F, A82S and A82T were solved and analyzed as shown in Table 3.3. All the four structures were solved in the space group P41212. The asymmetric unit consists of one molecule of the protease monomer. Protease dimers for each structure were calculated by applying the crystallographic two-fold symmetry. The 2|Fo|-|Fc| electron density maps around the mutation area are shown in Figure 3.3a for each mutant. Contiguous electron density was seen for all the mutants except for A82F, where partial electron density was seen for the side chain of Phe82. As shown in Figure 3.3b, the C α backbone of each mutant was superposed onto that of the wild type HIV-1 protease around the mutation area to search for any conformational changes. No significant conformational changes were observed at the point of additional mutation in each of the corresponding four crystal structures. The four crystal structures were further analyzed for conformational changes in comparison with the native MDR769 HIV-1 protease (PDB code: 1TW7). Most of the conformational changes were seen in

www.manaraa.com

the flaps and the 80s loops among the four mutants when compared to that of the native MDR769 protease.

3.2.2 The proline switch in the MDR769 I10V mutant HIV-1 protease

The crystal structure of the MDR769 I10V HIV-1 protease variant revealed unusual alternate conformations of Pro81 (proline switch) with a 3 Å root mean square difference of the Cα atom of the two conformations. In the wild type HIV-1 protease and MDR769 HIV-1 protease, the Pro81 formed a trans peptide bond, while both cis and trans peptide bonds were observed in MDR769 I10V mutant due to the Pro81 switch. We are the first group to report such a significant conformational change in the 80s loop. The alternate conformations of Pro81 are shown in Figures 3.4a (electron density) and 3.4b (refined model). Shown in Figure 3.4c is the protease dimer showing the conformational change in the 80s loop due to the proline switch. As shown in Figure 3.4d, we propose that Leu23, Glu21, Asn83 and Val84 are key players in transmitting the conformational ripple effect of the mutation at codon 10 (with an MDR769 background) to the 80s loop resulting in the proline switch. As shown in Figure 3.5a, the intra-monomer distortion is caused by the additional 3 Å increase in the distance between the Cα atoms of Ile50 and Pro81 within the same monomer.

www.manaraa.com

Table 3.3 Crystallographic table of four MDR769 HIV-1 protease variants

PDB entry	3PJ6	3OQD	3OQA	3OQ7
Crystal parameters				
Resolution range (Å)	27.66 – 2.25	27.62 – 1.71	27.28 – 2.25	27.23 – 1.71
Space group	P41212	P41212	P41212	P41212
Solvent content (%)	47	47	47.91	47
Data processing				
No. of unique reflections	5356 (768)	12231 (1628)	5466 (764)	11367 (1108)
I/σ(I) [b]	14.2 (5.3)	26.4 (2.8)	20.3 (4.8)	17.6 (2.5)
Rmerge[a] (%)	7.4 (24.7)	6.1 (45.9)	7.1 (37.1)	6.0 (36.8)
Data redundancy[c]	3.4 (3.4)	11.3 (4.4)	6.4 (6.5)	5.8 (3.0)
Completeness (%)	97.6 (99.5)	98.9 (93)	99.7 (99.7)	93.1 (65.2)
Refinement statistics				
No. of reflections used	5099	11591	5196	10806
Rworking[d] (%)	24.3	19.59	20.22	19.5
Rfree[e] (%)	30.41	22.81	23.77	22.39
Mean temperature factors				
Protein	15.82	19.75	26.32	19.59
Main chains	15.4	18.63	25.8	18.57
Side chains	16.28	20.98	26.89	20.71
Waters	8.59	30.61	28	34.49
R.m.s.d. bond lengths (Å)	0.021	0.013	0.019	0.013
R.m.s.d. bond angles	1.94	1.34	1.82	1.4
Ramachandran plot				
Most favored (%)	87.2	96.2	96.2	94.9
Additional allowed (%)	10.3	3.8	3.8	5.1
Generously allowed (%)	2.6	0	0	0
Disallowed (%)	0	0	0	0

[a]*Rmerge*= ΣhklΣi|I(hkl)i-{I(hkl)|/ ΣhklΣiI(hkl)i , where $I(hkl)_i$ is the intensity of an observation and I(hkl) is the mean value for its unique reflection. Summations cover all reflections.

[b]The values in the parethesis indicate the highest resolution shell.

www.manaraa.com

[c]N_{obs}/N_{unique}.

[d]*Rwork*= Σhkl|F(hkl)o- {F(hk|/ΣhklF(hkl)o

[e]R free was calculated same way as R working, but with the reflections excluded from refinement. The R free set was chosen using default parameters in Refmac 5

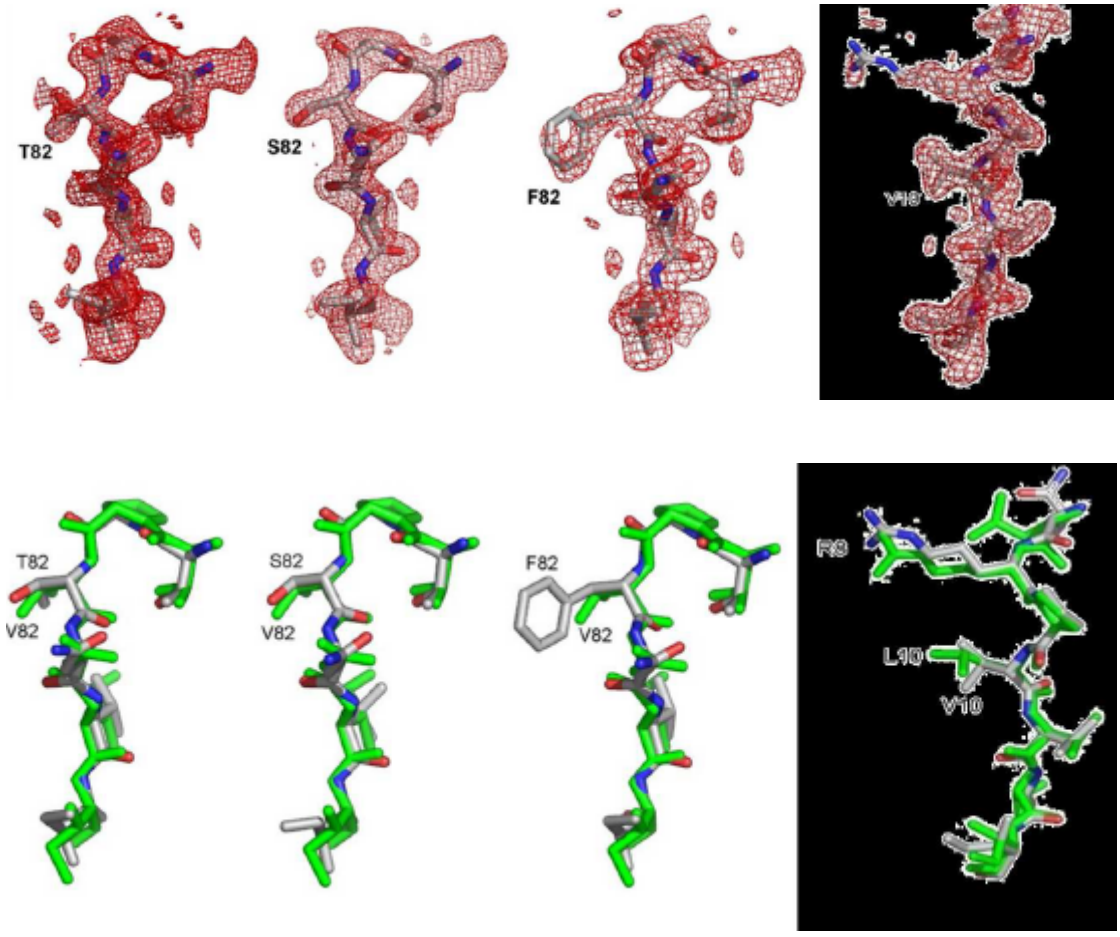


Figure 3.3 electron density maps and analysis of four MDR HIV-1 protease variants (a) Shown in red are the 2|Fo|-|Fc| electron density maps of the four MDR769 HIV-1 protease variants contoured at 1.0 sigma for mutants A82T, A82S and I10V and 0.7 sigma for A82F. Shown in stick models are the residues 80 to

www.manaraa.com

85 for mutants A82F, A82S and A82T and residues 7 to 13 for mutant I10V. (b) The C α atoms of residues 80 to 85 (for mutants A82F, A82S and A82T shown in white) and 7 to 13 (for I10V mutant shown in white) were superposed onto those of the wild type HIV-1 protease (shown in green).

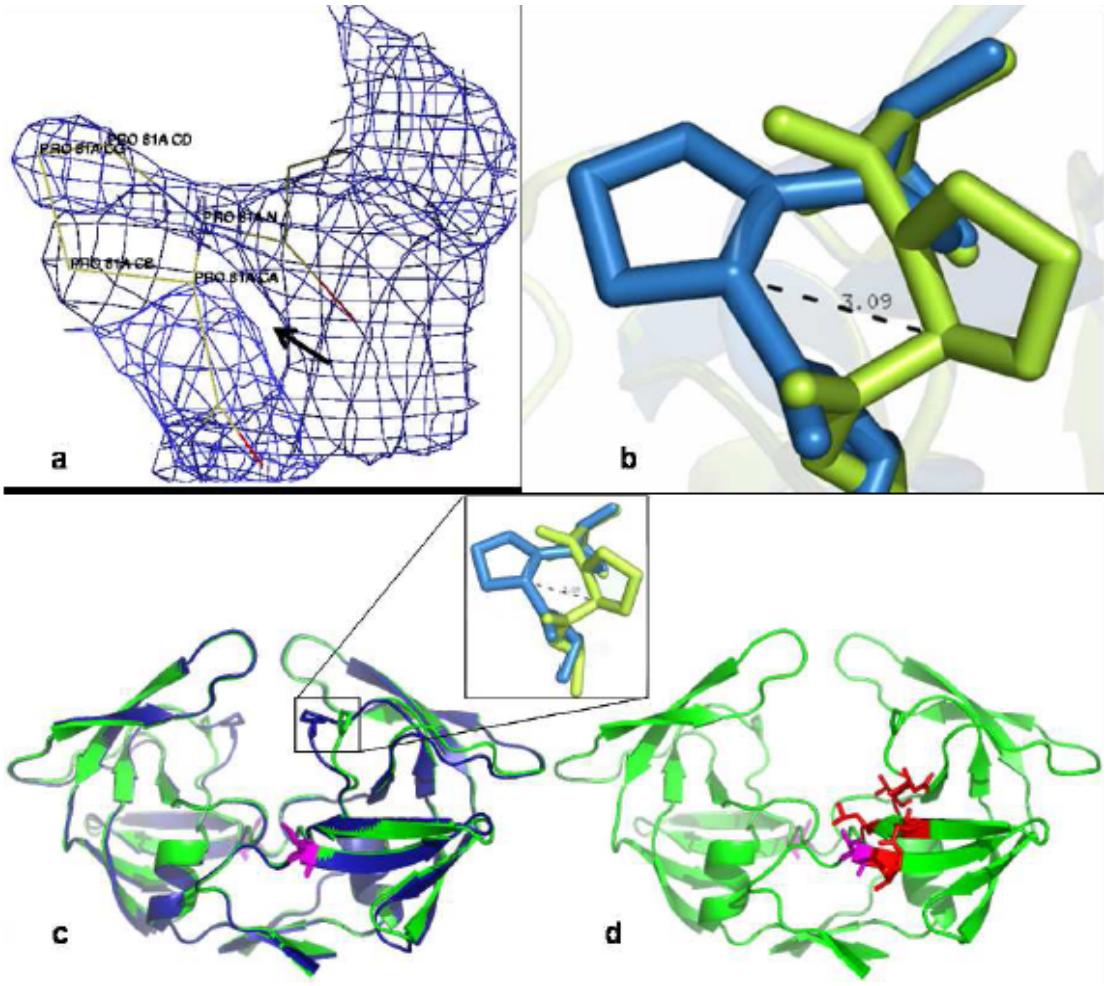
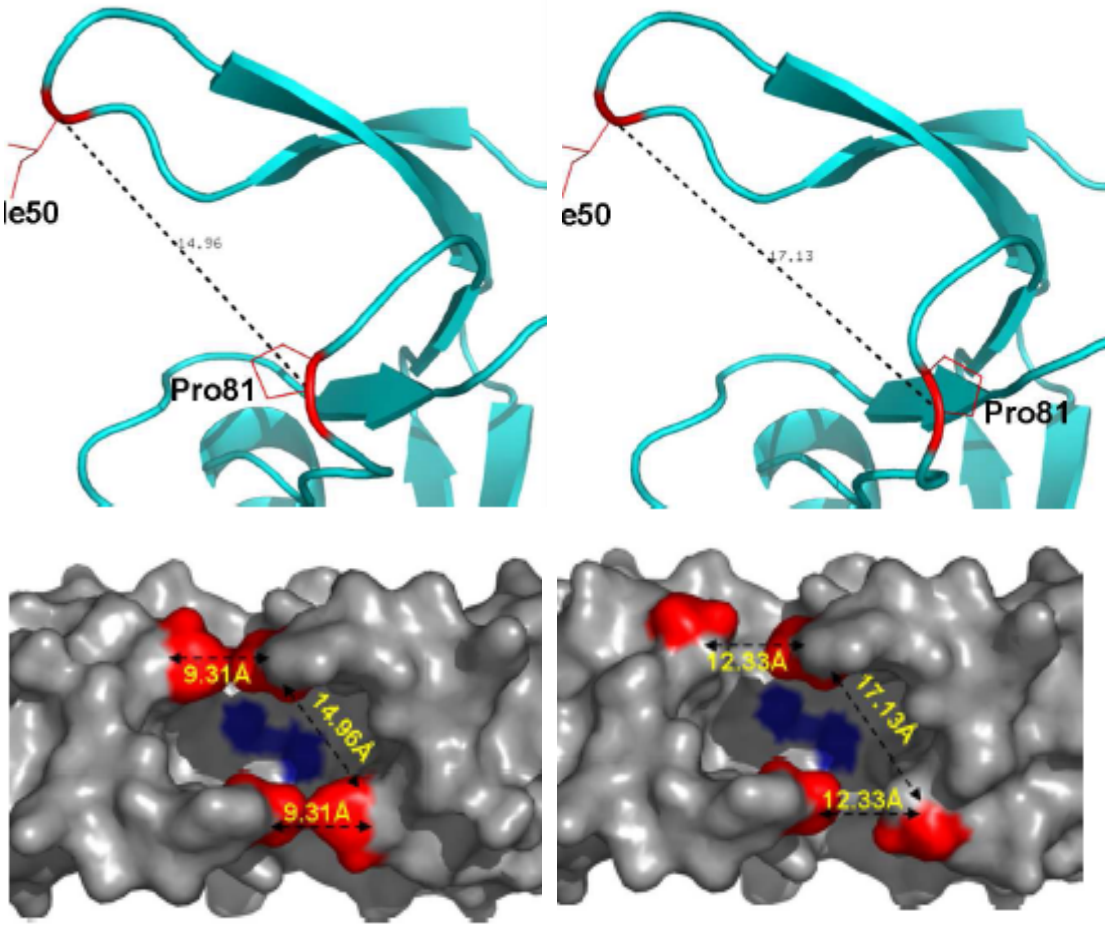


Figure 3.4 The proline switch in MDR769 I10V mutant. (a) 2|Fo| - |Fc| electron density map shown in blue for the alternate conformations of the Pro81 in the MDR769 I10V HIV-1 protease crystal structure. The black arrow points to the gap between the back bones of the two conformations of the Pro81. (b) Stick models of

www.manaraa.com

the Pro81 showing the alternate conformations with a 3 Å shift in the Cα atoms of the either conformations. Shown in blue is the more stable conformation pointing towards active site cavity and the green is relatively less stable conformation pointing away from the active site cavity. (c) MDR769 I10V HIV-1 protease dimer with alternate conformations of the Pro81 (shown in the inset). Mutation I10V is highlighted in pink. (d) MDR769 I10V mutant HIV-1 protease dimer showing residues Glu21, Leu23, Asn83 and Val84 highlighted in red color stick models that are key players in the transmission of structural changes from codon 10 to the 80s loop.



www.manaraa.com

Figure 3.5 The intra and inter monomer distances in I10V mutant. (a) The MDR769 I10V HIV-1 protease monomer is shown in cyan color with Ile 50 and Pro81 highlighted in red. On the left is shown the intramonomer distance (14.96 Å) between the Cα atoms of Ile50 and Pro81 within the monomer with regular conformation of the Pro81. Shown on the right is the same protease with proline switch where the intramonomer distance is 17.13 Å which is greater than 2 Å contributing to the distortion of S1/S1' pocket. (b) Surface diagram of the I10V mutant showing the top view looking into the active site cavity (Asn25 and Asn125 highlighted in blue). Shown on the left are the van der Waals interactions between the Ile50 (red) of one monomer and Pro81 (red) of the second monomer with a distance of 9.31 Å between the Cα atoms. Shown on the right is the same structure with the proline switch where the distance has increased to 12.33 Å. This change causes unstable binding of the inhibitors.

3.2.3 The flipped-out conformation of Phe82 side chain in the MDR769 A82F mutant

It was hypothesized that a mutation from short side chain to a bulky side chain should restore the lost contacts with inhibitors but instead, our crystal structure of MDR769 A82F variant showed a flipped-out conformation of the Phe82 side chain as shown in Figure 3.6. The electron density for the terminal ring portion of the Phe82 side chain was discontinuous indicating a possibility

www.manaraa.com

of flexible side chain. No alternate conformations were found for the side chain of Phe82 even at 0.7 sigma level contoured maps.

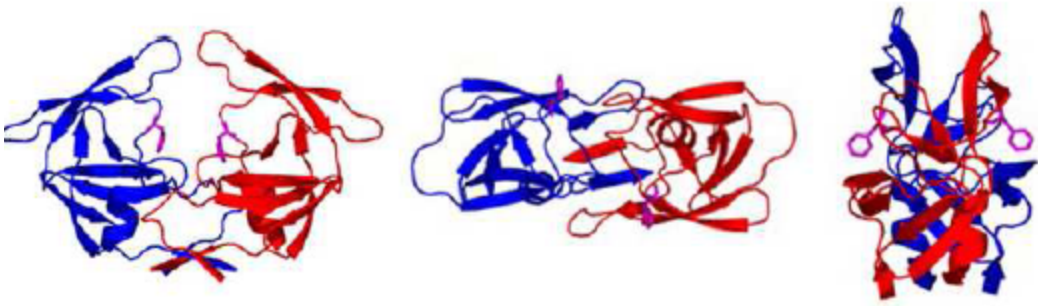


Figure 3.6 Crystal structure of the MDR769 A82F HIV-1 protease. (a) Front view of the protease dimer with Phe82 shown in red color stick model. (b) Top view of the protease dimer looking into the active site cavity. Phe82 is highlighted in red color stick model. The Phe82 ring points out of the active site cavity. (c) Side view of the protease dimer with Phe82 highlighted in red color stick model to show the characteristic flipped-out conformation of the Phe82 side chain.

3.2.4 The A82S and A82T mutants

Crystal structures of MDR769 A82S and A82T HIV-1 protease variants show that the hydroxyl group on the side chain is pointing away from the active site cavity in A82S while it is pointing towards Leu23 in A82T. Due to the Pro81 side chain next to Thr82 in the A82T mutant, the hydroxyl group of Thr82 is repelled away from Pro81 towards Leu23 which is at a farther distance from

www.manaraa.com

Thr82. As shown in Figure 3. 7, the introduction of polarity at position 82 results in the local change of electrostatics of S1/S1' binding pockets.

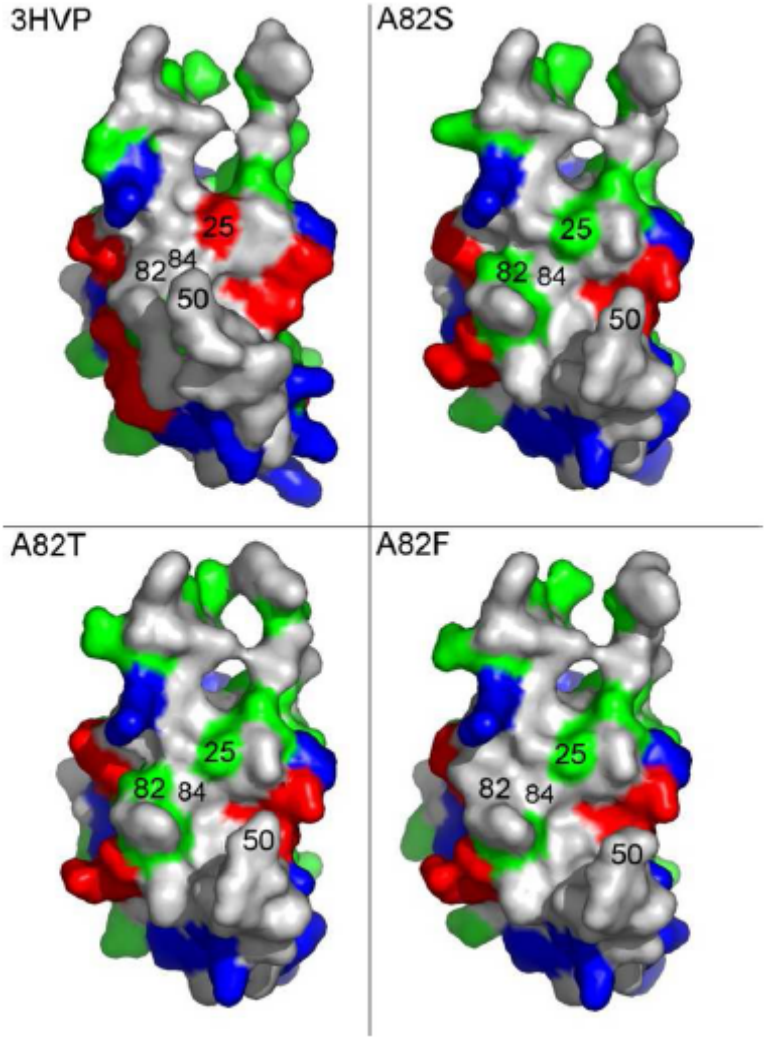


Figure 3.7 Electrostatic potential surface diagrams of wild type (PDB code: 3HVP) monomer protease compared to that of the MDR769 A82S, A82T and A82F mutant protease monomers. The red, blue, green and white colors represent the regions that are negatively charged, positively charged, uncharged polar and hydrophobic respectively. Mutants A82S and A82T show increased polar uncharged regions in the S1/S1' pockets that cause unstable binding

www.manaraa.com

of inhibitors with hydrophobic P1/P1' groups. The A82F mutant shows similar distribution as that of the wild type in the S1/S1' pocket due to the bulky Phe82 hydrophobic side chain.

3.2.5 Dimer interface calculation and analysis

Crystal structures of all the four mutants were solved as monomers in the space group P41212. Crystallographic two-fold symmetry was applied to obtain the dimers for each mutant. These dimers are perfectly symmetrical and represent the functional biological unit. Calculation and analysis of the dimer interface contacts and surface area suggested that all four mutants including the native MDR769 show fewer contacts compared to that of the wild type protease. As shown in Figure 3.8a, among the four mutants, the I10V and A82T show more contacts at the dimer interface compared to that of the native MDR769 while mutants A82F and A82S show fewer contacts compared to that of the native MDR769. Although all the mutants, including the native MDR769, showed similar molecular weight to surface area ratio compared to the wild type protease at the dimer interface, the distribution of the surface area was found to be different. The wild type and A82F mutant protease dimers showed relatively more surface area at the dimer interface (Figure 3.8b). The inter-flap distance was calculated between the Cα atoms of Ile50 of either flap for all the mutants of WT (pdb code 1KJ4), MDR769, MDR769 I10V, MDR769 A82F, MDR769 A82S and MDR769 A82T HIV-1 potease are 6.02 Å, 12.25 Å, 12.09

www.manaraa.com

Å, 12.21 Å, 11.17 Å, and 10.35 Å respectively . All the four mutants show a conserved trend of wide-open flaps with varying inter-flap distances. The overall dimer stability and wide-open flaps were comparable to that of the native MDR769 for all the four mutants.

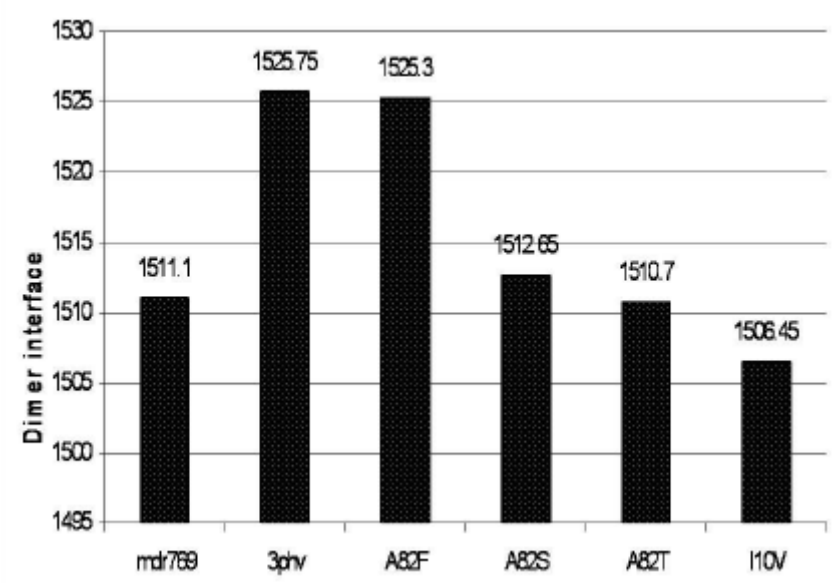


Figure 3.8 Protease dimer interface analysis. (a) Total number of contacts made by the atoms at dimer interface for the wild type and MDR mutants. All the four mutants including the native MDR769 dimer have relatively fewer contacts compared to those of the wild type HIV-1 protease. Among the four mutants, the I10V and A82T show relatively more contacts than those of the mutant A82S and A82F. The wild type protease dimer is more stable than that of the MDR variants. (b) Surface area at the dimer interface is relatively more for the wild type and MDR769 A82F HIV-1 protease variants compared to the native MDR769 and the other two mutants.

www.manaraa.com

3.3 Discussion

3.3.1 Expanded active cavity with wide open flaps

The crystal structures of all the four mutants showed an overall conserved trend of expanded active site cavity with wide-open flaps. In the case of the I10V mutant, due to the alternate conformation of the Pro81, when the side chain is pointing towards the active site cavity, the protease structure looks normal and comparable to the native MDR769 HIV-1 protease but when the side chain is pointing away from the active site cavity, it exerts conformational stress on the 80s loop leading to the distortion of the S1/S1' binding pocket besides. This results in unstable binding of the inhibitors due to loss of contacts. This could be one of the most important mechanisms by which MDR769 and other MDR clinical isolates are selected during the HAART. The MDR769 I10V mutant HIV-1 protease represents the most expanded active site cavity reported so far in the literature. The proline switch also causes the loss of inter-monomer van der Waals interactions between the flap and the 80s loop as shown in Figure 3.5b. In addition to the original expansion of the active site cavity of the native MDR769 strain, the I10V mutant with the proline switch shows an overall expansion of the active site cavity by 6 Å (Figure 3.5b) compared to that of the wild type HIV-1 protease. Loss of inter-monomer contacts in combination with the intra-monomer distortions caused due to proline switch may also weaken the protease dimer. Protease variants like

www.manaraa.com

MDR769 I10V can be very challenging to the most potent inhibitors available to date but on the other hand, such variants are lethal to the viral maturation unless there are compensatory mutations in the substrate for a productive catalysis.

3.3.2 The flipped-out conformation of Phe82 side chain in the MDR769 A82F mutant

The bulky side chain of Phe82 increases the conformational stress on the flexible 80s loop that causes its flipped-out conformation. This unusual conformation not only causes distortion in the S1/S1' binding pockets but also may create a pulling affect on the monomers apart resulting in a weaker dimer. Such protease variant can pose threat to the most potent inhibitors as well as can decrease the viral replication/maturation unless the substrate molecule has compensatory mutations to accommodate the distorted active site cavity of the protease.

3.3.3 Hydrophilic changes at the HIV-1 protease S/S1' binding pocket due to the mutations at codon 82

Crystallographic analysis of MDR769 A82S and A82T structures suggests that if residue 82 is mutated from a hydrophobic residue such as Val (wild type) or Ala (native MDR769) to a polar residue such as Ser or Thr, some of the critical hydrophobic contacts with the inhibitor especially at the P1/P1' region

www.manaraa.com

are disrupted. Considering the architecture of the side chain, since Thr mimics Val (wild type protease) more than Ser, the A82T mutant has less distorted S1/S1' pocket compared to that of the A82S mutant. Nevertheless, both mutants challenge the hydrophobic P1 and P1' functional groups of the inhibitors with A82S being relatively severe compared to A82T. Thus, the decreased hydrophobicity in the S1/S1' binding pocket together with the wide-open conformation of the flaps result in unstable binding of the inhibitors leading to drug resistance. Unfortunately protease variants such as the MDR769 A82T can be productive in the substrate catalysis aiding successful propagation of the MDR strains.

3.3.4 The clinical significance of the four mutants on HIV-1 protease

The clinical significance of the four mutants (I10V, A82F, A82S and A82T) was analyzed in detail from the Stanford HIV database. The frequency of various DMs seen in a pool of clinical isolates that consist of a representative mutation – I10V or A82F or A82S or A82T was analyzed. The A82F mutant is associated with at least 8 other DMs based on 177 clinical isolates from treated patients. Similarly, A82S and A82T mutations are associated with at least 6 and 7 other DMs respectively within their pools of clinical isolates from treated patients. On the other hand, the I10V mutation was only associated with two other mutations (L63P and L90M) based on 495 clinical isolates (1.5

www.manaraa.com

to 5 times the size of the isolate pools for the other three mutants). This indicates that the mutation I10V is very potent in inducing drug resistance probably by changing the three-dimensional structure of the protease resulting in the Pro81 switch. Thus, our current study provides new insights into the multidrug resistance mechanisms of MDR strains of HIV-1 protease.

3.4 Conclusions

In summary, we have crystallized, solved and analyzed the structures of four MDR769 HIV-1 protease variants. Crystal structure of I10V mutant showed that mutation at codon 10 causes a conformational ripple affect leading to the alternate conformations of Pro81 (proline switch). Crystal structures of A82F, A82S and A82T indicate that mutations at codon 82 may not necessarily restore the lost contacts and that the decreased hydrophobicity of S1/S1' pockets further causes loss of contacts with inhibitors.

Based on the four crystal structures presented in this article, we conclude that the flexibility caused in 80s loops (either directly due to mutations such as A82F, A82S and A82T or indirectly due to mutations such as I10V) not only cause loss of contacts with the inhibitors resulting in the multidrug-resistance problem but also result in a weaker HIV-1 protease dimer that is functionally not as active as the wild type enzyme. However, such protease variants can still function when there are compensatory mutations in the substrate cleavage site to accommodate the conformational distortions in the protease binding

www.manaraa.com

pockets and the active site cavity. This study further helps to understand the design of future inhibitors with enhanced potency.

The MDR769 HIV-1 protaese presents the end stage of the HIV-1 protease evolution in the respect of drug resistance. All the additional mutations fail to increase further the flap distance and to expand further the active site cavity. Without further worsening of these two structural elements, the MDR769 HIV-1 protease structure is an excellent study target for us to investigate the rapidly emerging drug resistance issue of HIV infection.

www.manaraa.com

CHAPTER 4: MODELING OF MDR HIV-1 PROTEASE GAG-CAPSID (MA/CA) COMPLEXES

4.1 Introduction

4.1.1 HIV-1 protease flap distance and drug resistance

To design the best inhibitors, it is very important to understand the structure and dynamics (219-220) of HIV-1 protease involved in binding and cleavage of bigger substrates (221). Previously we reported a high-resolution crystal structure (206) of multidrug-resistant (MDR) clinical isolate 769 (222) HIV-1 protease showing an expanded active site cavity (194) and wide-open flaps. According to the active site expansion hypothesis, mutations from longer side chain residues to shorter side chain residues decrease contact with the protease inhibitors (PIs), thereby, decreasing their binding affinity and potency. As illustrated in Figure. 4.1, the wild type protease shows semi-open (apo-enzyme) (211) and closed (ligand bound) (205) conformations of the flaps while the MDR protease shows wide-open flaps. Numerous crystal structures of wild type HIV-1 protease (analyzed from the Protein Data Bank (PDB)) showed closed, open and semi-open conformations of the protease flaps with varying inter-flap distances. The inter-flap distance of the MDR protease is roughly 6-8 Å more than that of the wild type protease. All the inter-flap distances were measured between the C α of Ile50 on either flap. To date, there is only one another group that has published crystal structures of HIV-1 protease variants showing altered conformation of the flaps (218).

www.manaraa.com

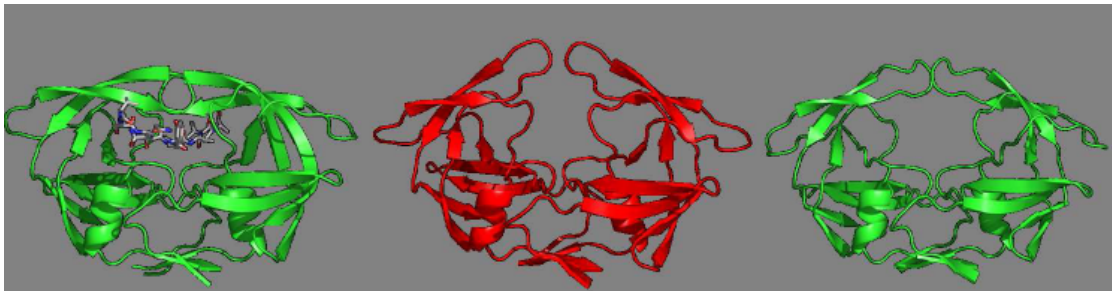


Figure 4.1 Crystal structures of HIV-1 protease with different flap conformations. The wild type (green) protease structures showing closed (PDB code: 1KJ4) on the left and semi-open (3PHV) on the right conformations of the flaps. The MDR769 HIV-1 protease (red) shows wide-open conformation of the flaps.

4.1.2 Wide open conformation of MDR769 HIV-1 protease flaps

This chapter is focused on answering the question, whether the wide-open conformation of the flaps in MDR769 HIV-1 is due to the conformational rigidity attained by the accumulation of mutations or if it represents one of the thermodynamically favorable substrate binding conformations. We hypothesized that: (a) if the wide-open conformation of the MDR protease is caused by mutation induced conformational rigidity, then the flaps should stay open upon energy minimization and/or molecular dynamics (MD) simulations (contrary to what is observed in the wild type protease(223)) due to the steric hindrance caused by the mutations in the protease gene (or) (b) if the wide-open conformation of the MDR protease is representative of one of the favorable substrate binding conformations, then the substrate molecule can be easily modeled into the active site cavity of the MDR protease without steric

www.manaraa.com

clashes with the protease flaps.

4.1.3 Energy minimization of MDR769 HIV-1 protease and WT control

In order to understand the flap movement in the MDR protease, the crystal structure 1TW7 (MDR769 HIV-1 apo-protease) was energy minimized in the absence of crystal contacts. 1000 steps of energy minimization showed that the flaps of the MDR protease moveed closer towards each other while still maintaining an inter-flap distance of at least 10 Å. As a control, we performed the same study using wild type (PDB code: 3PHV) apo-protease, which showed the normal flap closure with a final inter-flap distance of 4 Å. This clearly suggests that the stable wide-open nature of the MDR protease is not entirely due to the crystal contacts (224) as reported by other groups (225) because the energy minimization was done in the absence of crystallographic symmetry atoms. So we reasoned that the wide-open nature of the MDR protease could be due to the steric hindrance caused by individual mutations or by all the ten mutations (accumulated in the MDR protease gene) in concert. These mutations have changed the three dimensional structure of the MDR protease permanently such that the MDR protease shows a conformational threshold point for flap closure beyond which the conformational rigidity hinders the flap closure.

www.manaraa.com

4.1.4 Construction of protease – substrate models

In order to verify if the wide-open nature of the MDR protease represents one of the thermodynamically favorable substrate binding conformations, we modeled the natural substrate, Gag consisting of the Capsid (CA) domain only including the Matrix/Capsid (MA/CA) cleavage site, into the expanded active site cavity of the MDR protease and performed MD simulations. Initially we have constructed a similar model using wild type protease to use it as a control in our MD studies.

4.2 Results and Discussion

4.2.1 The characteristic wide-open conformation of the MDR769HIV-1 protease flaps

Our modeling studies revealed that the substrate (Gag-Capsid) could not be modeled into the expanded active site cavity of the MDR769 HIV-1 protease without steric clashes. This indicates that the wide-open flaps of MDR769 HIV-1 protease (as is from the crystal structure PDB code:1TW7) are not open enough to accommodate a natural substrate and needs at least two-fold further opening of the wide-open flaps of the MDR protease. On the other hand, the molecular dynamics simulations of the apo-protease structure of the MDR protease as well as protease-substrate complex model showed that the wide-open flaps of MDR769 HIV-1 protease never close beyond the threshold point set by the conformational rigidity of the MDR protease structure. Thus, the accumulation of mutations permanently change the

www.manaraa.com

three-dimensional structure of the MDR769HIV-1 protease resulting in a steric hindrance for the flap closure even in the presence of substrate modeled into the active site cavity. This steric hindrance gives the characteristic wide-open conformation to the MDR769HIV-1 protease flaps.

4.2.2 Molecular dynamics simulations show two-fold and three-fold opening of the flaps in the MDR and wild type HIV-1 proteases

Successful crystallographic docking of the substrate (Gag-Capsid) was achieved only after simulating the apo proteases until the flaps are completely open. As shown in Figure 4.2, the wild type and MDR apo-proteases were initially subjected to a 5 ns MD to achieve a completely open conformation of the flaps. The wild type protease showed a three-fold opening of the flaps compared to its starting structure (IKJ4) and the MDR protease showed a two-fold opening of the flaps compared to its starting structure (1TW7). With the flaps completely open (inter-flap distance ~ 20 Å) the substrate was successfully docked into the active site cavity of both wild type as well as the MDR proteases. These protease-substrate complex models were further subjected to 5 ns of MD as shown in Figure 4.3. Thus, the wide-open conformation of the MDR protease does not represent one of the thermodynamically stable substrate binding conformations because; the flaps had to open further to accommodate the substrate.

www.manaraa.com

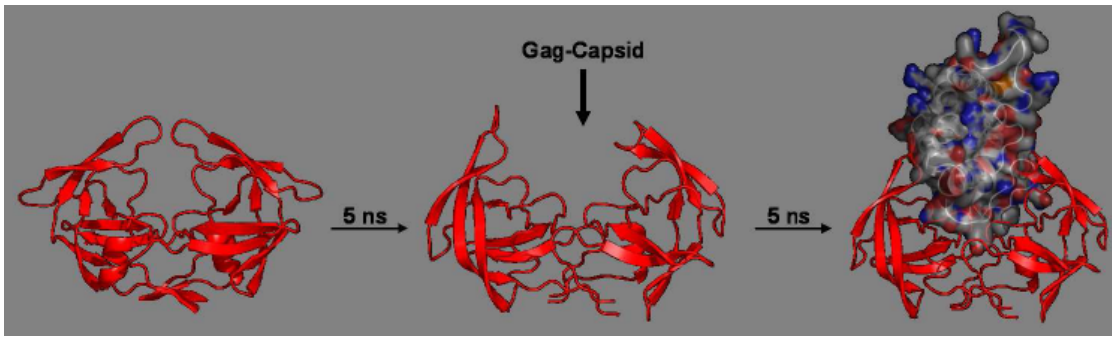


Figure 4.2 Substrate docking and the MD simulation method. Shown in red is the MDR769HIV-1 protease. On the left, the apo-protease before substrate docking, shows an inter-flap distance of 12.25 Å (characteristic wide-open flaps). In the middle, the same apo-protease after 5 ns of MD, shows a completely open conformation of the flaps with an inter-flap distance of 25 Å. The substrate (Gag-Capsid) is then docked into the active site cavity without any steric clashes and the enzyme-substrate model is further subjected to a 5 ns MD to obtain the final complex model shown on the right.

www.manaraa.com

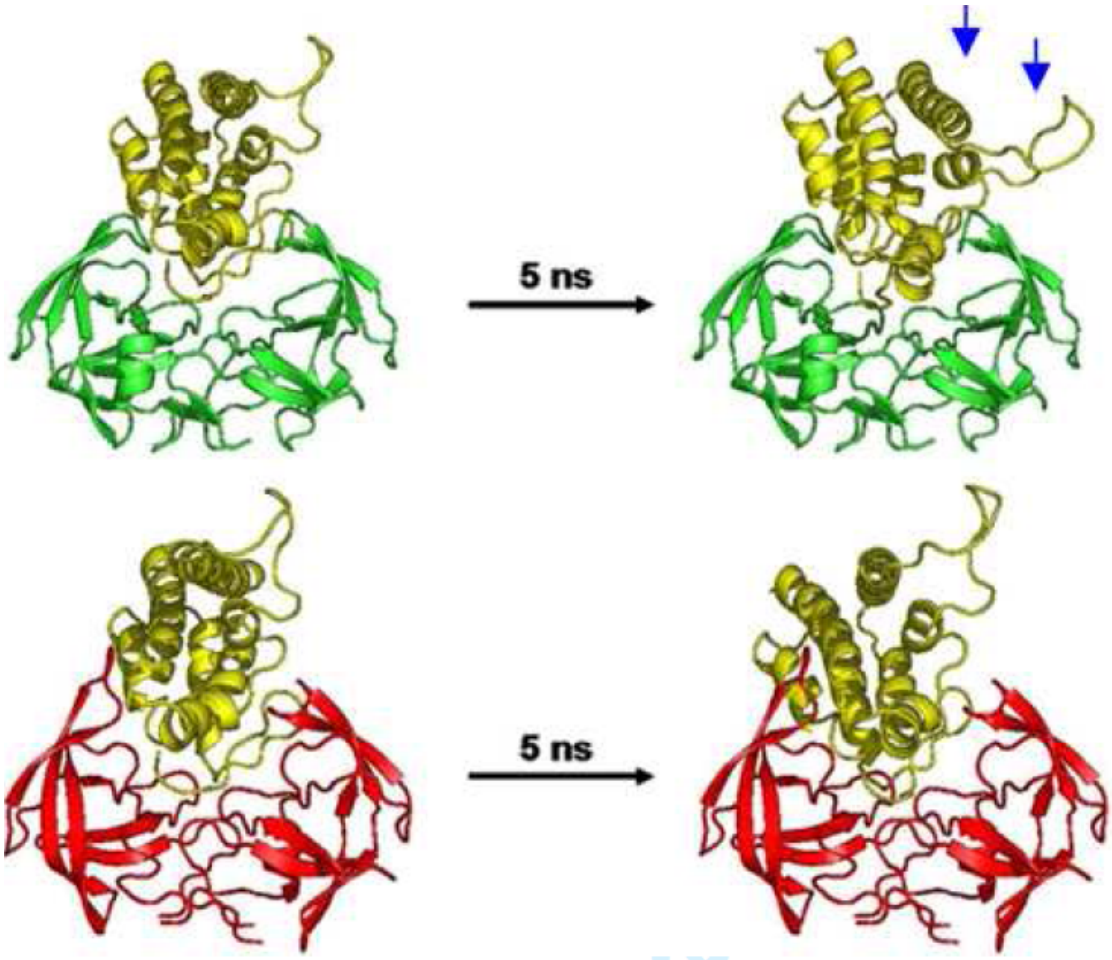


Figure 4.3 Molecular dynamics simulations of the protease-substrate complex models. On the left, are shown the substrate (yellow) docked wild type (green) and MDR769 (red) HIV-1 protease models without any steric clashes. After 5 ns of the MD simulations, the same models are shown on the right. Conformational changes in the substrate in the wild type protease-substrate model are pointed by the blue pointers.

www.manaraa.com

4.4.3 Substrate (Gag-Capsid) is less stable in the expanded active site cavity of the MDR769HIV-1 protease

The radius of gyration of the MDR protease-substrate complex model did not decrease with time during the MD in comparison to that of the wild type protease-substrate complex model as shown in the Figure 4.4 As evident from the fluctuating radius of gyration of the MDR protease-substrate complex model, the MDR protease failed to form a stable enzyme-substrate complex which could result in unproductive catalysis. The loss of stability for the substrate modeled into the expanded active site cavity of the MDR protease may be due to loss of contacts and uncooperative wide-open nature of the flaps. On the other hand, if the substrate has compensatory mutations, may be the MDR protease could form a stable enzyme-substrate complex. This mechanism of co-evolution (226) has to be further tested and is beyond the scope of the current study.

www.manaraa.com

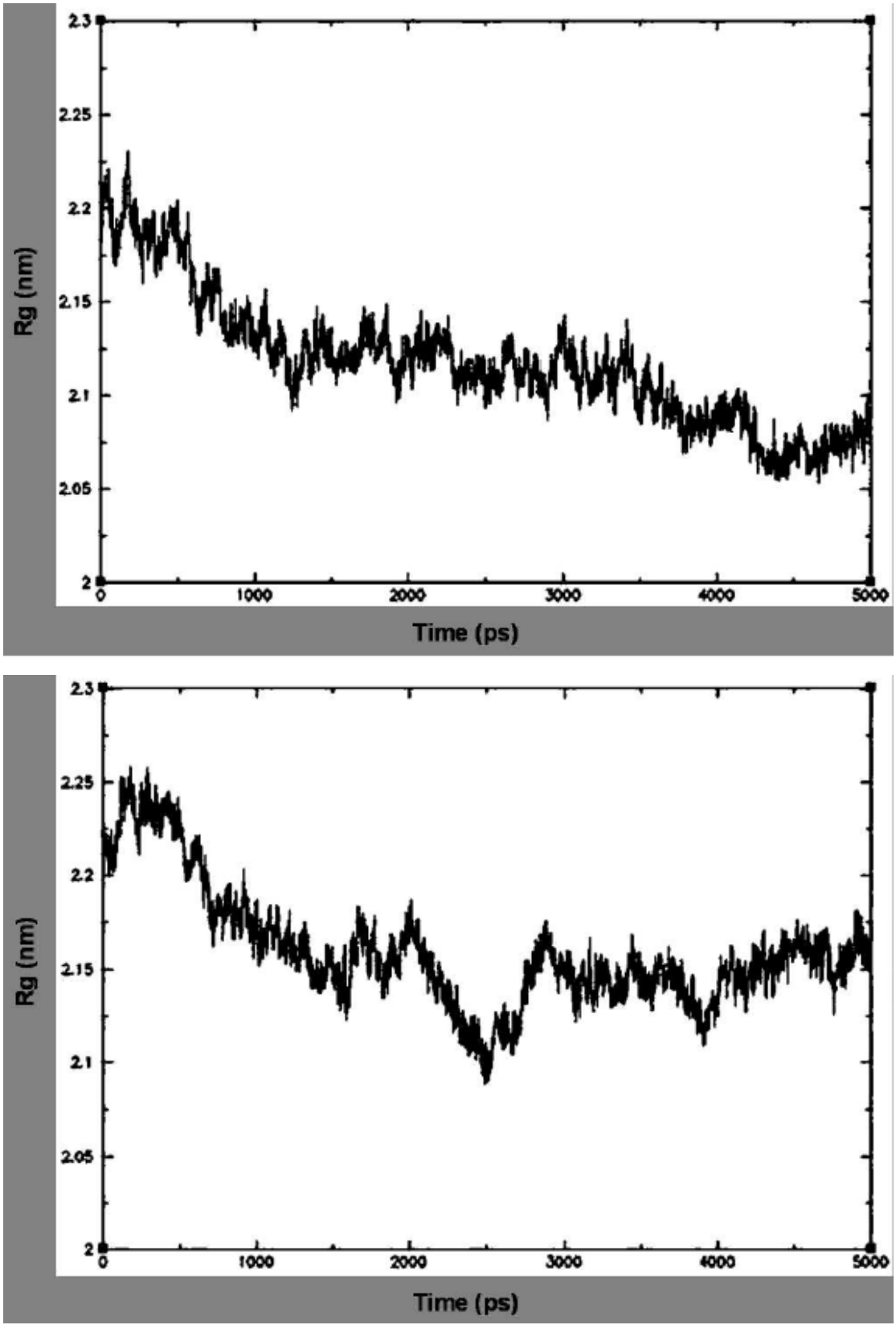


Figure 4.4 Radius of gyration of the models during the MD simulation.

www.manaraa.com

The radius of gyration (Rg - nano meters) is plotted on the y-aixs against the MD simulation time (pico seconds) on the x-axis. During a 5 ns simulation, the wild type protease-substrate complex shows a gradual decrease in Rg (shown in panel a) indicating an increase in the overall compactness of the model. On the other hand, the MDR769 HIV-1 protease-substrate complex shows a fluctuation in Rg (shown in panel b) indicating less compactness.

4.2.4 Conformational changes in the substrate (Gag-Capsid) during the MD simulations

Conformational changes in the substrate (Gag-Capsid) were observed during the MD simulations in the wild type protease-substrate complex model (Figure 4.5a) but not in the MDR protease-substrate complex Figure 4.5b). The substrate modeled into the wild type HIV-1 protease showed a movement of 13 Å in the Ala96 loop as shown in Figure 4.6. This could be due to the alignment or readjustment of the modeled substrate in the active site cavity of the wild type protease towards the formation of a successful enzyme-substrate complex. On the other hand, such conformational changes were not seen in the MDR protease-substrate complex due to the unstable binding of the modeled substrate in the expanded active site cavity of the MDR protease.

www.manaraa.com

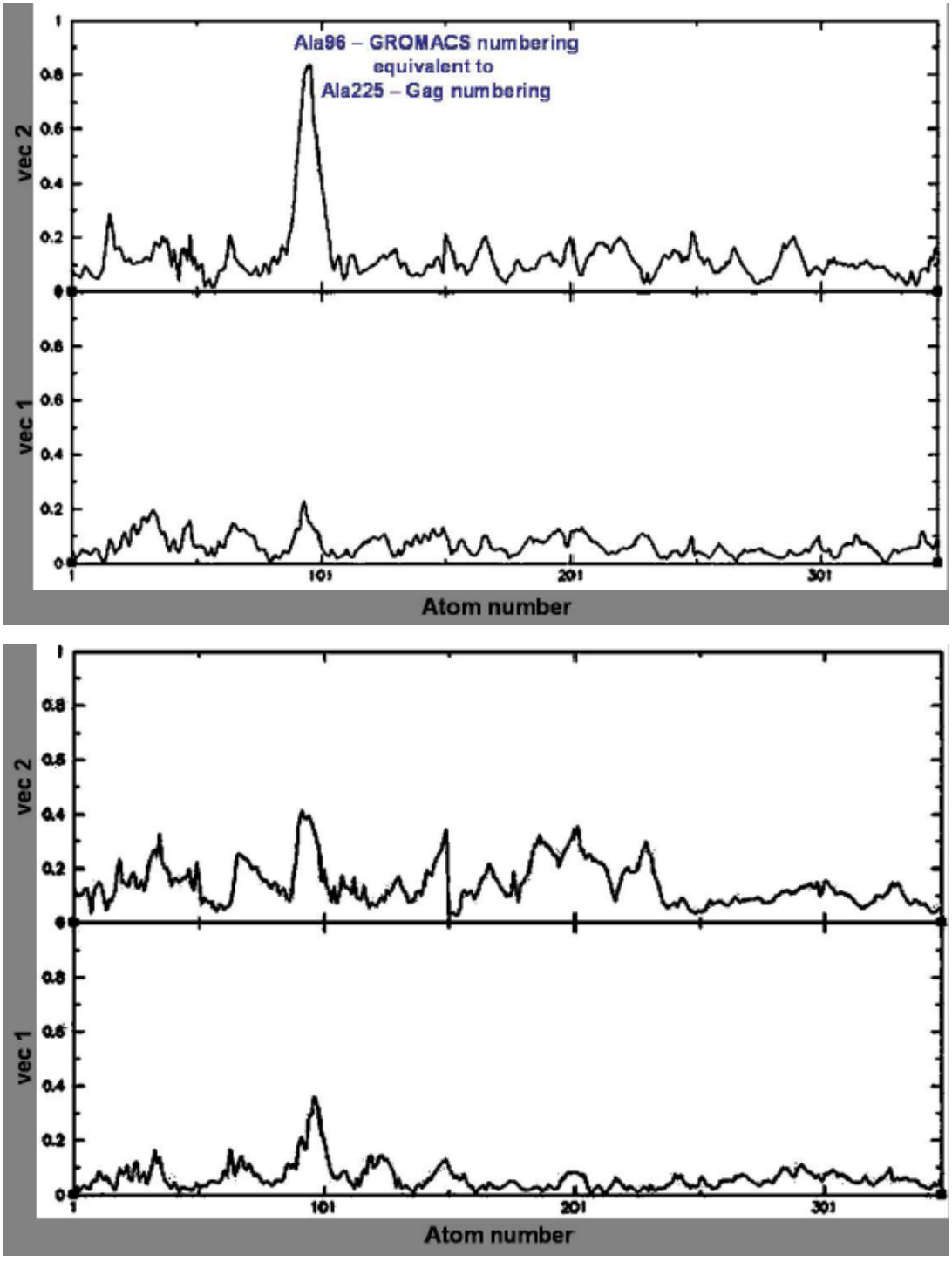


Figure 4.5 Root mean square deviation of the substrate molecule. The substrate docked into the wild type protease shows a significant RMSD at Ala225 indicating a major conformational change (shown in panel a) while no such significant conformational changes were noticed in the case of the MDR769 HIV-1 protease model (shown in panel b)

www.manaraa.com

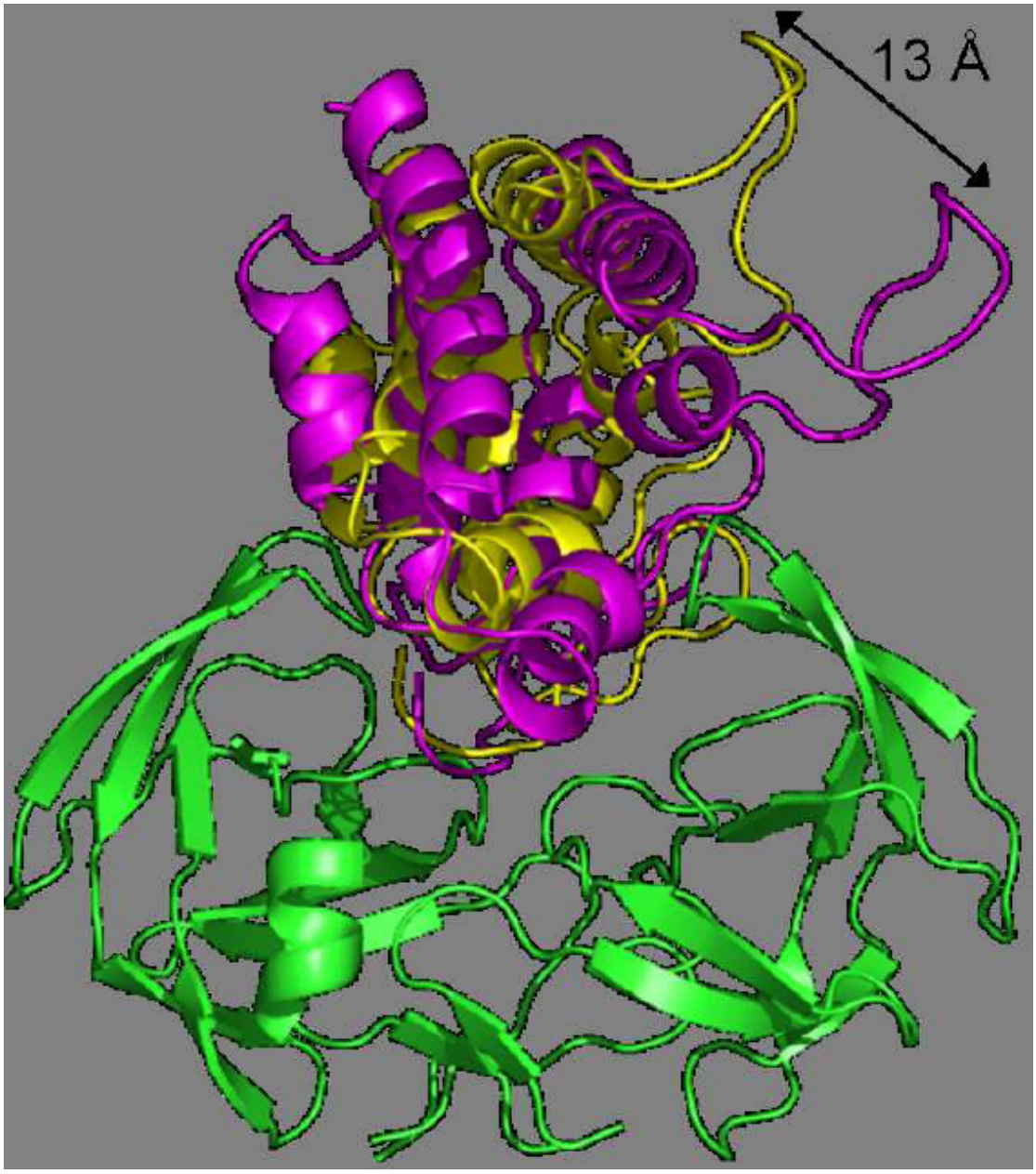


Figure 4.6 Conformational changes in the substrate Gag-Capsid docked into wild type HIV-1 protease model. Wild type HIV-1 protease (green) – substrate Gag-Capsid (yellow) before MD simulations is superposed onto the resultant (purple) model after 5 ns of MD. The Ala225 loop shows a movement of 13 Å.

www.manaraa.com

4.2.5 Alignment of the cleavage site between the catalytic aspartic acid residues

The orientation of the cleavage site (MA/CA) in the substrate (Gag-Capsid) modeled into the wild type protease active site cavity has changed during the MD simulations. In the MDR protease-substrate model, the cleavage site of the substrate appeared to be slipped out of the catalytic site. As shown in Figure 4.7 the substrate cleavage site in MDR protease-substrate complex failed to align the substrate cleavage site between the catalytic aspartic acid residues. Due to the expanded active site cavity and wide-open flaps of the MDR protease, the substrate modeled into the active site cavity was not stable enough to re-orient the MA/CA cleavage site towards a productive catalysis.

www.manaraa.com

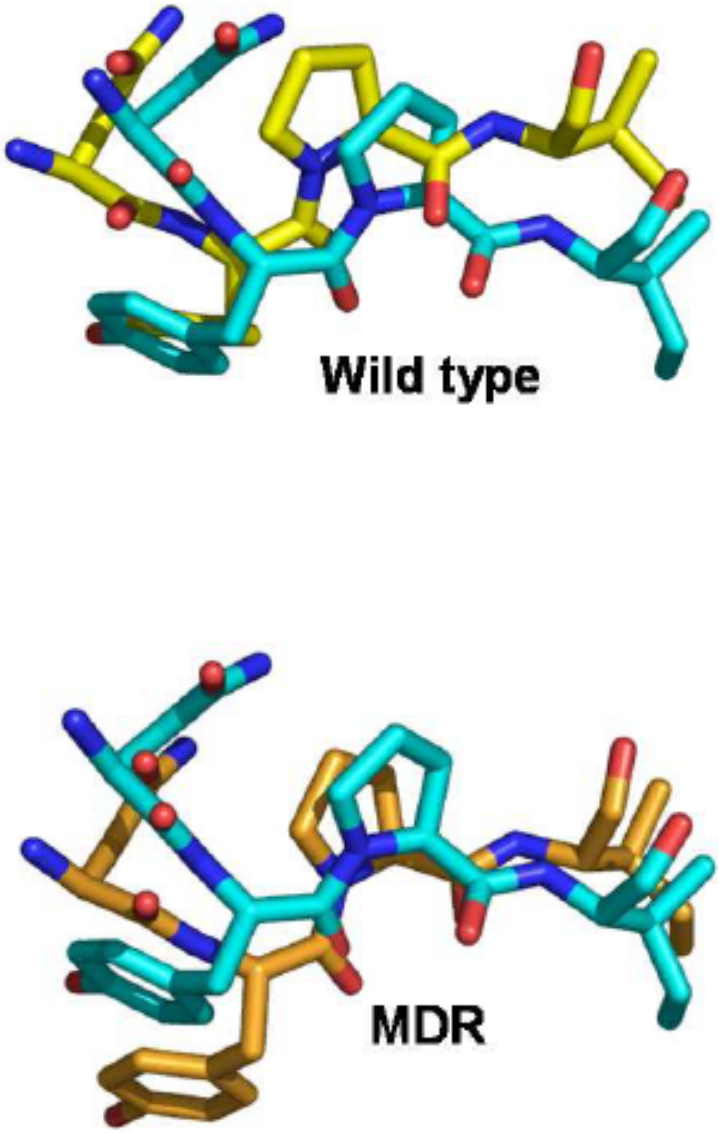


Figure 4.7 Substrate cleavage site alignment in the protease active site. Substrate cleavage site Asn-Tyr-/-Pro-Ile orientation before MD run is shown in cyan color and the same after 5 ns of MD run is shown in yellow for wild type HIV-1 protease complex model and orange for the MDR HIV-1 protease complex model. In the wild type model, the cleavage site is oriented favorably for efficient catalysis, on the other hand, the cleavage site slips out of the catalytic site indicating less efficient catalysis in the MDR model.

www.manaraa.com

4.2.6 Implication of MDR769HIV-1 protease substrate (Gag-Capsid) modeling studies to drug resistance

Based on our analysis, we report that the conformational rigidity of the MDR769 HIV-1 protease, in addition to the expanded active site cavity, prohibits formation of a stable enzyme-substrate complex, which would result in reduced catalysis. In the case of small molecule PIs, we predict a similar behavior where, if the PI is not potent enough to induce closure of the wide-open flaps, then that PI is not stable in the expanded active site cavity. Due to lost contacts resulting in unsuccessful enzyme inhibition, one can expect the emergence of multidrug-resistance. Thus, the MDR769 HIV-1 protease, due to its conformational rigidity, poses a great challenge to the current PIs that are used in HAART and also is a good model to study the MDR strains of HIV-1 protease.

4.3 Conclusions

In summary, our results supports the hypotheses that the conformational rigidity exhibited by the MDR769 HIV-1 protease hinders substrate binding and therefore reduces catalysis. These studies have implications for the design of potent competitive PIs against the MDR strains of HIV-1 protease as well as enhancing the potency of maturation inhibitors (227) such as PA457 (92). In addition, this work may have implications in studies related to viral replication in the presence and absence of compensatory mutations in the substrate-cleavage sites. Based on our findings, we conclude that the

www.manaraa.com

conformational rigidity of the MDR769 HIV-1 protease flaps could be one of the novel mechanisms for the multi-drug resistance nature of the MDR protease and that new approaches are needed for further studies using models like the MDR769 HIV-1 protease.

www.manaraa.com

CHAPTER 5: HIV-1 PROTEASE IC50 MEASUREMENTS WITH LIBRARIES OF REDUCED CA/P2 PEPTIDES AND FDA APPROVED INHIBITORS

5.1 Introduction

5.1.1 Selectivity and specificity of HIV-1 protease

It was reported that MDR HIV-1 proteases show reduced binding affinity to FDA approved protease inhibitors, while they still keep a certain level of enzymatic activity that is required for the viability of the virus itself (228-229). As a result, the HIV-1 protease is highly specific for its natural cleavage site peptides or their analogs (230-231). However, the selectivity of HIV-1 protease is not obvious considering the wide variability of its substrate sequences. Research has been performed to determine the best substrate for a particular peptide sequence through kinetic measurements of hydrolysis of peptides with single amino acids substitution at each position (232-234). *Poorman et al* demonstrated that residues at positions P2, P1 and P2' are critical for the cleavage (235). However, later studies indicated that substrate recognition by HIV-1 protease was achieved by shape of the substrate rather than the sequence (205,236).

5.1.2 Reduced CA/p2 peptide library

CA/p2 cleavage site is the rate-limiting step in all the nine substrate cleavage (237). HIV-1 protease can catalize the cleavage of as many as nine different natural substrates with various catalytic efficiencies. CA/p2 cleavage

www.manaraa.com

site is cleaved by HIV-1 protease with the least efficiency compared to other eight substrate peptides. As a result, we choose CA/p2 as our study model to make a library of the reduced CA/p2 peptide analogs with substitution of one amino acid only in each analog. In addition, the P4' side Met was replaced by NorLeucine to avoid the chemical synthetic obstacles. This library of reduced CA/p2 analogs are contributed by our collaborator, Dr. Mark Spaller from Dartmouth Medical School. The sequences of this reduced CA/p2 peptide library is shown in Table 5.1. The comparison between the regular CA/p2 peptide and reduced CA/p2 peptide is shown in Figure 5.1. We measured the IC50 of this library to both MDR769 HIV-1 protease and WT HIV-1 protease. Similar studies were conducted by Altman *et al* on RT/RH analogs with computational and thermodynamics techniques (238).

5.1.3 HIV-1 protease inhibitors

Similarly, the IC50 values of six FDA approved inhibitors against both MDR769 HIV-1 protease and WT HIV-1 protease were measured to determine the extent to which these inhibitors were resistant.

An HIV-1 protease inhibitor is one of the components of HAART. Currently, there are ten FDA approved HIV-1 protease inhibitors on the market. Most of these drugs contain a hydroxyethylene core and are peptidomimetic competitive inhibitors against the HIV-1 protease by mimicking the natural substrates (239-246) . Tipranavir, though, has a dihydropyrone ring as a central scaffold rather than a peptidomimetic hydroxyethylene core (247).

www.manaraa.com

We have six of all the FDA approved HIV-1 protease inhibitors in our lab, as shown in Figure 5.2. IC50 values of these inhibitors for both WT and MDR769 HIV-1 protease were measured to determine the binding affinity.

5.1.4 Lopinaivr

Lopinavir, approved by FDA in 2000, is the only protease inhibitor that is co-formulated with a low dose of ritonavir -- the combination is known as Kaletra (Abbott). *In vitro* assay showed IC50 of 6-17 nM range against various stains (245). Although drug resistance is relatively less likely to develop for lopinavir compared to other HIV-1 protease inhibitors (248), resistance mutations are noticed, including 32I, 47A, and 46I, or L33F, I54V, V82A, I84V, and L90M, or combinations among L76V, M46I, and V82A in the protease, or A431V in gag (249-253). In addition, it was found that the accumulation of mutations, including mutations of codons 82, 54, and 46, during prior protease inhibitor treatment facilitates the emerging of lopinavir resistant mutations (254).

We expect to have worse IC50 values of MDR769 HIV-1 protease against both reduced CA/p2 analog library and the FDA approved inhibitors compared to those of wild type HIV-1 protease. Based on this working hypothesis, it is proposed that reduced CA/p2 analog library and the FDA approved inhibitors bind to MDR769 HIV-1 protease with reduced affinity. In addition, the analysis of the reduced CA/p2 analog library sequence and FDA approved inhibitor structure, combined with the IC50 data, may lead to some hints in how to

www.manaraa.com

modify currently existing FDA approved HIV-1 protease to overcome the drug resistance issue.

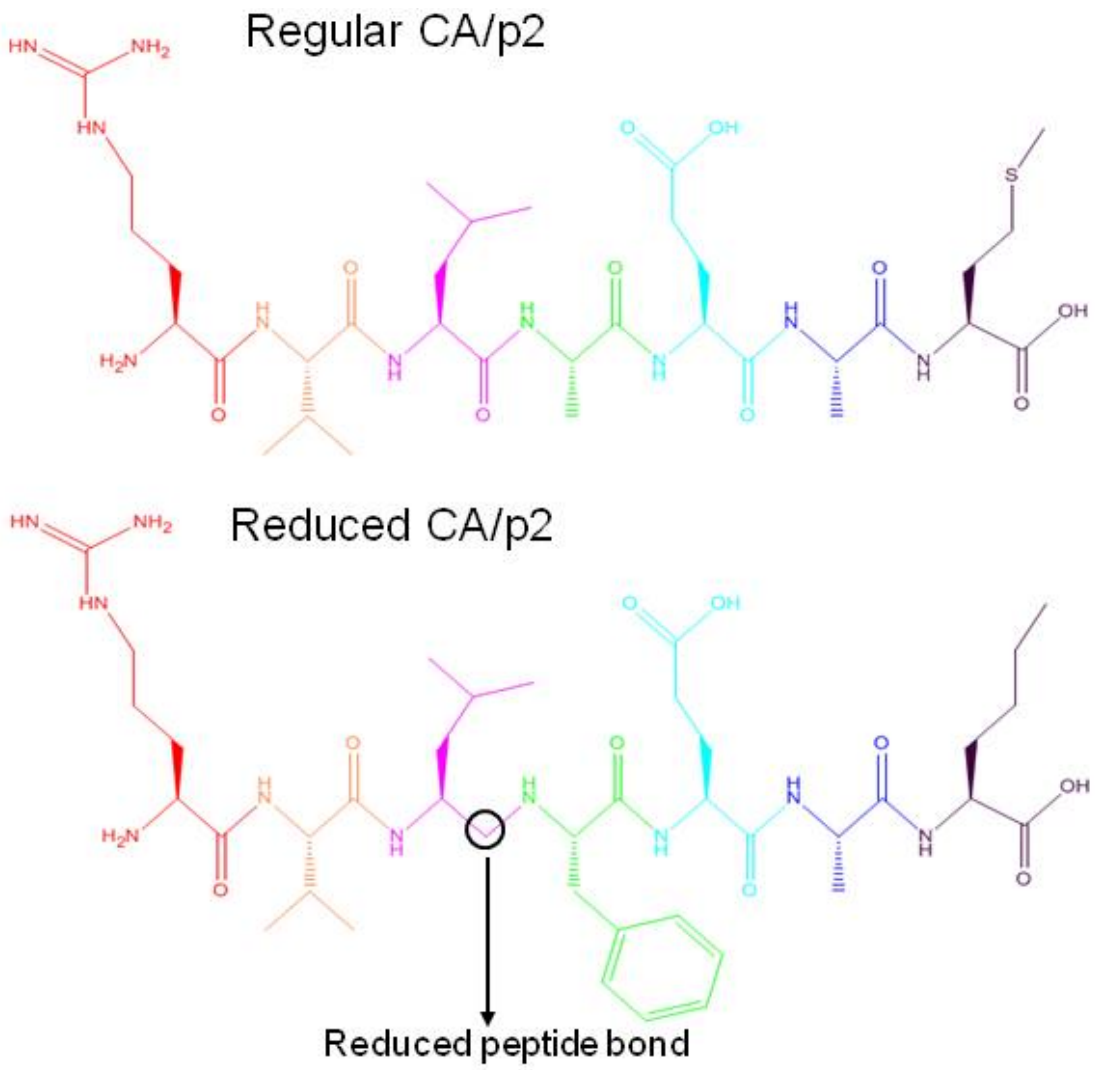
Table 5.1 CA/p2 reduced peptide library[a].

Reduced Peptide	P3	P2	P1		P1'	P2'	P3'	P4'
Original CA/p2	Arg	Val	Leu		Ala	Glu	Ala	Met
P1'F	Arg	Val	Leu	R	Phe	Glu	Ala	NorLeu
P2A	Arg	Ala	Leu	R	Phe	Glu	Ala	NorLeu
P2F	Arg	Phe	Leu	R	Phe	Glu	Ala	NorLeu
P1A	Arg	Val	Ala	R	Phe	Glu	Ala	NorLeu
P1F	Arg	Val	Phe	R	Phe	Glu	Ala	NorLeu
P2'A	Arg	Val	Leu	R	Phe	Ala	Ala	NorLeu
P2'F	Arg	Val	Leu	R	Phe	Phe	Ala	NorLeu
SP1	Arg	Val	Leu					
SP2					Phe	Glu	Ala	NorLeu

[a]The Sequnces of these peptides are listed from P3 site to P4' site. The cleavage site is between P1 and P1'. The first sequence represents the natural CA/p2 sequence with a regular peptide bond between P1 and P1'. The capitalized letter "R" stands for reduced peptide bond and makes the peptides uncleavable by active HIV-1 protease. The mutagenesis is highlighted by red. The last two peptides SP1 and SP2 stand for short peptide 1 and short peptide 2 respectively. They are the N and C terminal cleavage products for natural CA/p2 peptide. In addition, the P4' site is Methionine in natural CA/p2 sequence and is mutated to NorLeu to avoid the chemical synthetic

www.manaraa.com

obstacles.



Fgire 5.1 Comparison of regular and reduced CA/p2 peptide bond. The upper panel is the regular CA/p2 peptide and the lower panel is the reduced CA/p2 peptide. From the left to the right are P3, P2, P1, P1', P2', P3', P4' site respectively. The peptide bond is reduced between P1 and P1' site in the reduced CA/p2 peptide analogs. The P4' site is mutated from Methionine to NorLeucine to avoid chemical synthesis difficulties. P2 to P2' sites are subjected to mutations in order to explore the influences of the binding affinity to HIV-1 protease introduced by different amino acid

www.manaraa.com

residues.

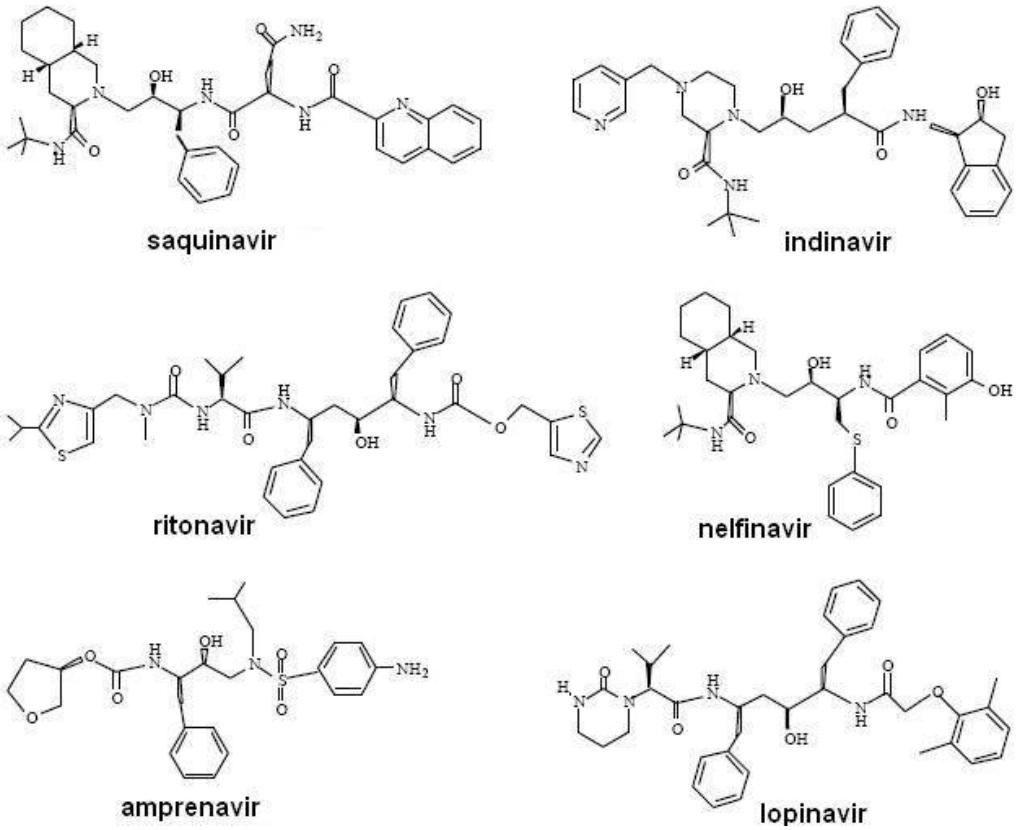


Figure 5.2 Structure of FDA approved HIV-1 protease inhibitors. The structures of six FDA approved inhibitors are shown in this panel. All of these inhibitor contain a peptidomimetic hydroxyethylene core. Among these, lopinavir and ritonavir demonstrate a symmetric P1/P1' conformation, while others have different P1/P1' functional groups.

5.2 Results

5.2.1 IC50 enzyme data with the reduced CA/p2 peptide analogs

www.manaraa.com

The IC50 values for the reduced CA/p2 peptide library are shown in Table 5.2. Similar IC50 values are observed between MDR769 and WT HIV-1 protease against the library of reduced CA/p2 peptide analogs. In addition, the SP1 (short peptide 1) and SP2 (short peptide 2), the cleavage products of CA/p2, both fail to show feedback inhibition against either MDR HIV-1 protease or WT HIV-1 protease.

However, the reduced peptide P1'F with P1' site Ala replaced by Phe shows IC50 4.3 nM against MDR 769 HIV-1 protease, despite being slightly higher than that of WT HIV-1 protease. This IC50 is even comparable to IC50 of some FDA approved HIV-1 protease inhibitors. P1'F with P1 Leu and P1' Phe is a good combination, and we will explore other similar combinations based on this.

Table 5.2 IC50 values of CA/p2 reduced peptide library against both MDR and WT HIV-1 protease

Peptide	WT IC50 (nM)	MDR769 IC 50(nM)
P1'F	2.6	4.3 ± 0.02
P2A	78.17	74.1 ± 10.49
P1F	30.67	142 ± 4
P2'A	232.33	233.67 ± 3.21
P2F	1143.33	770.33 ± 62.96
P2'F	1009.67	1263.33 ±106.93

www.manaraa.com

P1A	4873.33	5020 ± 434.86
SP 1	100000	103866.7 ± 7145.16
SP 2	N/A	664333.3± 72885.5

5.2.2 IC 50 enzyme data with the six FDA approved HIV-1 protease inhibitors

The IC50 values of six FDA approved protease inhibitors, amprenavir, indinavir, lopinaivr, ritonavir, nelfinavir, and saquinavir, against both MDR 769 and WT HIV-1 protease are given in Table 5.3. Analysis of the data shows significant drug resistance of MDR HIV-1 protease against all protease inhibitors except for lopinavir, compared to those of WT HIV-1 protease. The IC50 increase of lopinavir is only 4-folds from WT to MDR HIV-1 protease. As a result, lopinavir is a good starting point to modify the existing FDA approved HIV-1 protease inhibitors in order to restore the inhibitor binding affinity without dramatically altering the ADME (absorption, distribution, metabolism, and elimination) and toxicity of the potential drug candidates.

Table 5.3 IC50 values of FDA approved inhibitors against both MDR and WT HIV-1 protease

Inhibitor	WT (nM)	MDR769 (nM)	IC50 increase

www.manaraa.com

Amprenavir	0.4763 ±0.031	5.94 ±0.145	12.47
Indinavir	2.67 ±0.832	142.3 ±23.46	53.29
Lopinavir	1.237 ±0.139	5.39 ±0.225	4.357
Ritonavir	0.443 ±0.007	56.17 ±5.34	126.79
Nelfinavir	2.18 ±0.1	104 ±1.732	47.71
Saquinavir	0.514 ±0.00436	311.3 ±15.822	605.64

5.3 Discussion

5.3.1 Substrate cleaveage efficiency of WT and MDR HIV-1 proteases

The chemical cleavage step is disrupted by the mutations in MDR769 HIV-1 protease. Two independent steps are required for the process of HIV-1 protease substrates. The first step is the physical binding of the substrates to the HIV-1 protease. The second step is the chemical cleavage of the substrates, which is executed by the residues Asp 25 and Asp125. Our enzyme assay shows similar IC50 values of MDR769 and WT HIV-1 protease against the library of reduced CA/p2 peptide analogs. It indicates that the binding affinity of this library is similar regardless of the MDR or WT HIV-1 protease. Meanwhile, other studies indicate certain degree loss of catalytic acitivity of MDR HIV-1 protease, compared to that of WT HIV-1 protease, in

www.manaraa.com

cleaving the substrate peptides. As a result, it is expected that the chemical cleavage step rather than the physical binding step of the substrate peptide is disrupted by the drug resistance rendering mutations. This chemical cleavage disruption may result from the expanded active site cavity of the MDR HIV-1 protease. Due to the extra space in the active site cavity allowing relatively free movement of the substrate peptides, the substrate peptides are thought to be floating in the active site cavity. This makes the substrate peptide cleavage site less accessible to the two “active residues” Asp 25 and Asp 125.

5.3.2 CA/p2 P1’F reduced peptide shows highest IC50 against both WT and MDR HIV-1 protease

It was surprising to learn that the reduced peptide P1’F inhibits the WT and MDR HIV-1 protease with IC50 values comparable to those of the FDA approved HIV-1 protease inhibitors. Despite the fact that MDR HIV-1 protease was slightly resistant to reduced peptide P1’F, the absolute value of the IC50 indicated that reduced peptide P1’F is a potential drug candidate with P1/P1’ combination Leu and Phe. Other similar combinations of P1 and P1’ sites are explored to optimize the inhibitory activity of the reduced CA/p2 peptide analog. In addition, to further comfirm the importance of the P1/P1’ combination, a shorter peptide library, ranging from P2 to P2’ sites rather than P3 to P4’, will be synthesized and tested against WT and MDR HIV-1 protease.

5.3.3 P1/P1’ site symmetry of lopinaivr may contribute to the

www.manaraa.com

enhanced drug efficacy

Lopinavir shows a four-fold drug resistance by MDR HIV-1 protease. Further analysis of lopinavir and other protease inhibitors reveals a unique characteristic of lopinavir -- the symmetric P1 and P1' sites. Ritonavir, similar to lopinavir, shows a symmetric P1/P1' sites, but does not demonstrate similar pattern of drug resistance. However, it is possible that the bulky volume of ritonavir makes it less potent against MDR HIV-1 protease. The symmetry of P1 and P1' site of lopinavir is under further investigation to understand the potential application of the symmetric sidechain concept to HIV-1 protease structure based drug design, given the symmetric nature of the enzyme. In addition, lopinavir and ritonavir will be cocrystallized with MDR HIV-1 protease to understand in details the mechanism of different patterns of drug resistance.

5.3.4 Implication of the discrepancy between the library of reduced CA/p2 peptide analogs and inhibitor IC50 values

The IC50 values of the library of reduced CA/p2 peptide analogs surprisingly fail to show the different binding affinity of the WT and MDR HIV-1 protease to the substrates. However, the inhibitor enzymatic studies give the expected different binding affinities of these inhibitors to WT and MDR HIV-1 protease, although the difference varies in different inhibitors. The comparison between the reduced CA/p2 analogs and FDA approved HIV-1 protease inhibitors reveals the following differences:

1. The reduced CA/p2 peptides comprise of seven amino acid residues,

www.manaraa.com

while the inhibitor range from P2 to P' sites only, equal to four amino acid residues in length

2. There is one hydroxyl group between the P1 and P1' sites of the protease inhibitor mimicking the crystallographic water critical in substrate cleavage in the HIV-1 prtease substrate complex, while no such corresponding functional group is observed in the reduced CA/p2 peptide analogs
3. The difference in the environment of the hydroxyethylene carbon atom may make a difference in binding affinity. The inhibitors have an R1-CH2-CH(OH)-**C**H2-R2 structure whereas the reduced CA/p2 analogs have an R1-CH2-CH2-**N**H-R2.

Consequently, there are three possible ways to modify the FDA approved inhibitors to overcome the drug resistance issue by restoring the binding affinity to MDR HIV-1 protease. First, expand the existing protease inhibitors. The extent to which the expansion can reach needs to be carefully calibered. On one hand, if the volume is not big enough, the binding affinity is not increased enough to achieve the desirable results. One the other hand, an over-expanded inhibitor may lose its target specificity and lead to intolerable toxicity despite the high activity. In addition, the bioavlaibity may be unfavorable with large inhibitors. Second, the hydroxyl group between P1 and P1' sites may be removed or substituted by other functional groups to generate the modified inhibitors with better IC50 values and higher binding affinity to the

www.manaraa.com

MDR HIV-1 protease. Third, one of the carbon atoms flanking the hydroxyl group between P1 and P1' site may be substituted by a nitrogen atom. Which carbon atom to be replaced cannot be determined based on current functional studies. Further experiment with either of these two carbons replaced by nitrogen atom separately will be helpful in determining the significance of each flanking carbon atoms to the ligand binding affinity to the MDR HIV-1 protease.

www.manaraa.com

CHAPTER 6: CRYSTAL STRUCTURE OF THE EXTRA-CELLULAR DOMAIN OF HUMAN MYELIN PROTEIN ZERO

6.1 Introduction

6.1.1 Function of the myelin protein zero

Myelin protein zero (MPZ) is the causative gene for CMT1B, although it is much less common than CMT1A, accounting for less than five percent of all CMT1 cases (142,151-152).

Myelin protein zero (MPZ), associated with CMT1B, is a transmembrane protein of 219 amino acid residues. Human MPZ consists of three domains: first, 125 residues constitute the glycosylated immunoglobulin-like extracellular domain; second, 27 residues span the membrane; and third, 67 residues represent the highly basic intracellular domain (160-161). MPZ, a homophilic adhesion molecule (162-164), is a member of the immunoglobulin super-family (165) and is essential for normal myelin structure and function.

6.1.2 Structure of the extracellular domain of human MPZ

We report the crystal structure of the extracellular domain of human MPZ (hP0ex) fused with the maltose binding protein (MBP) at resolution 2.1 Å. Despite the fact that the crystal structure of rat MPZ extracellular domain (rP0ex) is available (166), the crystal structure of the human counterpart is indispensible considering several critical amino acid residue variations between the two homologues . There are three sequence differences at amino

www.manaraa.com

acid residues 10, 16 and 77. Residue 10 is a clinically important site where a mutation may lead to phenotypic change. MPZ residues 16 and 77 are at the four-fold interface and head-to-head interface according to the packing of the rat P0ex model. Consequently, the crystal structure of the human P0ex is necessary to understand the mechanism of CMT disease.

6.1.3 Functional studies of clinically important mutants of MPZ

In addition, based on the crystal structure of hP0ex, we investigate the molecular mechanism of pathogenesis of MPZ mutants H10P, R69C, R69H, T95M, and H52R. First, *Li et al* demonstrated that mutant H10P, causes late onset CMT1B and presents severe length-dependent axonal loss without segmental demyelination. Material either in the inner wraps of myelin and/or periaxonal space coexisted with the accumulated MPZ. In addition, the molecular organization of the axolemma (Mauthner sheath or plasma membrane of the axon) showed abnormalities (255). Second, R69C mutations caused an early onset CMT1B (256-258). Further investigation showed regions of uncompacted myelin, one of the two types of pathology as seen in other early onset CMT1B (259-260). Onion bulb formation, thinly myelinated axons (261), and amyelinated fibers were also noticed in this mutant. Third, T95M, a mutation leading to later onset CMT1B, did not interrupt the normal intracellular traffic of MPZ and affected the intercellular adhesion moderately. However, this mutation interfered with the glycosylation of MPZ, which is important for MPZ function. Fourth, mutation H52R, unlike other early onset

www.manaraa.com

CMT1B causing mutations, showed no adverse affects on the intracellular traffic of MPZ. The MPZ with mutation H52R could reach the cell surface but it significantly reduced the intercellular adhesion (262).

6.1.4 Amyloidosis and CMT1B

Recent studies have linked amyloidosis to CMT1B caused by certain mutations such as T95M (263). In amyloidosis, amyloid fibrils, derived from naturally soluble protein or protein fragments, are insoluble and accumulate in a variety of organs and tissues (264-267). The orginal naturally soluble proteins are unrelated and may be rich in β sheet, α helix, or both α helices and β sheets. They may be rigid globular proteins or natively unfolded proteins (268-274). However, the amyloid fibrils contain a common core structure of cross folded β-sheets, consisting of continous β sheets with the β strands running perpendicular to the long axis of the fibrils (275). Immunoglobulin light chain amyloidosis (AL amyloidosis), associated with a plasma-cell dysfunction (276), is the most common amyloidosis in the US (277).The immuolglobulin light chains, light chains and heavy chains, or fragments of heavy chains cannot be broken down to constituent amino acids due to the misfolding (278). The light chains assemble into protofibril and then into amyloid fibril in the form of an amyloid fribril subunit (279). The light chain of immunoglobulin contains two immulogolulin fold structures, while the extra cellular domain of human MPZ has a single immunoglobulin fold structure. It is possible that the X-ray diffraction studies will reveal an amyloid structure of the hP0ex mutants that

www.manaraa.com

causes CMT1B. Further studies of the β-pleated sheet binding to Congo red with its tinctorial properties of green birefringence under polarized light may provide an alternative way to study the hP0ex mutants besides the the unique X-ray diffraction pattern. Amyloid protein diffraction patterns show characteristic perpendicular reflections along the meridional direcction at 4.7 Å and along the equatorial direction at 10 Å (280).

6.2 RESULTS

6.2.1 Crystal structure of the MBP hP0ex fusion

hP0ex fused with MBP crystallized in space group $P2_1$. The crystallographic parameters are shown in Table 6.1 with the final structure refined to 2.1 Å resolution. The electron density maps, contoured at 2 sigma, of the maltose in the MBP and the intra-molecular disulfide bond in the hP0ex are shown in Figure 6.2. Two fusion proteins are included in one crystallographic asymmetric unit, with one maltose ligand in each MBP molecule (Fig. 6.1). The MBP adopts the “closed” conformation in the presence of the maltose molecule. The connecting region (22 amino acid residues, part of the pMal™-p2X vector) between the C-terminus of MBP and the N-terminus of hP0ex was difficult to trace and therefore was omitted due to the poor electron density of these residues. As a result, one fusion protein molecule was treated as two independent chains in the deposited PDB file. Two intact hP0ex (from residue 1 to residue 121) were present in one crystallographic asymmetric unit. The last four residues of the C-terminus, 122 - 125 (Pro, Thr, Arg, and Tyr),

www.manaraa.com

were disordered due to flexibility and therefore omitted. The major interface in the crystal was between two MBP interfaces involving residues Asp 120 (Asn 205, Lys 202 of molecule two), Lys 239 (Asn201 of molecule two), Asp 314 (Lys 142 of molecule two), Pro 315 (Gly 143 of molecule two), and Arg 316 (Ala141 of molecule two). hP0ex Arg 77 only contacted MBP via Glu 359 because of the flexible connecting region; hence, no significant obstruction of the quaternary structure was induced by the close proximity of the large MBP domain.

Table 6.1 Data Collection and Crystallographic Refinement Statistics

Data Collection		
Beam Line	APS (LS-CAT 21)	
Wavelength (Å)	0.9779	
Space group	P2(1)	
Cell dimension (Å) a, b, c	a=61.91	b=54.77
	c=145.85	
α, β, γ	α=90.00	β=98.49
	γ=90.00	
Resolution (Å)	30.00-2.00	
R merge (%)[a]	8.7/53.9[b]	
I/σ(I)	21.6/4.3	
Redundancy	3.5/3.4	
Completeness (%)	99.9/100	
Unique reflection[c]	77053	
Refinement		
Resolution range (Å)	19.96-2.1	
R work (%)[d]	19.2	
R free (%)[e]	25.9	
No. of protein residue	976	
No. of water molecule	507	
B factor		
Whole molecules	28.995	
Side chain (include water)	30.311	
Main chain	27.543	

www.manaraa.com

Ligand (Maltose E/F)	18.109/17.607
Solvent	36.411
RMSD	
Bond length	0.012
Bond angle	1.597
Ramachandran Plot	
Favorable (%)	89.3
Additional (%)	9.1
Generous (%)	1.7
Forbidden (%)	0
PDB accession ID	3OAI

[a]*Rmerge*= ΣhklΣi|I(hkl)i-{I(hkl)|/ ΣhklΣiI(hkl)i , where $I(hkl)_i$ is the intensity of an observation and I(hkl) is the mean value for its unique reflection. Summations cover all reflections.

[b]The values after the hyphens indicate the highest resolution shell.

[c]$N_{obs}/N_{unique.}$

[d]*Rwork*= Σhkl|F(hkl)o- {F(hk|/ΣhklF(hkl)o

[e]R free was calculated same way as R working, but with the reflections excluded from refinement. The R free set was chosen using default parameters in Refmac 5

6.2.2 Crystal Structure of hP0ex

It was predicted and found that hP0ex domain shows significant similarity to the members of the immunoglobulin family (281), specifically, a v-type immunoglobulin fold (282). The structure of rP0ex, a homolog of hP0ex, has also been assigned to the immunoglobulin superfamily (166). Our crystal structure further confirmed this finding (Figure 6.1 and Figure 6.3). The structure consists of nine antiparallel β strands forming a barrel-like shape. However, as with other immunoglobulin superfamily members, the hydrogen bonds do not go around the barrel, the structure is physically a β-sandwich

www.manaraa.com

comprised of two distinct β-plated sheets (283). Strands D, E, B, and A make up one β sheet, while strands D, F, C, C', C'' constitute the other β sheet. In spite of the proximity between strands A and B, the entire length of strand A forms hydrogen bonds with strand G. A similar conformation is seen in molecules CD2 (284) and CD4 with a truncated strand A (285-286). In some other immunoglobulin superfamily members such as CD8 (287) and VCAM-1 (288), the first half of the strand A forms hydrogen bonds with the strand B, while the second half forms hydrogen bonds to strand G. Strands B, C, E, F, and G present a rigid hydrophobic core characteristic of immunoglobulin variable domains. In addition, one conserved disulfide bond is also found bridging strands B and F.

www.manaraa.com

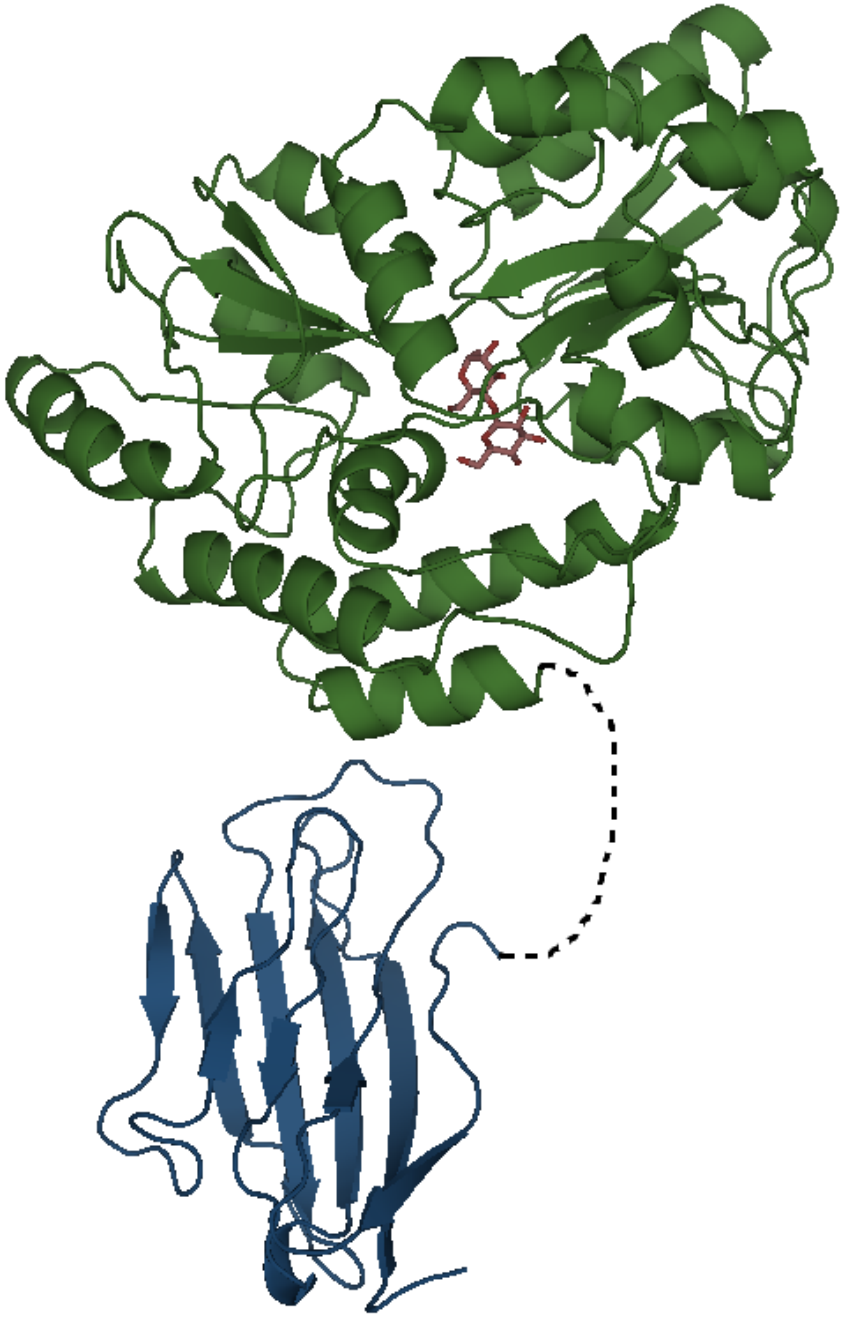


Figure 6.1. Crystal structure of the fusion protein MBP- hP0ex. In green color is the closed-form of MBP with a maltose molecule. In blue color is the hP0ex and the linker connecting MBP and hP0ex is presented by dotted line due to disordered structure.

www.manaraa.com

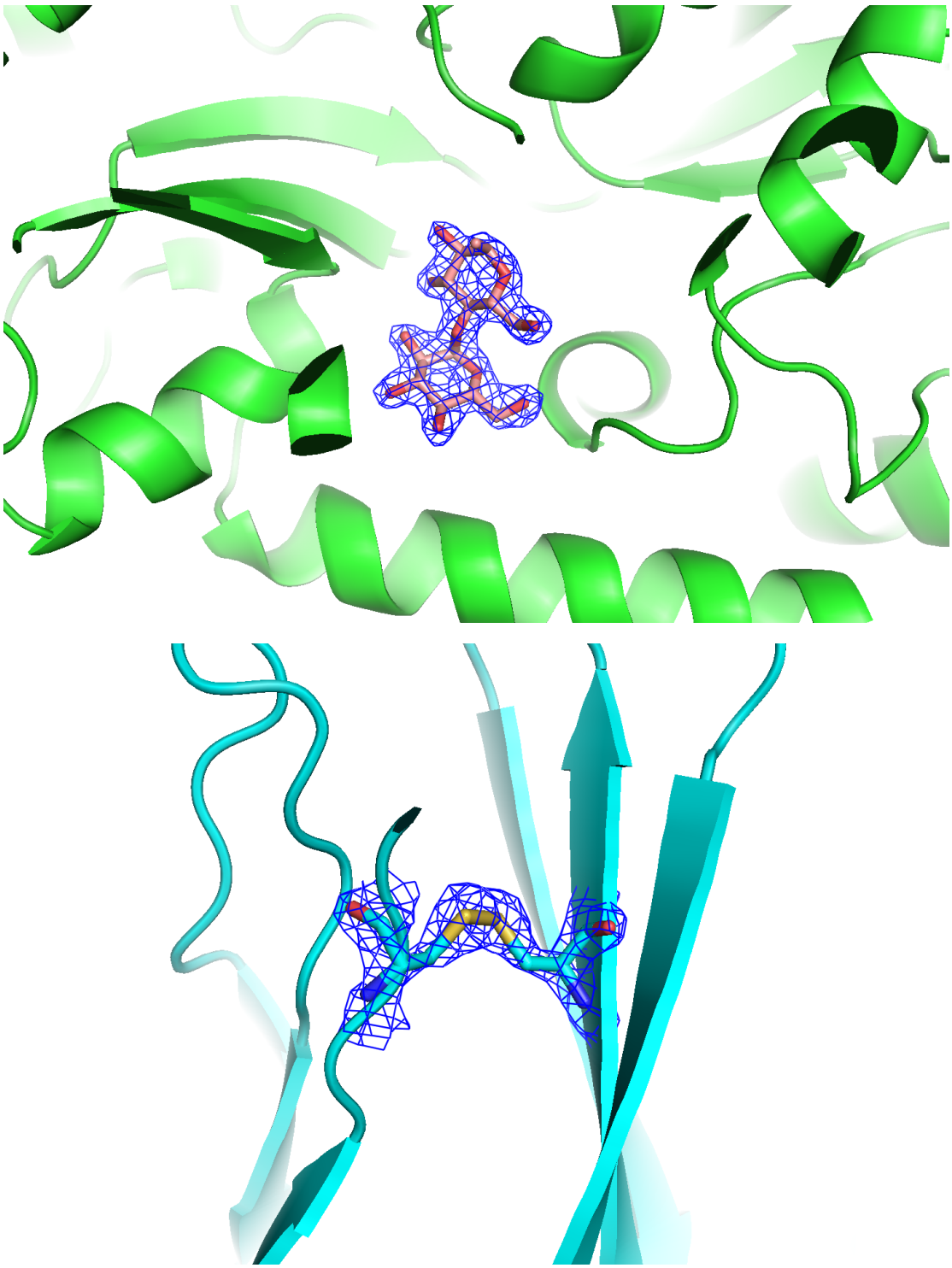


Figure 6.2 The electron density map of maltose and disulfide bond at 2.0 sigma. Upper panel shows the MBP with maltose molecule. The lower panel shows the intra-molecular disulfide bond in hP0ex. The electron density map is rendered by a blue mesh.

www.manaraa.com

6.2.3 The overall structures of hP0ex and rP0ex are similar

Although three residues differ between hP0ex and rP0ex (H10Y, R16Q, and R77S), the overall structures of hP0ex and rP0ex are quite similar (Figure 6.4). These mutations are located on the surface of the hP0ex, making hP0ex more basic compared to rP0ex. The overall more basic characteristic of hP0ex may be important in understanding the differences between human and rat Schwann cell functions. The two crystal structures also differ at the FG loop which is continuous in hP0ex but broken in rP0ex due to four disordered residues (Pro103, Pro104, Asp105, and Ile106, Figure 6.4).

Secondary structure analysis based on the hP0ex and rP0ex structures rendered with the program PyMOL, revealed almost identical secondary structure compositions. Both hP0ex and rP0ex represent a typical v-type immunoglobulin fold (282). The difference between the two structures is that some β strands in the hP0ex are missing. The loss of some secondary structure elements in hP0ex makes it relatively more flexible. For instance, strands A' and B' are not seen and helix 1 (residues 76 to 78) is missing in hP0ex . The absence of helix 1 in hP0ex may be partially due to the change of

www.manaraa.com

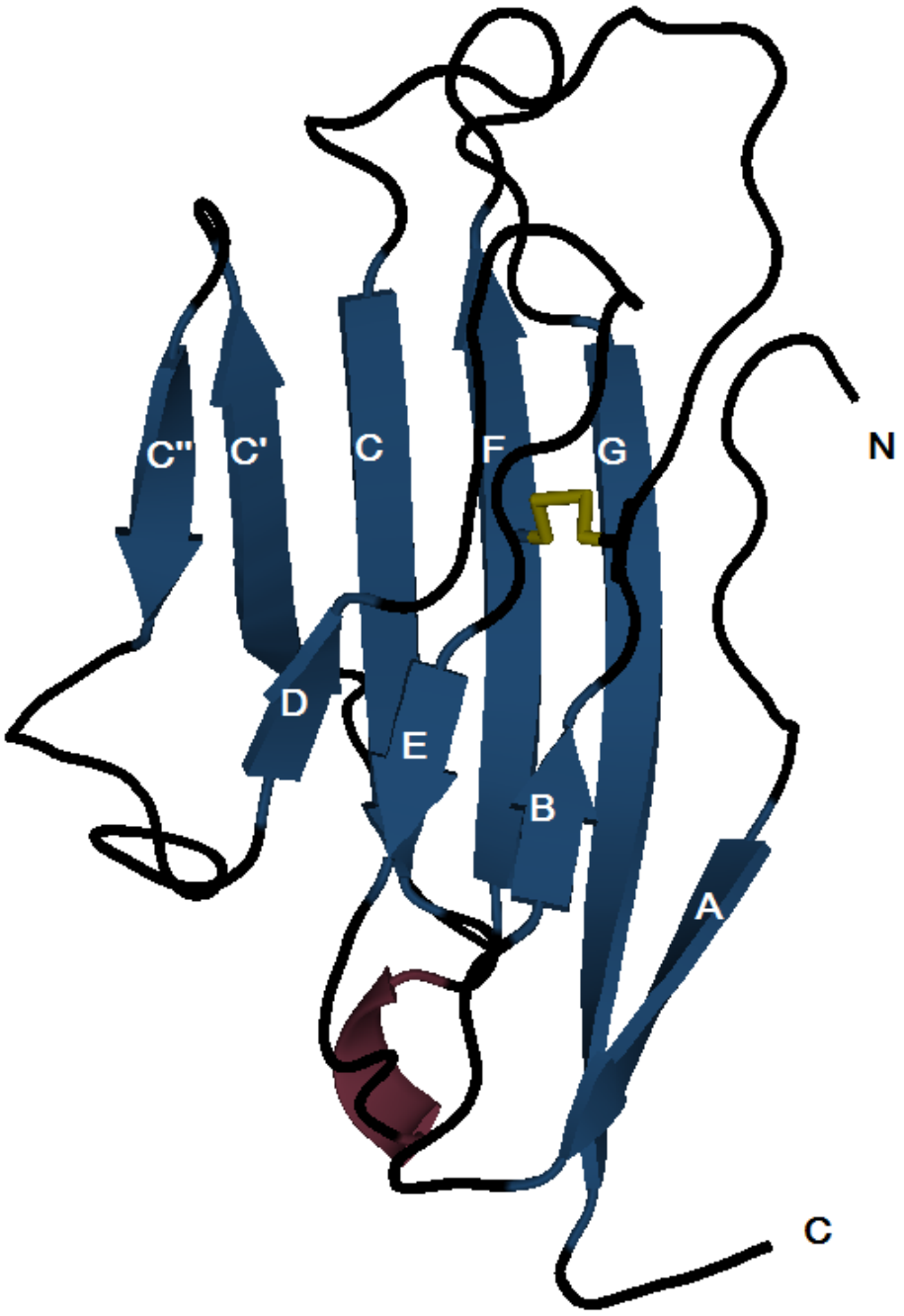


Figure 6.3. hP0ex forms a v-type immunoglobulin fold. The secondary structure elements are highlighted by different colors. β strands are shown in blue and the helices are in brown color and loops are in black. The intra-molecular disulfide bond is shown in dark yellow.

www.manaraa.com

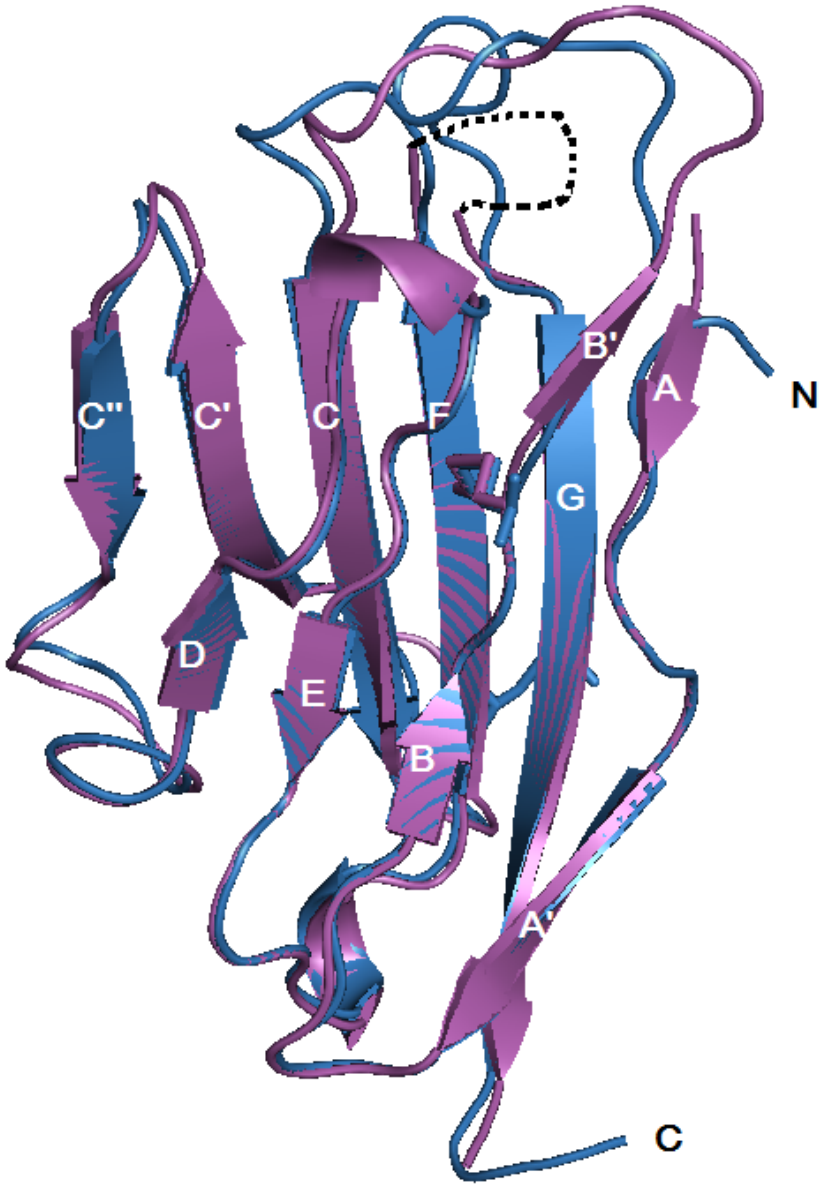


Figure 6.4. Comparison between hP0ex (blue) and rP0ex (purple and black dots). The overall structures are highly conserved between the two species, considering the similarity in the amino acid sequence. However, the hP0ex structure contains the amino acid sequences that are difficult to track due to the broken electron density map for rP0ex.

www.manaraa.com

codon 77 from Arg in hP0ex to Ser in rP0ex. The other two variations between hP0ex and rP0ex, H10Y and R16Q, appear not to alter the secondary structure

6.2.4 Some regions in hP0ex show high flexibility relative to rP0ex

Superposition based on Cα of the hP0ex and rP0ex structure shows high root mean square deviations (RMSD) in some regions. RMSD is more than 1 Å for both main chain and side chain in the following regions: the N-terminus (residues 1 and 2), loop BC (residues from 25 to 32), loop C''D (residues 62, 63 and 64), both ends of the disordered loop in rP0ex (residues 102 and 107). RSMD analysis shows abnormally high RMSD regions (side chain only) in the following: N-terminus (residue 4), Strand A (residues 7, 8 and 10), loop AB (residue14), loop BC (residues 23,24, and 33), loop CC' (residues 45 and 46), strand C' (residues 48 and 49), loop C'C'' (residues 54 and 55), strand C'' (residues57, 58 and 60), loop C''D (residues 61, 65, 67 and 70), loop DE (residues 78, 79 and 83), the broken ends in rP0ex (residues 101 and 109), strand G (residue112), and C-terminus (residue119). Areas surrounding the three variation sites show high RMSD values (Figure 6.5).

www.manaraa.com

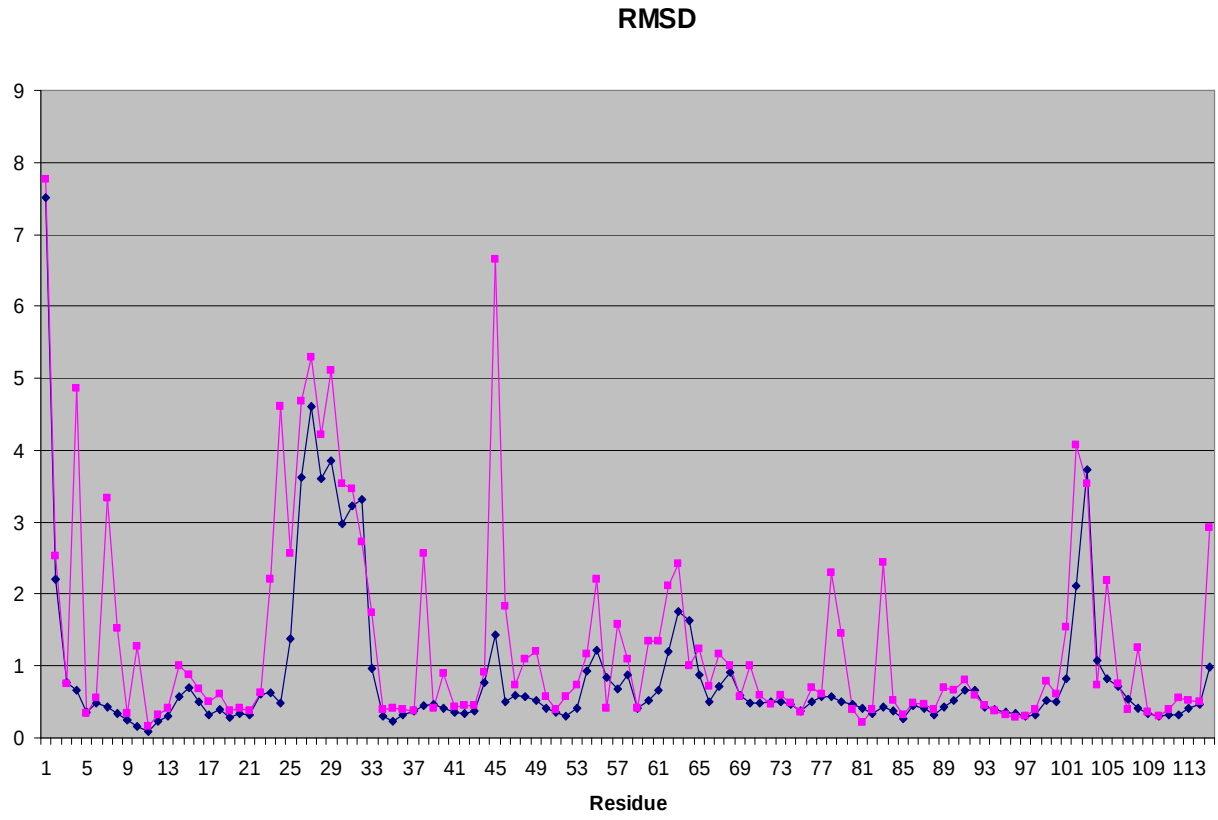


Figure 6.5 The RMSD digram based on the superposition of hP0ex and rP0ex on the C α . The X axis is the residue number and the Y axis the RMSD value. The series 1 (blue) represents the main chain RMSD and the series 2 (purple) is the RMSD of the side chain. The corresponding missing residues in hP0ex from residues 103 to 106 are omitted in the comparison. In addition, to facilitate the comparison, residues after 102 are renumbered.

6.2.5 The ß-strand A in hP0ex is relatively shorter than the matching ß-strand in rP0ex

Although the core of the immunoglobulin fold between hP0ex and rP0ex, consisting of strands B, C, E, F, and G, is highly conserved, the most significant difference among β strands is located in strand A. In hP0ex, strand A is

www.manaraa.com

relatively shorter than that in rP0ex, forming hydrogen bonds only with strand G. In contrast, strand A in rP0ex is divided into two small strands, A and A'. In rP0ex, strand A forms hydrogen bonds with the strand B, and strand A' forms hydrogen bonds with strand G (Figure 6.4), as seen in CD8 (287) and VCAM-1 (288). As a result, the stand A in hP0ex is relatively flexible compared with that in rP0ex. Subsequent mutations on strand A, such as H10P, may further destabilize strand A in hP0ex.

6.2.6 Surface area of hP0ex is larger than the surface area of rP0ex

The surface area in hP0ex (with 552 surface atoms from 115 residues) is increased by 155.9 $Å^2$ compared to that of rP0ex (with 561 surface atoms from 115 residues). The individual amino acid residue surface areas are shown in Figures 6.6 and 6.7. The surface areas of residues 1, 2, 7, 8, 16, 24, 32, 57, 65, 77, 101, and 119 are increased each by at least 20 $Å^2$, while surface areas of residues 10, 62, 64, 97, 102, and 107 are each decreased by more than 20 $Å^2$. The surface area changes at residues 10, 16, and 77 are due to the sequence variation between the human and rat protein. The surface area changes at residues 1, 2, 101, 102, 107, and 119 may be caused by the flexibility of N- and C- termini and the broken ends within rP0ex. The increased extracellular surface area of hP0ex may strengthen the interactions among hP0ex molecules on the surface of Schwann cells, leading to higher affinity between homophilic molecules.

www.manaraa.com

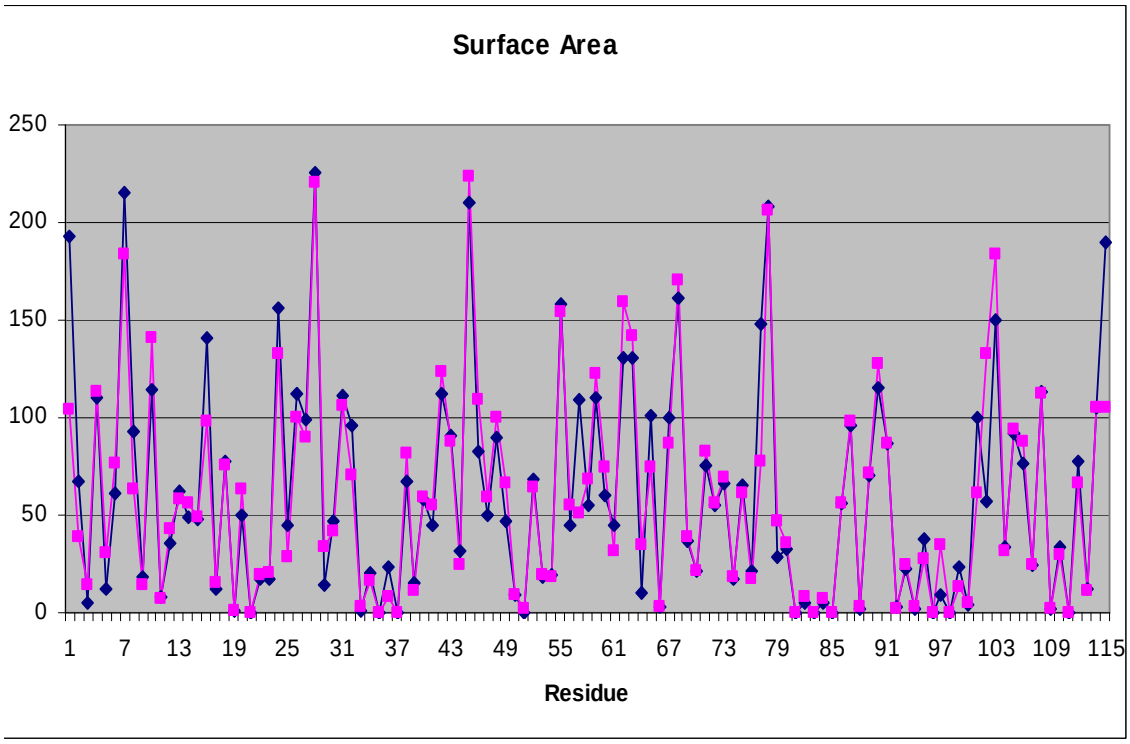


Figure 6.6 The surface area of individual amino acid residues in both hP0ex and rP0ex. The X axis is the amino acid residue number and the Y axis is the surface area of individual amino acid residues in square Angstroms. The blue series represents the surface area of hP0ex residues and the purple series represents the surface area of rP0ex amino acid residues. The overall surface area distribution is similar between hP0ex and rP0ex, however certain amino acid residues show large surface area differences that may affect the overall stability of the molecules and inter-molecule interactions. The corresponding missing residues in hP0ex from residues 103 to 106 are omitted in the comparison. In addition, to facilitate the comparison, residues after 102 are renumbered.

www.manaraa.com

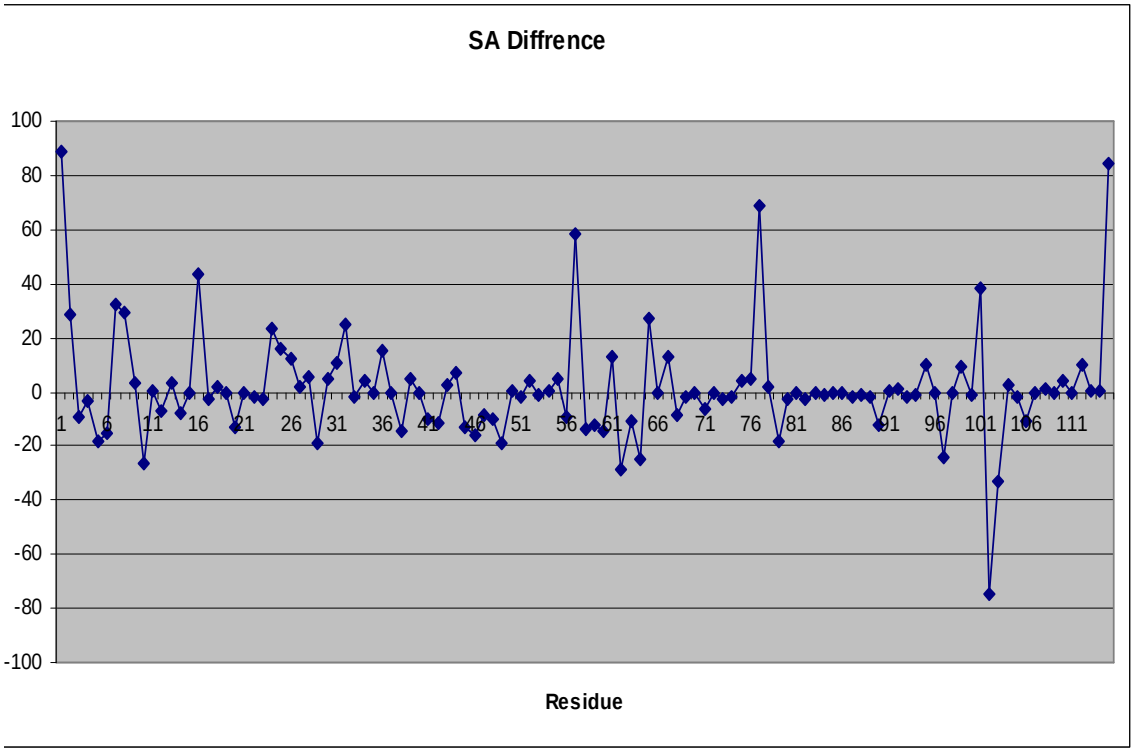


Figure 6.7. The difference of surface area of individual amino acid residues between hP0ex and rP0ex. The X axis is the amino acid residue number and the Y axis is the surface area difference of individual amino acid residues. The overall surface area distribution is simlar between hP0ex and rP0ex, however, certain amino acid residues show large surface area differences that may affect the overall stability of the molecules and inter-molecule interactions. The corresponding missing residues in hP0ex from residues 103 to 106 are omitted in the comparison. In addition, to facilitate the comparison, residues after 102 are renumbered.

6.3 Discussion

www.manaraa.com

6.3.1 hP0ex tetramer assembly based on the rP0ex tetramer

An hP0ex tetramer was assembled to investigate the hP0ex homophilic adhesion capacity based on the rP0ex tetramer conformation. Although the hP0ex tetramer does not exist in the crystals, due to steric hindrance from the MBP chain, the hP0ex tetramer is plausible both from the aspect of proposed hP0ex homophilic adhesion capacity and from the crystal structure of rP0ex showing a tetramer conformation.

The rP0ex tetramer is generated through symmetry operations based on the rat tetramer model (166). Four individual hP0ex molecules are superposed onto the rP0ex tetramer. The four resultant hP0ex molecules from the superposition are combined to give a hypothetical hP0ex tetramer (Figure 6.8). The hP0ex and rP0ex tetramers are highly conserved, which is essentially determined by the similarity between hP0ex and rP0ex monomer molecules.

However, there are three key differences between the human and rat tetramer model. First, in the hP0ex tetramers, the interaction between loop BC of one molecule and loop C''D and EF of the next molecule in the four-fold interface as defined by Shapiro *et al*.(166) is stronger than in the rP0ex tetramers. In the rP0ex tetramer, a stacking interaction is formed between Trp 24 in loop BC (of one monomer) and His 86 in loop EF (of the adjacent rP0ex monomer) due to the relatively long distance. However, in the modeled hP0ex tetramer the side chains of Trp 24 in loop BC (of one monomer) and His 86 in loop EF (of the adjacent monomer) are too close to one another and they

www.manaraa.com

appear to clash. Second, the N-terminus of hP0ex is more flexible and forms a loop, while the N-terminus of rP0ex is more organized in the form of a β strand. As a result, the hP0ex N-terminus of one molecule is bent toward the loop EF of the next hP0ex, forming two extra hydrogen bonds between Ile 1 and Asn 87. In contrast, the N-terminus of rP0ex forms hydrogen bonds with strand B within the same molecule. Third, the C-termini of hP0ex are thought to form anchors by which the hP0ex tetramers are attached to the Schwann cell membrane. It is anticipated that the C-termini point away from the tetramer structure and toward the membrane. However, the structure we solved does not support the above model. The four C-termini from the four monomers curve back into the donut-like hole within the tetramer. This curving of the C-terminus may be explained by the lack of membrane interaction and molecular flexibility.

6.3.2 Five clinically important mutations of hP0ex

Five clinically important mutations on hP0ex are under investigation, specifically, H10P, R69C, R69H, T95M, and H52R. These mutations are dispersed throughout the entire hP0ex molecule, as shown in Figure 6.9. Mutation H10P is located on strand A and mutations R69C and R69H are located on loop C''D. Mutation T95M is on strand F and mutation H52R is on stand C'. All these mutations lead to early or late onset neuro-degenerative disease.

www.manaraa.com

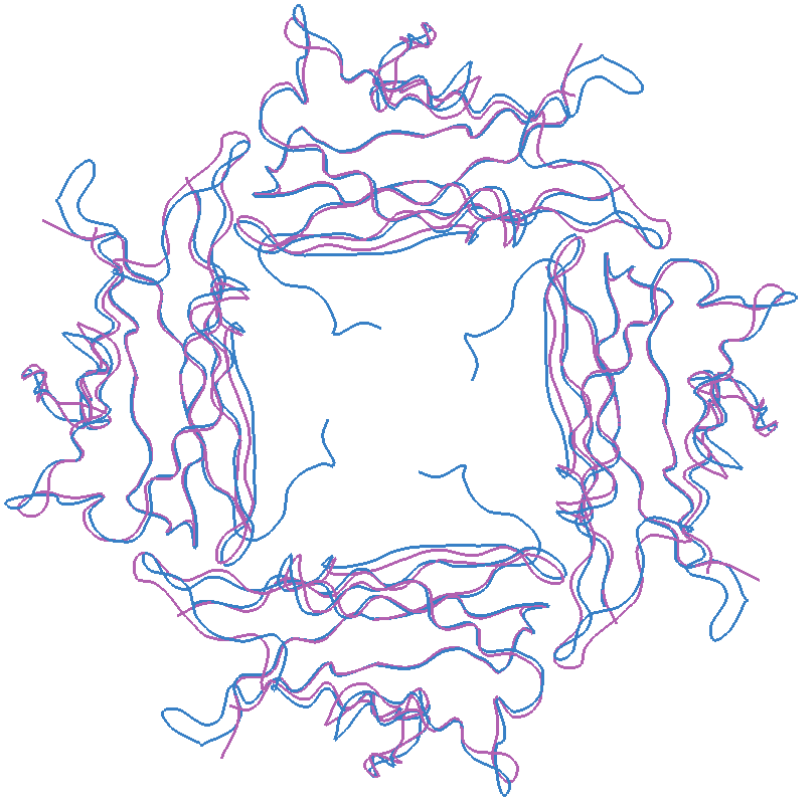


Figure 6.8. Superposition of hP0ex tetramer (blue) and rP0ex tetramer (purple). A tetramer does not exist in the hP0ex crystal structure. However, a tetramer has been identified as the “biological structure” in the rP0ex cystal. The same crystallographic symmetry operation is performed to generate the tetramer structure of hP0ex. The two tetramer structures of hP0ex and rP0ex then are superposed based on the Cα atoms. As shown in this figure, the two tetramer structures overlay each other well. However, further investigation reveals that some steric conflicts exist among the hP0ex monomers on the inter-molecule surface.

www.manaraa.com

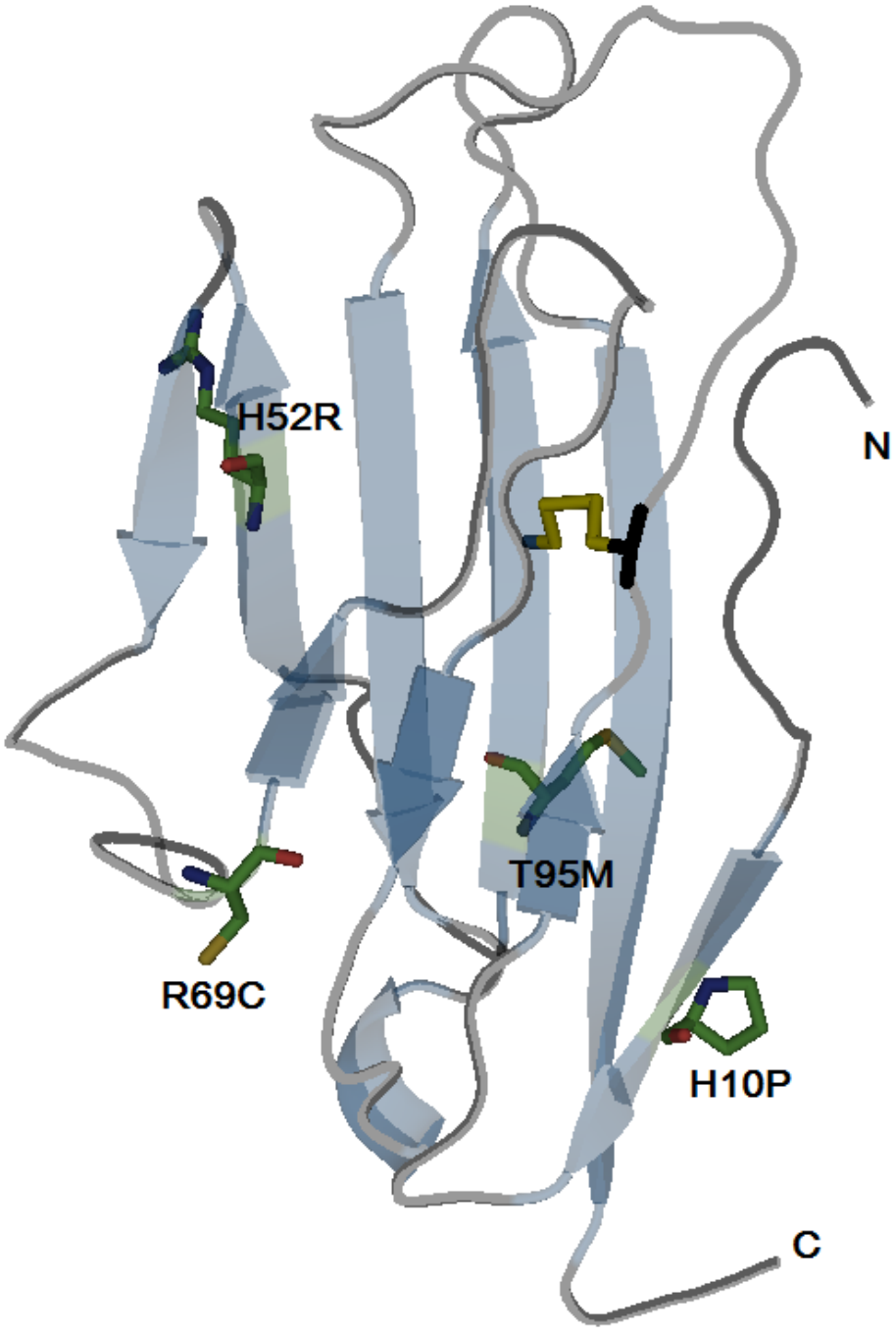


Figure 6.9 hP0ex monomer with H10P, R69C, T95M, and H52R mutations.

The clinically important mutations are shown by a stick model. The intra-disulfide bond is shown in dark yellow color.

www.manaraa.com

The five clinically important mutations are highlighted in the hP0ex tetramer to investigate the structural influences imposed by these mutations on hP0ex function (Figure. 6.10). While mutation H10P is not on any inter-molecular interface within or between the hP0ex tetramers, it might disrupt the hP0ex monomer structure. In the H10P mutant, no obvious interruption in the tetramer structure is predicted. However, as the modeling prediction of the H10P structure indicates, the introduction of mutation H10P may disrupt the intra-molecular disulfide bond and thus lead to structural instability of the hP0ex monomers (unpublished results). The disruption of the hP0ex monomer structure, in turn, may affect the tetramer formation and its function. A discussion of the other four clinically important mutations follows in the next paragraph.

Mutations R69C and R69H may reduce the homophilic adhesion capacity of hP0ex through disruption or weakening of hP0ex tetramers. These mutations are located on loop C''D, which is an important inter-molecular interface for tethering four hP0ex monomers together in the hP0ex tetramer. The mutation from Arg to either Cys or His shortens the side chain of the residue 69 and may result in weakened interaction between the loop C''D from one monomer and the loop BC from the adjacent monomer. Mutations T95M and H52R are on the outer surface of the hP0ex tetramer and might influence the hP0ex function by disrupting inter-tetramer interactions.

www.manaraa.com

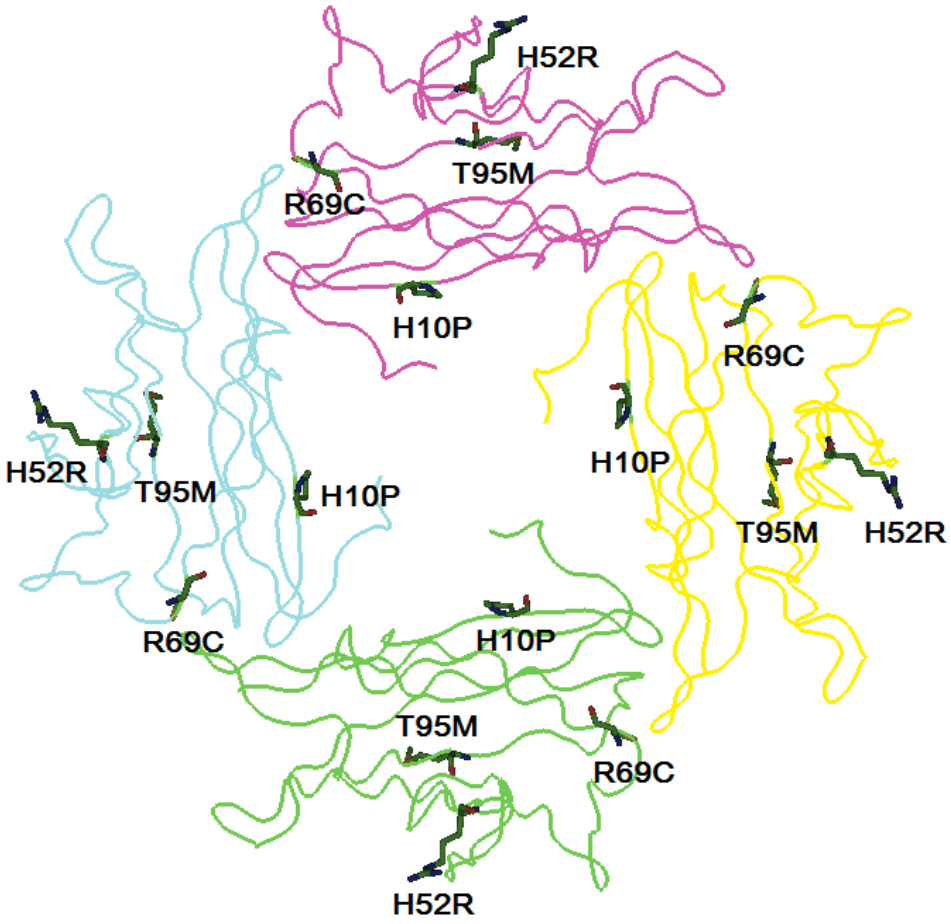


Figure 6.10. Important mutations shown in the hP0ex tetramer. The ribbons in different color are individual hP0ex monomers. Mutations are shown by stick model. The anaylsis of the structural influences introduced by the mutations is limited if the analysis is confined to the hP0ex monomers. Visualization of the mutations in the hP0ex tetramer takes into account the possible intra-monomer, inter-monomer, and inter-tertramer structural disruptions induced by these mutations, thus more comprehensive in the analysis.

www.manaraa.com

6.3.3 Prediction of the structure of H10P hP0ex

The histidine residue at position 10 of the human MPZ coding sequence is one of three amino acids that differ between the human and rat MPZ extracellular domains, and is one of 8 amino acids located within a ß-strand in which the root mean square deviation between the human and rat proteins is greater than 1 Å. Mutation of this residue to proline (H10P) causes a late onset form of CMT1B in several large families (155). Nerve conduction velocities in affected individuals are preserved, suggesting that myelin can functional normally. The peripheral nerve pathology associated with the H10P mutation, however, demonstrates accumulation of proteinaceous material between Schwann cells and axons, as well as focal areas of tomaculous-like myelin disruption (functional disturbance or pathological change), suggesting abnormal myelin breakdown and/or turnover.

In order to investigate a structural contribution to the H10P phenotype, we have used the crystal structure of the hMPZex to predict the hMPZex H10P structure by performing molecular modeling. The result of the simulation demonstrates several important changes in the MPZ structure caused by the H10P mutation. Because of the proline at residue 10, the N-terminus now forms an A' ß-strand similar to that in the rat MPZ which gives rise to an anti-parallel ß-sheet with the downstream segment of the B strand pulling the A' strand toward the N-terminus and away from the F strand. As a

www.manaraa.com

consequence, the distance between Cys21 and Cys98, located on the B and F strands, is increased to 3.7 Å, thus the disulfide bond between these two residues can no longer form. Absence of the disulfide bond between the two ß-sheets, a highly conserved element of Ig domain structure, probably causes the H10P mutant to be less stable than its wild type counterpart. In addition, the two cysteine residues could interact with free cysteines from other proteins, leading to the formation of disulfide-linked MPZ complexes. The substitution of proline at position 10 also directly alters the protein backbone by producing a “kink” at that location. A comparison of the wild type and H10P mutant forms of hMPZex is shown in Figure 6.11.

www.manaraa.com

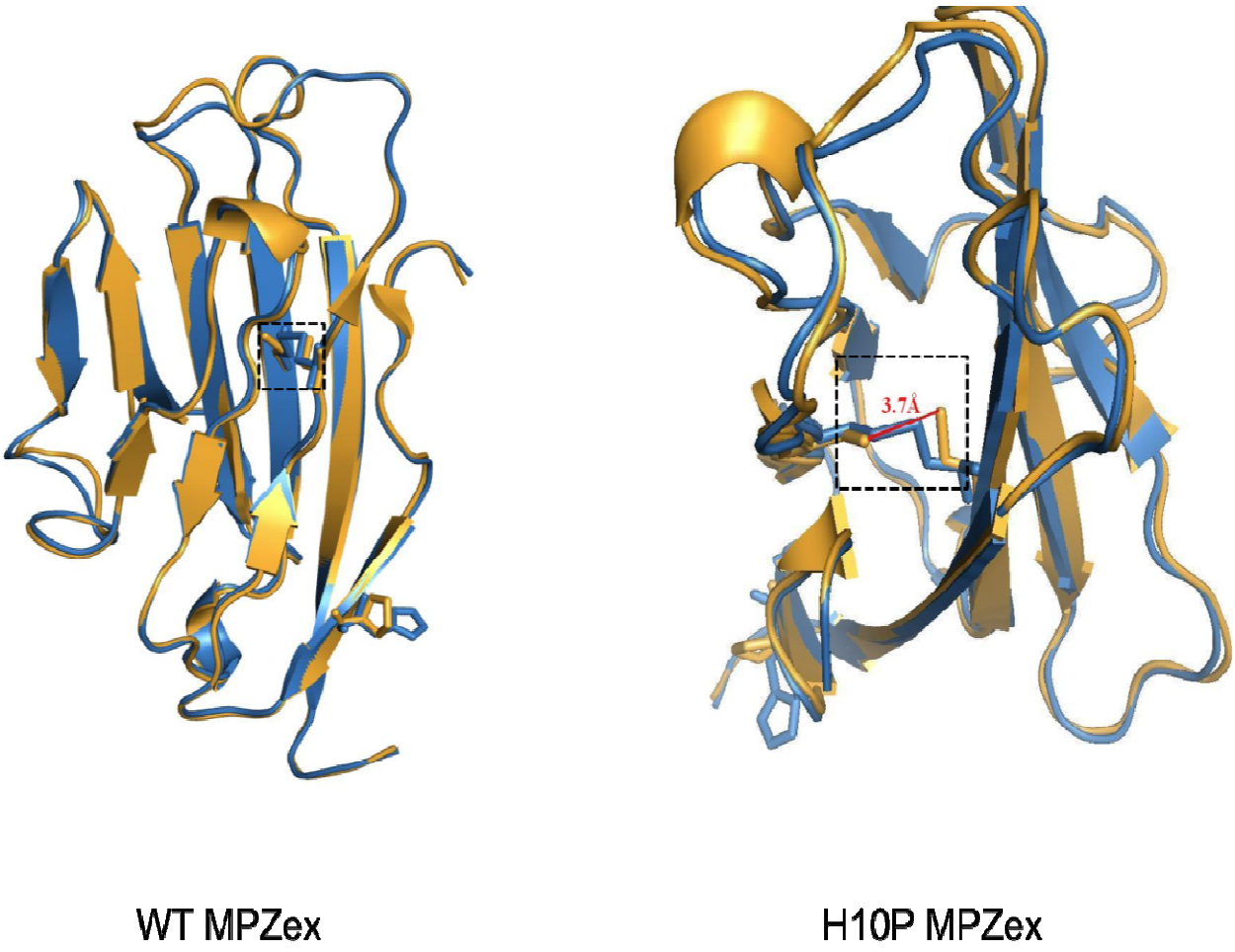


Figure 6.11. Structural comparison of the wild type and H10P mutant forms of MPZex. The left panel shows the superposition of the wild type (sky blue) and predicted H10P-MPZex (gold). The right panel more closely focuses on the disulfide bond region, showing the broken disulfide bond in H10P-MPZex. The overall structure of wild type MPZex and the H10P-MPZex are similar. However, the broken disulfur bond indicates the possibility of inter-meluclar disulfur bonds, which may be the reason of oligomerization of the H10P-MPZex.

www.manaraa.com

6.3.4 Crystallization of proteins fused to the maltose binding protein

According to a recent report from the Joint Center for Structural Genomics (http://www.jcsg.org), among all the proteins that are expressed, only 9.8% could be crystallized with favorable size and quality for X-ray screening. 74% of these crystals diffract X-ray for further data collection. Furthermore, only 50% of these crystals with diffraction data can be solved. As a result, the most serious obstacle for crystallographers is to grow high quality protein crystals.

Several different methods were developed recently to facilitate the crystallization of proteins that were previous refractory to crystallization (289). These methods include optimization of target constructs, noncovalent crystallization chaperones, removal of post-translational modifications, stabilization of protein targets, surface-entropy reduction (SER), improvement of crystal quality, and the use of fusion proteins for crystallization.

Different carrier fusion proteins have be used, a tag protein fused with the target protein, to facilitate protein crystallization, such as maltose binding protein (MBP), thioredoxin (TRX) (290-291), glutathione-S-transferase (GST) (292-295) and lysozyme (296-299). Usually the presence of the tag proteion helps increasing both the expression and solubility of the target protein (300-302), in addition to the simplification of protocol of target protein purification if the tag proteins are affinity tags (303-304).

Among all the potential tag proteins, the protein with highest success rate

www.manaraa.com

is the maltose binding protein. It is usually added on the N-terminus of the target protein and can act as a chaperone-like protein to facilitate the proper folding of the targets (305-306). The protein crystals obtained through the fusion of MBP are listed in Table 6.2.

However, it is not uncommon that the target proteins are still refractory to crystallization despite of the presence of MBP to increase the expression and solubility (307-310). In these cases, we may introduce some mutations on MBP itself in order to achieve the goal of maximal surface entropy reductions (SER) (311-313).

However, the addition of the MBP tag to the target proteins may itself cause some undesirable outcomes. The MBP itself is more than 300 amino acid residues, thus limit the size of target proteins that are suitable for *E. c*oli expression system. The presence of MBP on the N-terminus of the target protein may disrupt the native structure of the target protein, especially when the N-terminus of the target protein is functionally important and is involved in certain structural interactions. In addition, the linker region connecting the MBP and the target proteins may introduce significant structural heterogeneity, compromising the quality of the crystals. To partially solve this problem, the most effective method is to minimize the length of the linker region.

www.manaraa.com

Table 6.2. Crystal structures with maltose binding protein as the large affinity tags. Addapted from Moon *et. al.(314)*

Target protein	Size (AA)	Resolution (A °)	Linker sequence
HBS antigen(315)	14	2.7-2.9	GS
IAP(316)	22	1.75-1.86	AAA
MutS (C-term)(317)	34	2.0	HM
MATa(187)	50	2.1-2.3	AAAAA
Huntingtin (N-term)(318)	66	3.5-3.7	AAA
TIM40/MIA40(319)	82	3.0	SSSVPGRGSIEGRPEF
gp21(186)	88	2.5	AAA
Monobody YSX1(320)	91	2.0	GSSGSS
Monobody YS1(320)	93	1.8	GSS
IPS1 CARD domain(321)	93	2.1	SAMA
CRFR1 ECD(322)	96	1.96-3.4	AAAEF
ZP3 (ZP-N domain)(323)	102	2.3-3.1	AAA
L30(324)	103	2.31	SSSVPGRGSIEGRA
gp21(325)	108	-	AAA
designed helical protein(326)	108	1.9	SSSNNNNNNNNNN
SarR(185)	115	2.3	AAAEF

www.manaraa.com

Argonaute2 PAZ (327)	136	2.8	AAAEF
PTH1R ECD(328)	174	1.95	AAAEF
CD38(329)	256	2.4	N/A[b]
2OST(330)	298	2.65	AAA
RACK1A(331)	324	2.4	AAA
NEDD8(332)	431	2.8	AAA
hP0ex[c]	121	2.1	AAA

[a] All target proteins listed in this table are fused C-terminal to the carrier protein, unless otherwise stated.

[b] Information not available.

[c] The structure discussed in this dissertation.

6.3.5 Amyloidosis and CMT1B

Crystallization attempts, including the in-house and high-throughput screens, have been performed to obtain diffration quality crystals of mutants of hP0ex, such as H10P and R69H. However, no crystals were obtained despite the thorough screening. The fact that mutants are refractory to crystallization may indicate that a single mutation introduces a major structural disturbance into the fusion protein. The target hP0ex mutants may be disoriented in the crystal lattice established by the malose binding protein. Another possibility is that the hP0ex mutants form amyloids depite the presence of maltose binding protein. If the second hypothesis is correct, the protein sample is expected to show characteristic tinctorial properties of green birefringence under polarized

www.manaraa.com

light after it is stained with the dye Congo red. Futher experiments might include expression and purification of hP0ex mutant without maltose binding protein to enhance the possibility of amyloid formation compared to the protein fused with maltose binding protein.

Theoretically, it is possible that amyloid formation is deferred or even blocked if a certain small molecule binds to the altered structure of hP0ex without introducing any major structural or functional changes, *In silico* screening is currently performed to identify small molecules with high binding affinity to the hP0ex mutants. The identified molecules could be evaluated for efficacy in a transgenic mouse model with a MPZ knockout. These molecules are potential drug candidates to alleviate the signs and symptoms, and/or blocking the progression of CMT1B disease.

www.manaraa.com

CHAPTER 7: SUMMARY AND FUTURE DIRECTIONS

7.1 Summary

7.1.1 Summary of the MDR HIV-1 protease project

- No significant conformational changes are observed at the point of additional mutation in each of the corresponding MDR 769 HIV-1 protease variants.
- The proline 81 switch in the MDR769 I10V mutant HIV-1 protease leads to signifant conformational changes of the active site cavity.
- The Phe82 side chain in the MDR769 A82F mutant reveals the flipped-out conformation.
- The hydroxyl group on the side chain of residue 82 points away from the active site cavity in A82S while it is pointing towards residues Leu23.
- The wild type and A82F mutant protease dimers display a larger surface area at the dimer interface.
- Molecular dynamics simulations show two-fold and three-fold opening of the flaps in the MDR and wild type HIV-1 proteases.
- The substrate (Gag-Capsid) is less stable in the expanded active site cavity of the MDR769HIV-1 protease.
- The conformationa of the substrate (Gag-Capsid) is changed during the MD simulations.
- The substrate cleavage site in MDR protease-substrate complex fails to

www.manaraa.com

align between the catalytic aspartic acid residues after the MD simulations.

- No significant differences are noticed for the IC50 values of the reduced CA/p2 peptide library against MDR and WT HIV-1 protease.
- Reduced peptide P1'F shows IC50 value against MDR HIV-1 protease comparable to that of FDA approved HIV-1 protease inhibitors.
- All six FDA approved inhibitors (amprenavir, indinavir, lopinaivr, ritonavir, nelfinavir and saquinavir) are resisted by MDR HIV-1 protease.
- Lopinavir is the least resisted HIV-1 protease inhibitor among the six inhibitors.

7.1.2 Summary of the human myelin protein zero project

- The overall crystal structures of hP0ex and rP0ex are similar.
- Certain regions in hP0ex show high flexibility relative to rP0ex.
- The surface area of hP0ex is larger than the surface area of rP0ex.
- Molecular dynamics simulation of mutant H10P indicates the interrupted intra-molecular disulfide bonds.
- The presence of the maltose binding protein facilitates the crystallization of hP0ex.

7.2 Future directions

7.2.1 Future directions of the HIV-1 protease project

- Measure the Km and kcat of MDR769 HIV-1 protease with nine different

www.manaraa.com

natural substrate peptides.

- Chemically synthesize more potential HIV-1 protease inhibitors to overcome the drug resistance issue.
- Test the inhibitory activity of the synthesized drug candidates against both the WT and MDR HIV-1 protease.
- Perform *in silico* screening to identify potent drug candidates against MDR769 HIV-1 protease.
- Crystallize other three strains of MDR HIV-1 proteases in the lab.
- Cocrystallize the nine natural substrate peptides to reveal the substrate envelope of the MDR769 HIV-1 protease.
- Cocrystallize the FDA approved HIV-1 protease inhibitors with the MDR769 HIV-1 protease to reveal the structural basis of drug resistance against the FDA approved inhibitors.
- Based on the crystal structure of the MDR769 HIV-1 protease inhibitor complexes, further molecular dynamics simulation is performed to reveal the mechanism of drug resistance.

7.2.2 Future directions of the human myelin protein zero project

- Perform *in silico* molecular dynamical simulations of mutants based on the extracellular domain of WT human myelin protein zero to reveal the molecular mechanism of Charcot-Marie-Tooth disease 1B.
- Run *in silico* screening and docking to look for small molecules with high binding affinities to the extracellular domain of WT human myelin

www.manaraa.com

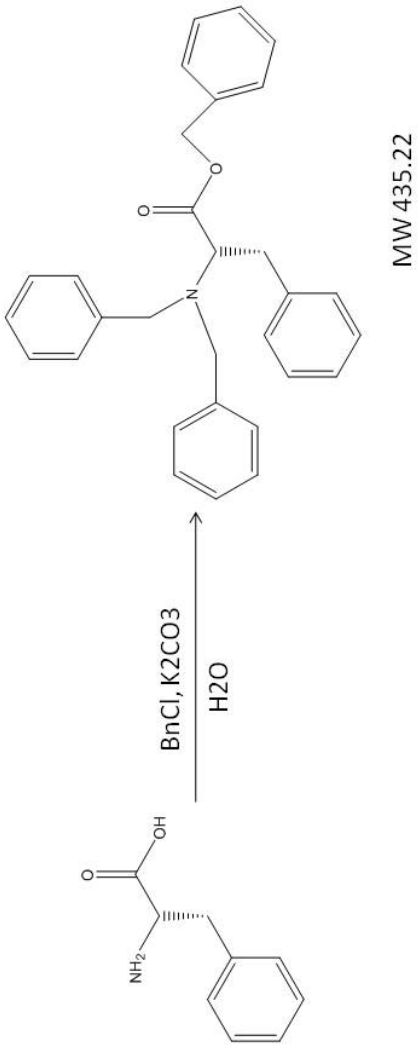
protein zero. The binding of these small molecules may alter the conformation of MPZ and lead to alleviation of the signs and symptoms of Charcot-Marie-Tooth disease 1B.

- Continue crystallization trials of the extracellular domain of the human myelin protein zero with mutations at H10P and R69H.
- Stain the human myelin protein zero with mutations H10P and R69H using Congo red to examine the presence of amyloid that shows characteristic green birefringence under polarized light.
- Perform the X-ray scattering diffraction studies on human myelin protein zero with mutations H10P and R69H to look for the signagure diffraction pattern of amyloid at wavelength 4.7 Å and 10 Å.

www.manaraa.com

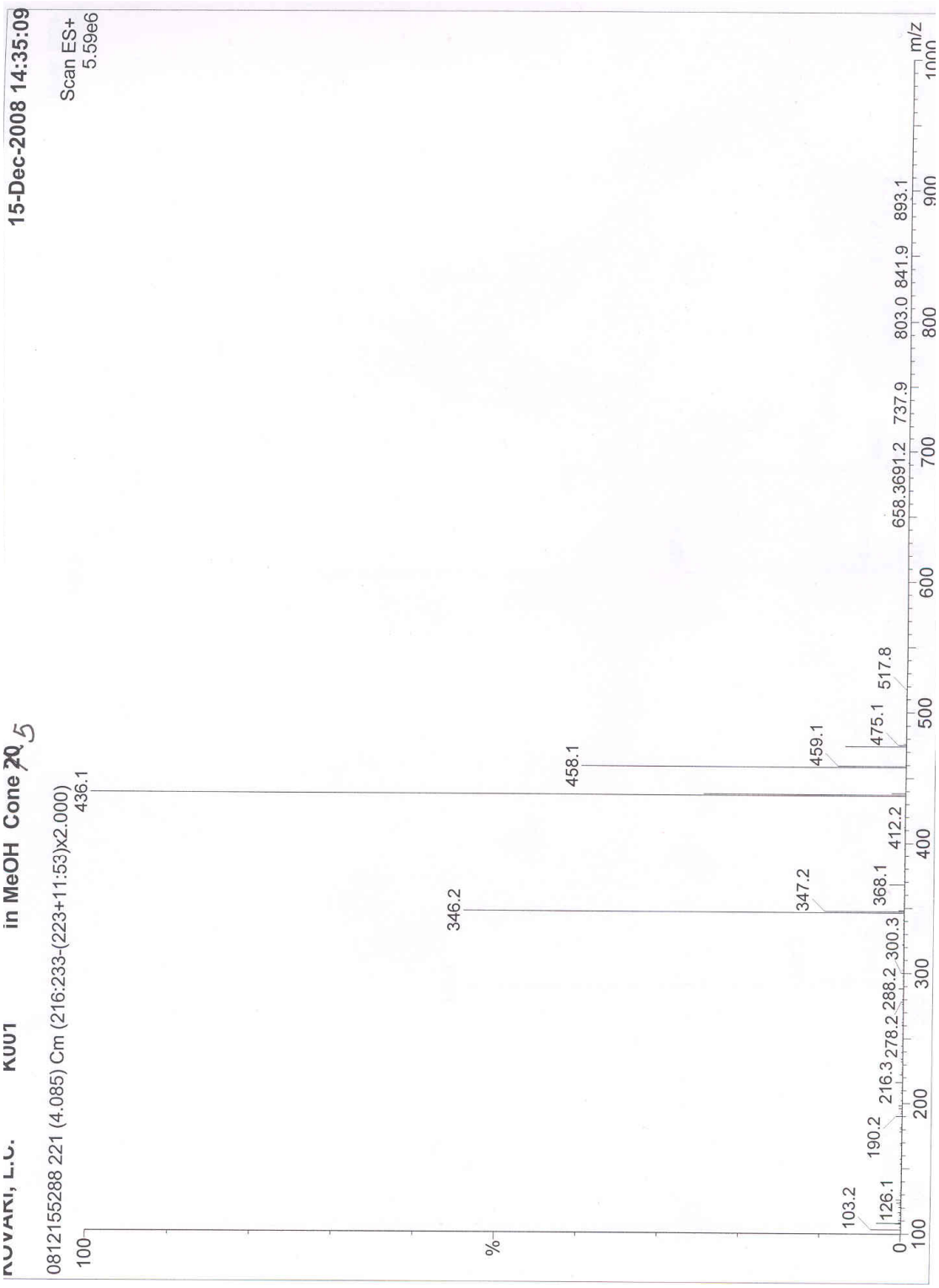
APPENDIX

Reaction 1



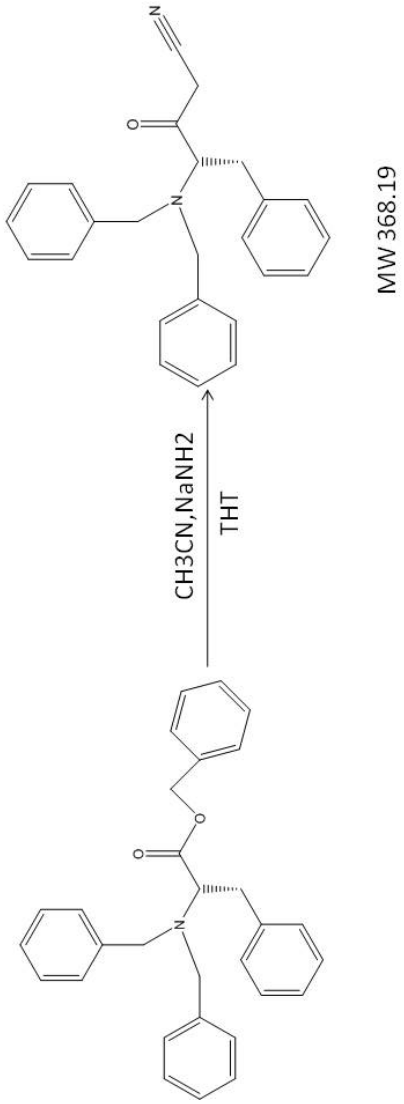
www.manaraa.com

Mass spectrometry result



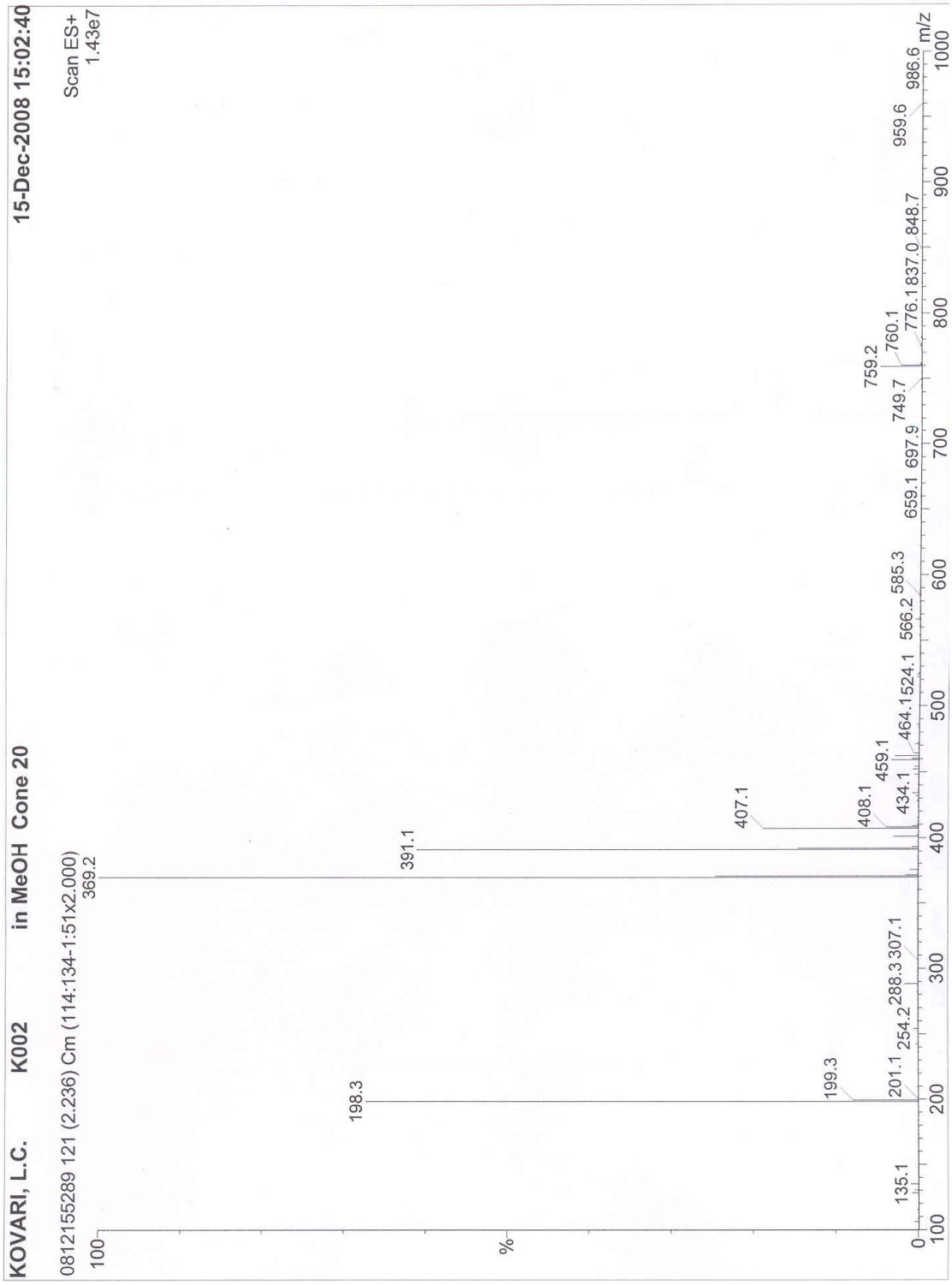
www.manaraa.com

Reaction 2



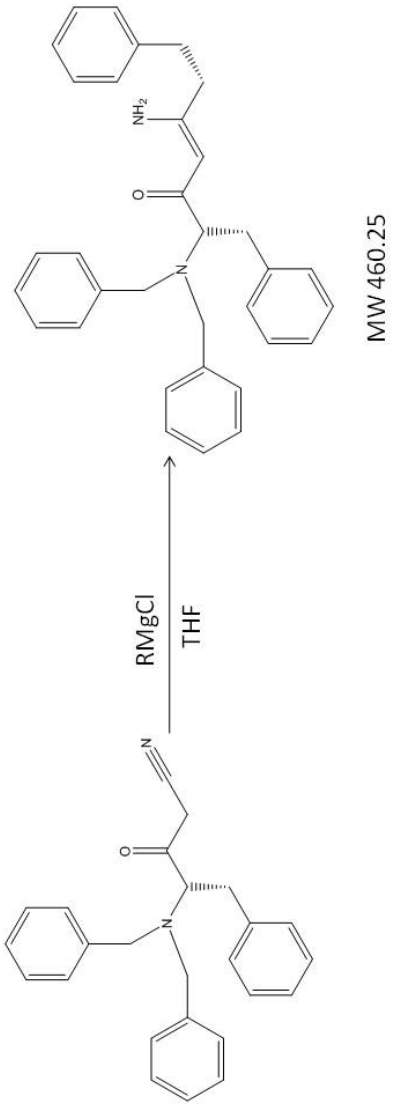
www.manaraa.com

Mass spectrometry result



www.manaraa.com

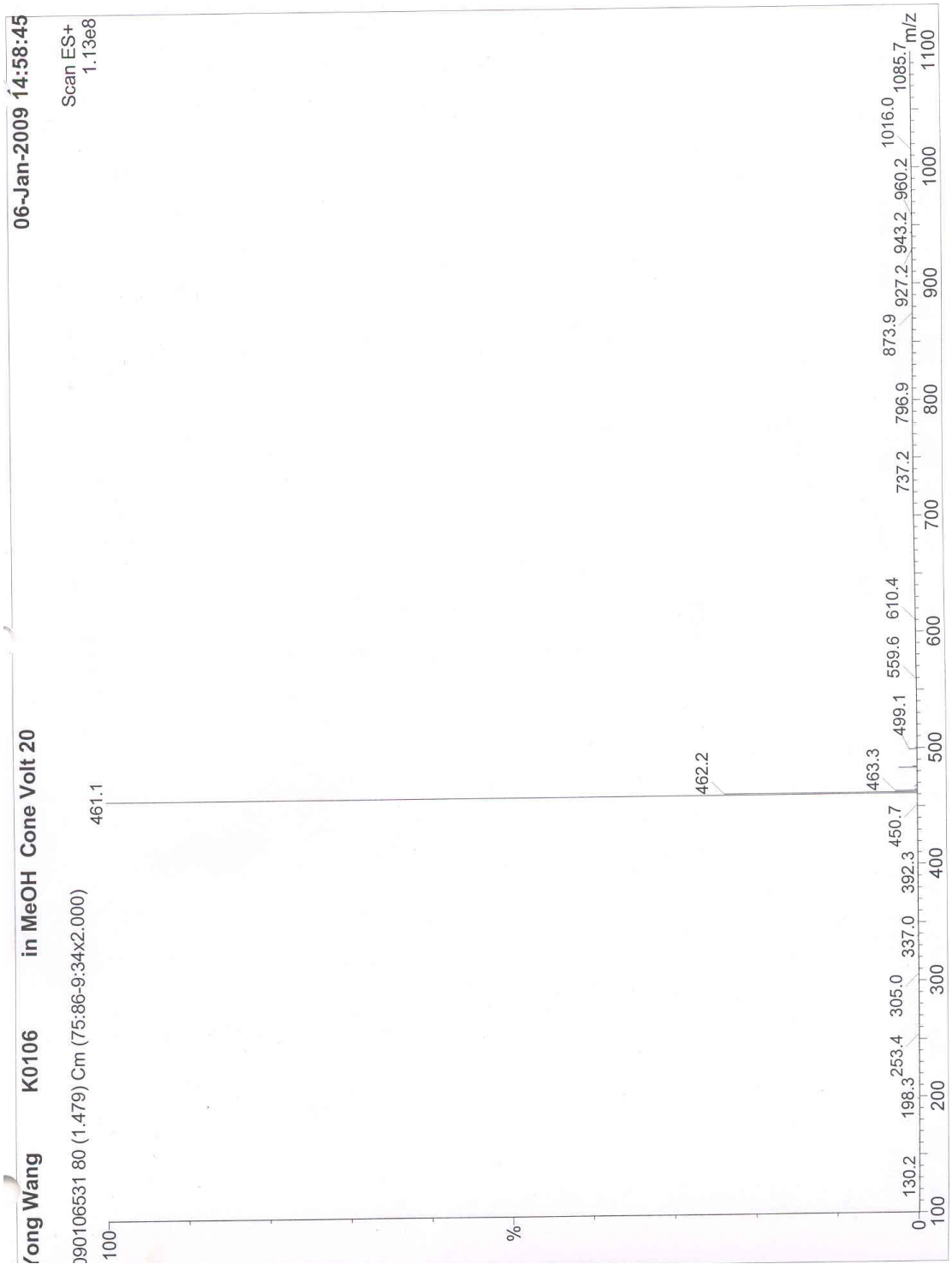
Reaction 3



MW 460.25

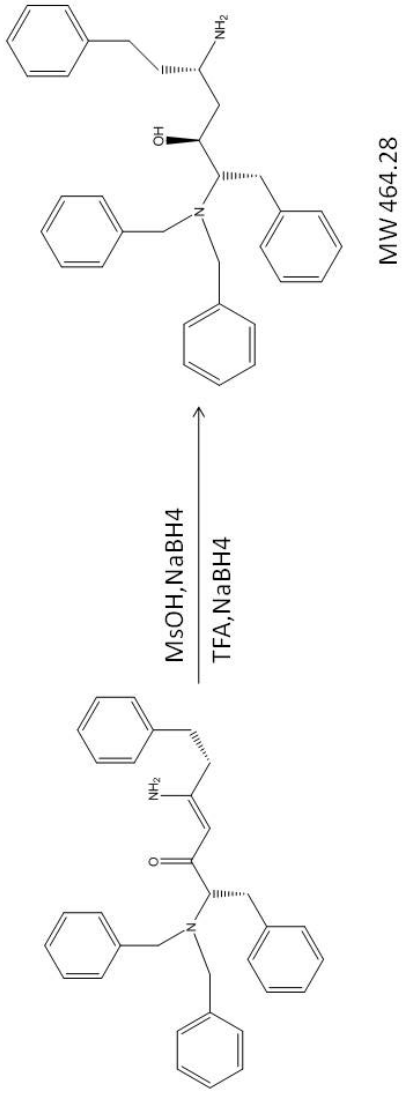
www.manaraa.com

Mass spectrometry results



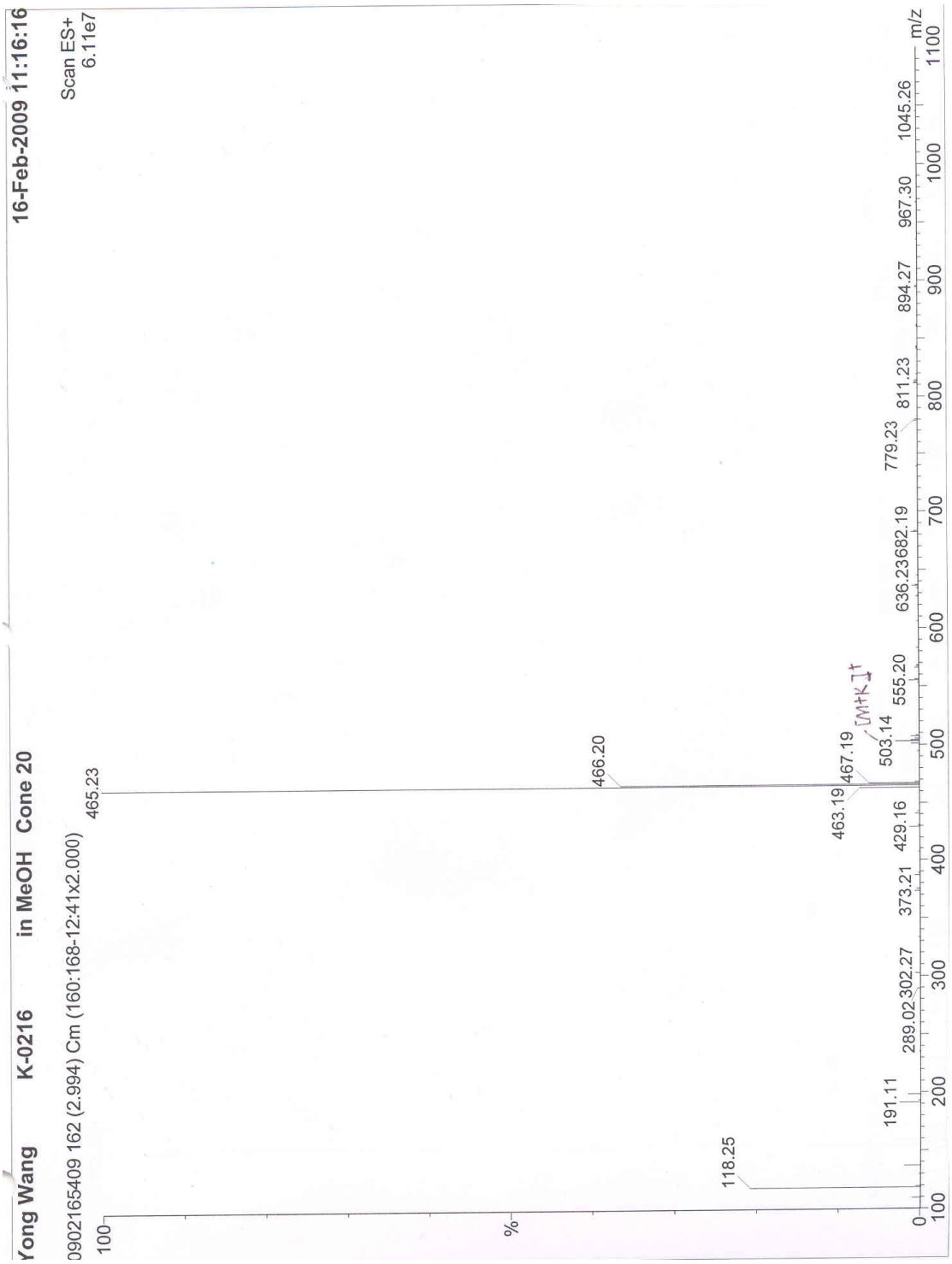
www.manaraa.com

Reaction 4



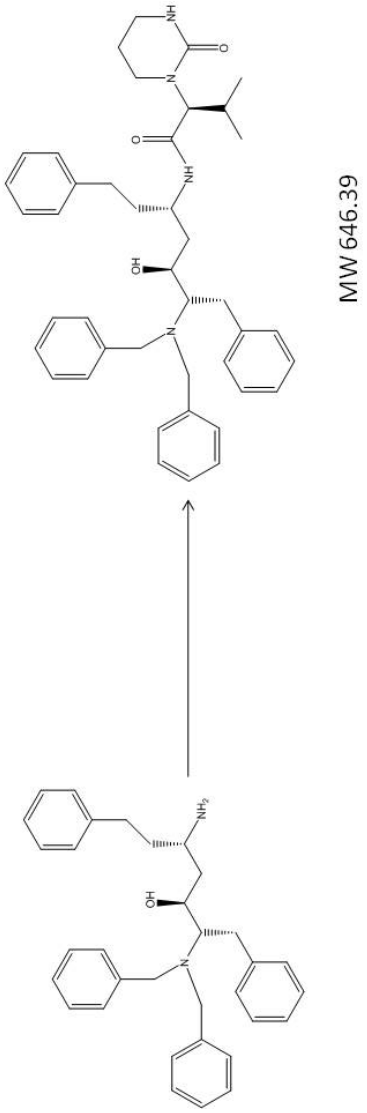
www.manaraa.com

Mass spectrometry results



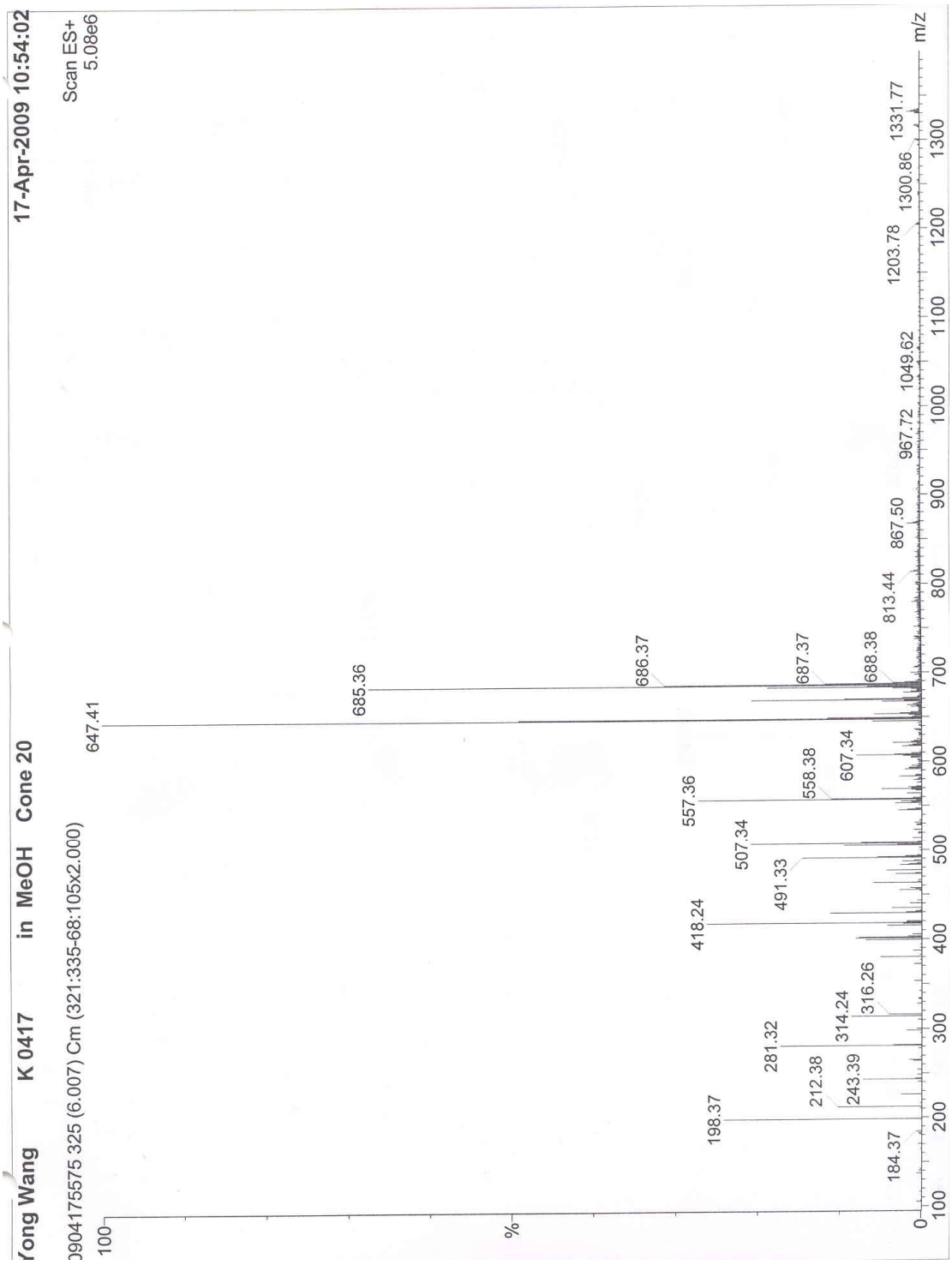
www.manaraa.com

Reaction 5



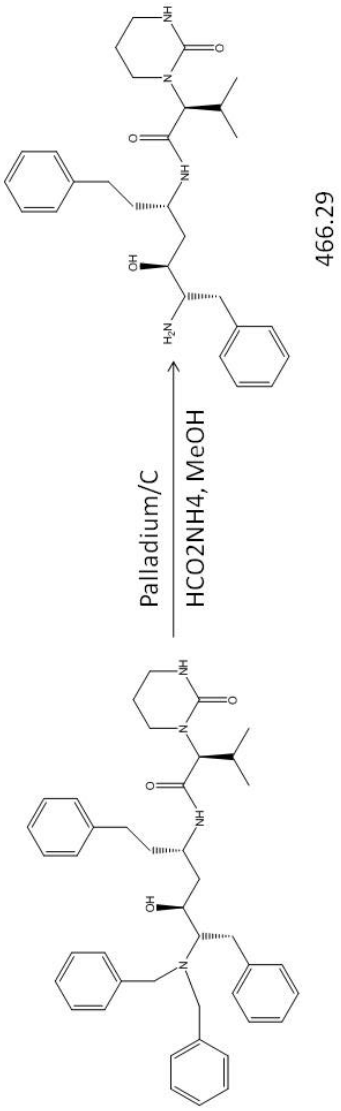
www.manaraa.com

Mass spectrometry results



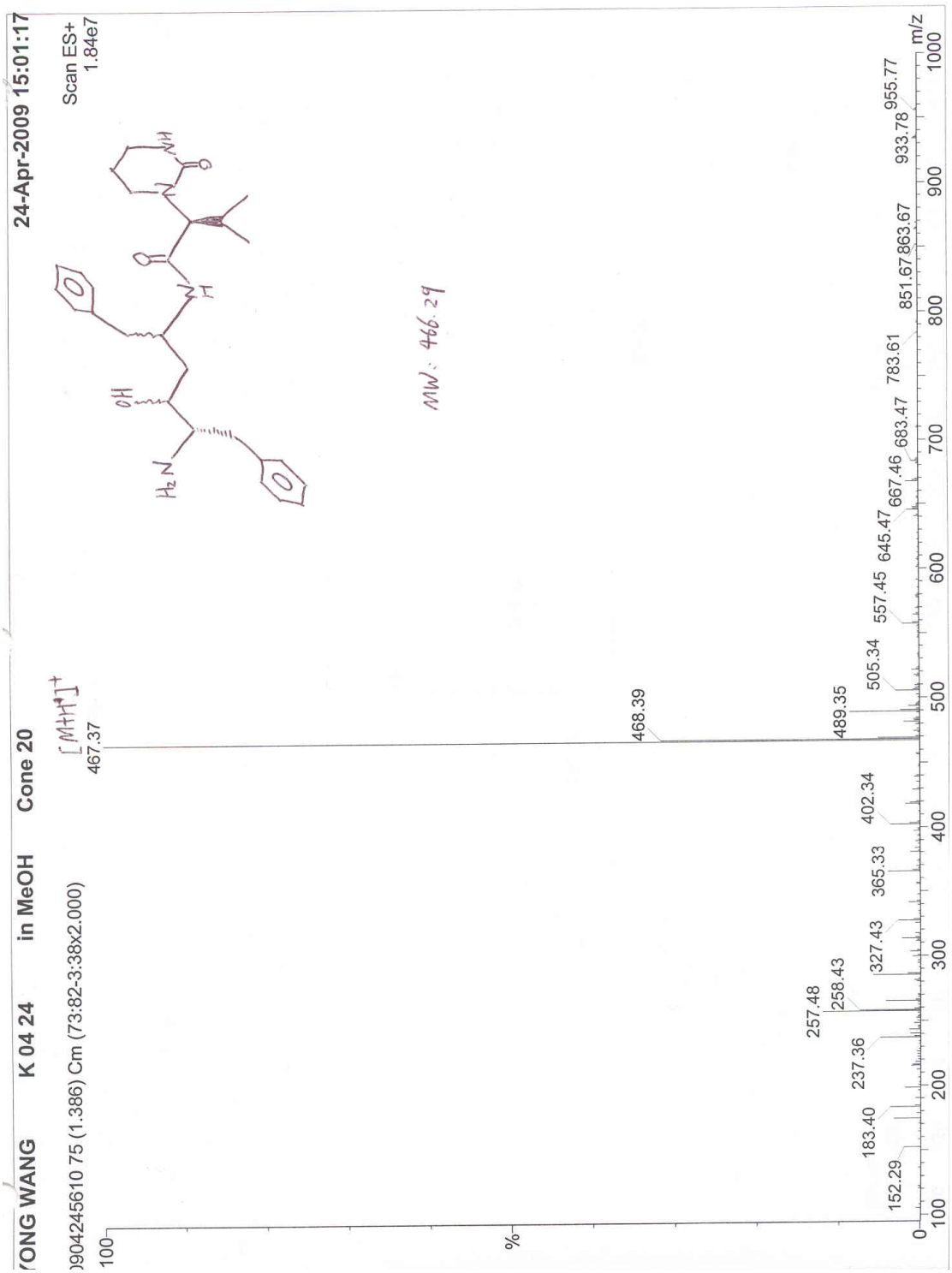
www.manaraa.com

Reaction 6



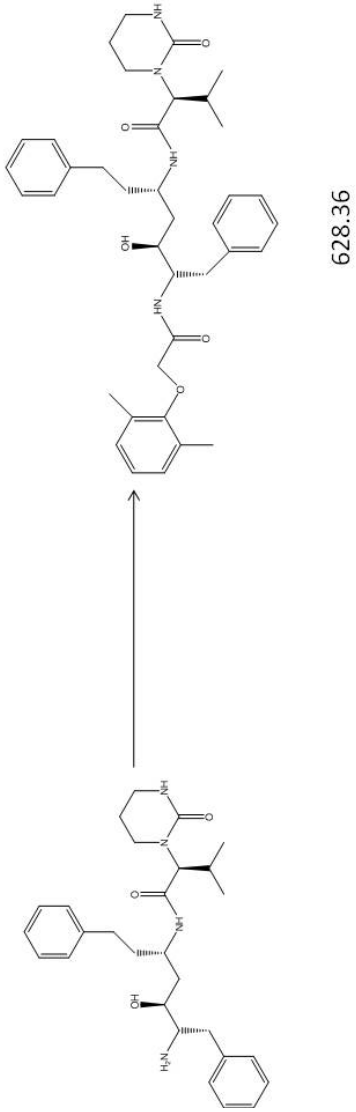
www.manaraa.com

Mass spectrometry result



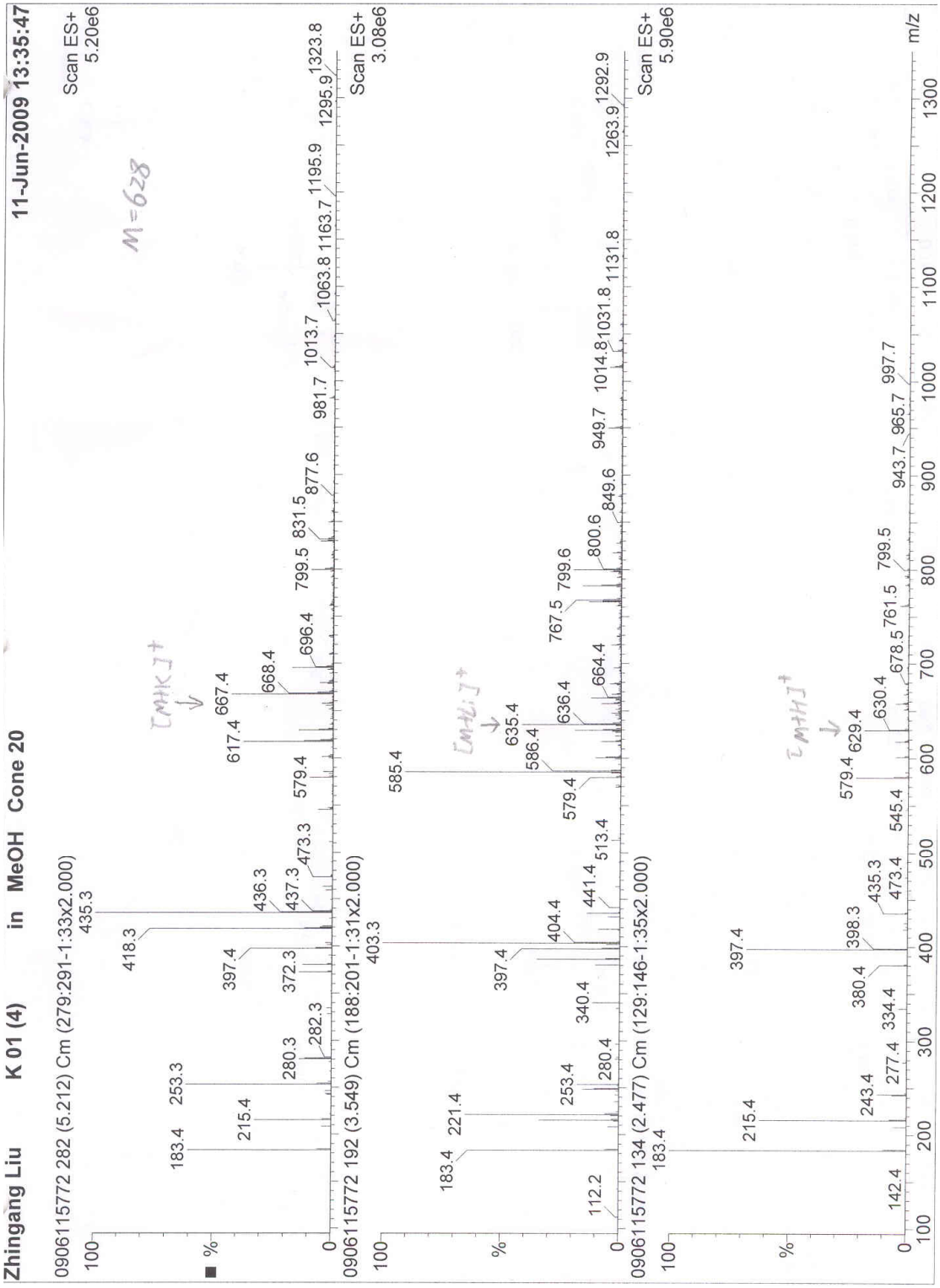
www.manaraa.com

Reaction 7



www.manaraa.com

Mass spectrometry results



www.manaraa.com

REFERENCES

1. Coffin, J. M. (1992) *Curr Top Microbiol Immunol* 176, 143-164
2. Levy, J. A. (2007) *HIV and the pathogenesis of AIDS*, 3rd ed., ASM Press, Washington, D.C.
3. Haase, A. T. (1986) *Nature* 322, 130-136
4. Gottlieb, M. S., Schroff, R., Schanker, H. M., Weisman, J. D., Fan, P. T., Wolf, R. A., and Saxon, A. (1981) *N Engl J Med* 305, 1425-1431
5. Chiu, I. M., Yaniv, A., Dahlberg, J. E., Gazit, A., Skuntz, S. F., Tronick, S. R., and Aaronson, S. A. (1985) *Nature* 317, 366-368
6. Gonda, M. A., Wong-Staal, F., Gallo, R. C., Clements, J. E., Narayan, O., and Gilden, R. V. (1985) *Science* 227, 173-177
7. Levy, J. A., Kaminsky, L. S., Morrow, W. J., Steimer, K., Luciw, P., Dina, D., Hoxie, J., and Oshiro, L. (1985) *Ann Intern Med* 103, 694-699
8. Rabson, A. B., and Martin, M. A. (1985) *Cell* 40, 477-480
9. Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. (1983) *Science* 220, 868-871
10. Gallo, R. C., Sarin, P. S., Gelmann, E. P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V. S., Mann, D., Sidhu, G. D., Stahl,

www.manaraa.com

R. E., Zolla-Pazner, S., Leibowitch, J., and Popovic, M. (1983) *Science* 220, 865-867

11. Montagnier, L., Gruest, J., Chamaret, S., Dauguet, C., Axler, C., Guetard, D., Nugeyre, M. T., Barre-Sinoussi, F., Chermann, J. C., Brunet, J. B., and et al. (1984) *Science* 225, 63-66

12. Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M., Kaplan, M., Haynes, B. F., Palker, T. J., Redfield, R., Oleske, J., Safai, B., and et al. (1984) *Science* 224, 500-503

13. Popovic, M., Sarngadharan, M. G., Read, E., and Gallo, R. C. (1984) *Science* 224, 497-500

14. Sarngadharan, M. G., Popovic, M., Bruch, L., Schupbach, J., and Gallo, R. C. (1984) *Science* 224, 506-508

15. Schupbach, J., Popovic, M., Gilden, R. V., Gonda, M. A., Sarngadharan, M. G., and Gallo, R. C. (1984) *Science* 224, 503-505

16. Levy, J. A., Hoffman, A. D., Kramer, S. M., Landis, J. A., Shimabukuro, J. M., and Oshiro, L. S. (1984) *Science* 225, 840-842

17. Coffin, J., Haase, A., Levy, J. A., Montagnier, L., Oroszlan, S., Teich, N., Temin, H., Toyoshima, K., Varmus, H., Vogt, P., and et al. (1986) *Science* 232, 697

18. Clavel, F., Guetard, D., Brun-Vezinet, F., Chamaret, S., Rey, M. A., Santos-Ferreira, M. O., Laurent, A. G., Dauguet, C., Katlama, C., Rouzioux, C., and et al. (1986) *Science* 233, 343-346

www.manaraa.com

19. UNAIDS. (2009)

20. Haase, A. T. (2005) *Nat Rev Immunol* 5, 783-792

21. Miller, C. J., Li, Q., Abel, K., Kim, E. Y., Ma, Z. M., Wietgrefe, S., La Franco-Scheuch, L., Compton, L., Duan, L., Shore, M. D., Zupancic, M., Busch, M., Carlis, J., Wolinsky, S., and Haase, A. T. (2005) *J Virol* 79, 9217-9227

22. Zhang, Z., Schuler, T., Zupancic, M., Wietgrefe, S., Staskus, K. A., Reimann, K. A., Reinhart, T. A., Rogan, M., Cavert, W., Miller, C. J., Veazey, R. S., Notermans, D., Little, S., Danner, S. A., Richman, D. D., Havlir, D., Wong, J., Jordan, H. L., Schacker, T. W., Racz, P., Tenner-Racz, K., Letvin, N. L., Wolinsky, S., and Haase, A. T. (1999) *Science* 286, 1353-1357

23. Embretson, J., Zupancic, M., Ribas, J. L., Burke, A., Racz, P., Tenner-Racz, K., and Haase, A. T. (1993) *Nature* 362, 359-362

24. Abel, K., Rocke, D. M., Chohan, B., Fritts, L., and Miller, C. J. (2005) *J Virol* 79, 12164-12172

25. Reynolds, M. R., Rakasz, E., Skinner, P. J., White, C., Abel, K., Ma, Z. M., Compton, L., Napoe, G., Wilson, N., Miller, C. J., Haase, A., and Watkins, D. I. (2005) *J Virol* 79, 9228-9235

26. Phillips, A. N. (1996) *Science* 271, 497-499

27. Lang, W., Perkins, H., Anderson, R. E., Royce, R., Jewell, N., and Winkelstein, W., Jr. (1989) *J Acquir Immune Defic Syndr* 2, 63-69

www.manaraa.com

28. McDonald, R. A., Mayers, D. L., Chung, R. C., Wagner, K. F., Ratto-Kim, S., Birx, D. L., and Michael, N. L. (1997) *J Virol* 71, 1871-1879

29. Levy, J. A., Mackewicz, C. E., and Barker, E. (1996) *Immunol Today* 17, 217-224

30. Rowland-Jones, S., Tan, R., and McMichael, A. (1997) *Adv Immunol* 65, 277-346

31. Ogg, G. S., Jin, X., Bonhoeffer, S., Dunbar, P. R., Nowak, M. A., Monard, S., Segal, J. P., Cao, Y., Rowland-Jones, S. L., Cerundolo, V., Hurley, A., Markowitz, M., Ho, D. D., Nixon, D. F., and McMichael, A. J. (1998) *Science* 279, 2103-2106

32. Landay, A. L., Mackewicz, C. E., and Levy, J. A. (1993) *Clin Immunol Immunopathol* 69, 106-116

33. Mackewicz, C. E., Ortega, H. W., and Levy, J. A. (1991) *J Clin Invest* 87, 1462-1466

34. Folkvord, J. M., Armon, C., and Connick, E. (2005) *AIDS Res Hum Retroviruses* 21, 363-370

35. Badley, A. D., Pilon, A. A., Landay, A., and Lynch, D. H. (2000) *Blood* 96, 2951-2964

36. van der Ende, M. E., Schutten, M., Raschdorff, B., Grossschupff, G., Racz, P., Osterhaus, A. D., and Tenner-Racz, K. (1999) *AIDS* 13, 1015-1019

www.manaraa.com

37. Pantaleo, G., Graziosi, C., and Fauci, A. S. (1993) *N Engl J Med* 328, 327-335

38. Kuznetsov, Y. G., Victoria, J. G., Robinson, W. E., Jr., and McPherson, A. (2003) *J Virol* 77, 11896-11909

39. McCune, J. M., Rabin, L. B., Feinberg, M. B., Lieberman, M., Kosek, J. C., Reyes, G. R., and Weissman, I. L. (1988) *Cell* 53, 55-67

40. Yu, X., Yuan, X., Matsuda, Z., Lee, T. H., and Essex, M. (1992) *J Virol* 66, 4966-4971

41. Gomez, C., and Hope, T. J. (2005) *Cell Microbiol* 7, 621-626

42. Camaur, D., and Trono, D. (1996) *J Virol* 70, 6106-6111

43. Liu, H., Wu, X., Newman, M., Shaw, G. M., Hahn, B. H., and Kappes, J. C. (1995) *J Virol* 69, 7630-7638

44. Pandori, M. W., Fitch, N. J., Craig, H. M., Richman, D. D., Spina, C. A., and Guatelli, J. C. (1996) *J Virol* 70, 4283-4290

45. Lu, Y. L., Spearman, P., and Ratner, L. (1993) *J Virol* 67, 6542-6550

46. Yu, X., Matsuda, Z., Yu, Q. C., Lee, T. H., and Essex, M. (1993) *J Virol* 67, 4386-4390

47. Ott, D. E., Coren, L. V., Kane, B. P., Busch, L. K., Johnson, D. G., Sowder, R. C., 2nd, Chertova, E. N., Arthur, L. O., and Henderson, L. E. (1996) *J Virol* 70, 7734-7743

48. Feinberg, M. B., Jarrett, R. F., Aldovini, A., Gallo, R. C., and Wong-Staal, F. (1986) *Cell* 46, 807-817

www.manaraa.com

49. Oroszlan, S., and Luftig, R. B. (1990) *Curr Top Microbiol Immunol* 157, 153-185

50. Hallenberger, S., Bosch, V., Angliker, H., Shaw, E., Klenk, H. D., and Garten, W. (1992) *Nature* 360, 358-361

51. Greene, W. C., and Peterlin, B. M. (2002) *Nat Med* 8, 673-680

52. Sodroski, J., Goh, W. C., Rosen, C., Dayton, A., Terwilliger, E., and Haseltine, W. (1986) *Nature* 321, 412-417

53. Sawai, E. T., Baur, A., Struble, H., Peterlin, B. M., Levy, J. A., and Cheng-Mayer, C. (1994) *Proc Natl Acad Sci U S A* 91, 1539-1543

54. von Schwedler, U., Song, J., Aiken, C., and Trono, D. (1993) *J Virol* 67, 4945-4955

55. Dalgleish, A. G., Beverley, P. C., Clapham, P. R., Crawford, D. H., Greaves, M. F., and Weiss, R. A. (1984) *Nature* 312, 763-767

56. Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J. C., and Montagnier, L. (1984) *Nature* 312, 767-768

57. Klatzmann, D., Barre-Sinoussi, F., Nugeyre, M. T., Danquet, C., Vilmer, E., Griscelli, C., Brun-Veziret, F., Rouzioux, C., Gluckman, J. C., Chermann, J. C., and et al. (1984) *Science* 225, 59-63

58. Arthos, J., Deen, K. C., Chaikin, M. A., Fornwald, J. A., Sathe, G., Sattentau, Q. J., Clapham, P. R., Weiss, R. A., McDougal, J. S., Pietropaolo, C., and et al. (1989) *Cell* 57, 469-481

www.manaraa.com

59. Jameson, B. A., Rao, P. E., Kong, L. I., Hahn, B. H., Shaw, G. M., Hood, L. E., and Kent, S. B. (1988) *Science* 240, 1335-1339

60. Mizukami, T., Fuerst, T. R., Berger, E. A., and Moss, B. (1988) *Proc Natl Acad Sci U S A* 85, 9273-9277

61. Moore, J. P., McKeating, J. A., Weiss, R. A., and Sattentau, Q. J. (1990) *Science* 250, 1139-1142

62. Sattentau, Q. J., and Moore, J. P. (1991) *J Exp Med* 174, 407-415

63. Schulz, T. F., Jameson, B. A., Lopalco, L., Siccardi, A. G., Weiss, R. A., and Moore, J. P. (1992) *AIDS Res Hum Retroviruses* 8, 1571-1580

64. Hatziioannou, T., Perez-Caballero, D., Cowan, S., and Bieniasz, P. D. (2005) *J Virol* 79, 176-183

65. Cullen, B. R. (2001) *Cell* 105, 697-700

66. Luban, J., Bossolt, K. L., Franke, E. K., Kalpana, G. V., and Goff, S. P. (1993) *Cell* 73, 1067-1078

67. Bushman, F., Lewinski, M., Ciuffi, A., Barr, S., Leipzig, J., Hannenhalli, S., and Hoffmann, C. (2005) *Nat Rev Microbiol* 3, 848-858

68. Luban, J. (1996) *Cell* 87, 1157-1159

69. Bukrinsky, M. I., Haggerty, S., Dempsey, M. P., Sharova, N., Adzhubel, A., Spitz, L., Lewis, P., Goldfarb, D., Emerman, M., and Stevenson, M. (1993) *Nature* 365, 666-669

www.manaraa.com

70. Heinzinger, N. K., Bukinsky, M. I., Haggerty, S. A., Ragland, A. M., Kewalramani, V., Lee, M. A., Gendelman, H. E., Ratner, L., Stevenson, M., and Emerman, M. (1994) *Proc Natl Acad Sci U S A* 91, 7311-7315

71. Bouyac-Bertoia, M., Dvorin, J. D., Fouchier, R. A., Jenkins, Y., Meyer, B. E., Wu, L. I., Emerman, M., and Malim, M. H. (2001) *Mol Cell* 7, 1025-1035

72. Fassati, A., and Goff, S. P. (2001) *J Virol* 75, 3626-3635

73. Reil, H., Bukovsky, A. A., Gelderblom, H. R., and Gottlinger, H. G. (1998) *EMBO J* 17, 2699-2708

74. Lehmann-Che, J., and Saib, A. (2004) *AIDS Rev* 6, 199-207

75. Bor, Y. C., Miller, M. D., Bushman, F. D., and Orgel, L. E. (1996) *Virology* 222, 283-288

76. Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. (2002) *Cell* 110, 521-529

77. Jacque, J. M., and Stevenson, M. (2006) *Nature* 441, 641-645

78. Murakami, T., and Freed, E. O. (2000) *Proc Natl Acad Sci U S A* 97, 343-348

79. Wyma, D. J., Kotov, A., and Aiken, C. (2000) *J Virol* 74, 9381-9387

80. Nguyen, D. H., and Hildreth, J. E. (2000) *J Virol* 74, 3264-3272

81. Ono, A., and Freed, E. O. (2001) *Proc Natl Acad Sci U S A* 98, 13925-13930

www.manaraa.com

82. Manes, S., del Real, G., Lacalle, R. A., Lucas, P., Gomez-Mouton, C., Sanchez-Palomino, S., Delgado, R., Alcami, J., Mira, E., and Martinez, A. C. (2000) *EMBO Rep* 1, 190-196

83. Gould, S. J., Booth, A. M., and Hildreth, J. E. (2003) *Proc Natl Acad Sci U S A* 100, 10592-10597

84. Chan, D. C., and Kim, P. S. (1998) *Cell* 93, 681-684

85. Tsibris, A. M., and Kuritzkes, D. R. (2007) *Annu Rev Med* 58, 445-459

86. Schinazi, R. F., Lloyd, R. M., Jr., Nguyen, M. H., Cannon, D. L., McMillan, A., Ilksoy, N., Chu, C. K., Liotta, D. C., Bazmi, H. Z., and Mellors, J. W. (1993) *Antimicrob Agents Chemother* 37, 875-881

87. Steigbigel, R. T., Cooper, D. A., Kumar, P. N., Eron, J. E., Schechter, M., Markowitz, M., Loutfy, M. R., Lennox, J. L., Gatell, J. M., Rockstroh, J. K., Katlama, C., Yeni, P., Lazzarin, A., Clotet, B., Zhao, J., Chen, J., Ryan, D. M., Rhodes, R. R., Killar, J. A., Gilde, L. R., Strohmaier, K. M., Meibohm, A. R., Miller, M. D., Hazuda, D. J., Nessly, M. L., DiNubile, M. J., Isaacs, R. D., Nguyen, B. Y., and Teppler, H. (2008) *N Engl J Med* 359, 339-354

88. Ortiz, R., Dejesus, E., Khanlou, H., Voronin, E., van Lunzen, J., Andrade-Villanueva, J., Fourie, J., De Meyer, S., De Pauw, M., Lefebvre, E., Vangeneugden, T., and Spinosa-Guzman, S. (2008) *AIDS* 22, 1389-1397

www.manaraa.com

89. Molina, J. M., Andrade-Villanueva, J., Echevarria, J., Chetchotisakd, P., Corral, J., David, N., Moyle, G., Mancini, M., Percival, L., Yang, R., Thiry, A., and McGrath, D. (2008) *Lancet* 372, 646-655

90. Wen, Z., Stern, S. T., Martin, D. E., Lee, K. H., and Smith, P. C. (2006) *Drug Metab Dispos* 34, 1436-1442

91. Holzgrabe, U. (2004) *Pharm Unserer Zeit* 33, 160

92. Li, F., Goila-Gaur, R., Salzwedel, K., Kilgore, N. R., Reddick, M., Matallana, C., Castillo, A., Zoumplis, D., Martin, D. E., Orenstein, J. M., Allaway, G. P., Freed, E. O., and Wild, C. T. (2003) *Proc Natl Acad Sci U S A* 100, 13555-13560

93. Li, F., Zoumplis, D., Matallana, C., Kilgore, N. R., Reddick, M., Yunus, A. S., Adamson, C. S., Salzwedel, K., Martin, D. E., Allaway, G. P., Freed, E. O., and Wild, C. T. (2006) *Virology* 356, 217-224

94. Takeuchi, Y., Nagumo, T., and Hoshino, H. (1988) *J Virol* 62, 3900-3902

95. Roberts, J. D., Preston, B. D., Johnston, L. A., Soni, A., Loeb, L. A., and Kunkel, T. A. (1989) *Mol Cell Biol* 9, 469-476

96. Weber, J., and Grosse, F. (1989) *Nucleic Acids Res* 17, 1379-1393

97. Clavel, F., and Hance, A. J. (2004) *N Engl J Med* 350, 1023-1035

98. Sukasem, C., Churdboonchart, V., Sukeepaisarncharoen, W., Piroj, W., Inwisai, T., Tiensuwan, M., and Chantratita, W. (2008) *Int J Antimicrob Agents* 31, 277-281

www.manaraa.com

99. Grabar, S., Pradier, C., Le Corfec, E., Lancar, R., Allavena, C., Bentata, M., Berlureau, P., Dupont, C., Fabbro-Peray, P., Poizot-Martin, I., and Costagliola, D. (2000) *AIDS* 14, 141-149

100. Gulick, R. M., Mellors, J. W., Havlir, D., Eron, J. J., Meibohm, A., Condra, J. H., Valentine, F. T., McMahon, D., Gonzalez, C., Jonas, L., Emini, E. A., Chodakewitz, J. A., Isaacs, R., and Richman, D. D. (2000) *Ann Intern Med* 133, 35-39

101. Condra, J. H., Schleif, W. A., Blahy, O. M., Gabryelski, L. J., Graham, D. J., Quintero, J. C., Rhodes, A., Robbins, H. L., Roth, E., Shivaprakash, M., and et al. (1995) *Nature* 374, 569-571

102. Schooley, R. T., and Mellors, J. W. (2007) *J Infect Dis* 195, 770-772

103. Thompson, M. A., Aberg, J. A., Cahn, P., Montaner, J. S., Rizzardini, G., Telenti, A., Gatell, J. M., Gunthard, H. F., Hammer, S. M., Hirsch, M. S., Jacobsen, D. M., Reiss, P., Richman, D. D., Volberding, P. A., Yeni, P., and Schooley, R. T. (2010) *JAMA* 304, 321-333

104. Natarajan, V., Bosche, M., Metcalf, J. A., Ward, D. J., Lane, H. C., and Kovacs, J. A. (1999) *Lancet* 353, 119-120

105. Bartlett, J. A., DeMasi, R., Quinn, J., Moxham, C., and Rousseau, F. (2001) *AIDS* 15, 1369-1377

106. Barbaro, G., Lucchini, A., and Barbarini, G. (2005) *Minerva Cardioangiol* 53, 153-154

107. Struble, K., Murray, J., Cheng, B., Gegeny, T., Miller, V., and Gulick,

www.manaraa.com

R. (2005) *AIDS* 19, 747-756

108. Coffin, J. M. (1995) *Science* 267, 483-489

109. Riddler, S. A., Haubrich, R., DiRienzo, A. G., Peeples, L., Powderly, W. G., Klingman, K. L., Garren, K. W., George, T., Rooney, J. F., Brizz, B., Lalloo, U. G., Murphy, R. L., Swindells, S., Havlir, D., and Mellors, J. W. (2008) *N Engl J Med* 358, 2095-2106

110. Wilson, D. P., Law, M. G., Grulich, A. E., Cooper, D. A., and Kaldor, J. M. (2008) *Lancet* 372, 314-320

111. Deeks, S. G., Hoh, R., Neilands, T. B., Liegler, T., Aweeka, F., Petropoulos, C. J., Grant, R. M., and Martin, J. N. (2005) *J Infect Dis* 192, 1537-1544

112. Dournon, E., Matheron, S., Rozenbaum, W., Gharakhanian, S., Michon, C., Girard, P. M., Perronne, C., Salmon, D., De Truchis, P., Leport, C., and et al. (1988) *Lancet* 2, 1297-1302

113. Mitsuya, H., Weinhold, K. J., Furman, P. A., St Clair, M. H., Lehrman, S. N., Gallo, R. C., Bolognesi, D., Barry, D. W., and Broder, S. (1985) *Proc Natl Acad Sci U S A* 82, 7096-7100

114. Deeks, S. G. (2003) *Lancet* 362, 2002-2011

115. Semenova, E. A., Marchand, C., and Pommier, Y. (2008) *Adv Pharmacol* 56, 199-228

116. Grobler, J. A., Stillmock, K., Hu, B., Witmer, M., Felock, P., Espeseth, A. S., Wolfe, A., Egbertson, M., Bourgeois, M., Melamed,

www.manaraa.com

J., Wai, J. S., Young, S., Vacca, J., and Hazuda, D. J. (2002) *Proc Natl Acad Sci U S A* 99, 6661-6666

117. Lataillade, M., and Kozal, M. J. (2006) *AIDS Patient Care STDS* 20, 489-501

118. Engelman, A., and Craigie, R. (1992) *J Virol* 66, 6361-6369

119. Jiang, S., Lin, K., Strick, N., and Neurath, A. R. (1993) *Nature* 365, 113

120. Wild, C., Oas, T., McDanal, C., Bolognesi, D., and Matthews, T. (1992) *Proc Natl Acad Sci U S A* 89, 10537-10541

121. Kilby, J. M., Lalezari, J. P., Eron, J. J., Carlson, M., Cohen, C., Arduino, R. C., Goodgame, J. C., Gallant, J. E., Volberding, P., Murphy, R. L., Valentine, F., Saag, M. S., Nelson, E. L., Sista, P. R., and Dusek, A. (2002) *AIDS Res Hum Retroviruses* 18, 685-693

122. Lalezari, J. P., Eron, J. J., Carlson, M., Cohen, C., DeJesus, E., Arduino, R. C., Gallant, J. E., Volberding, P., Murphy, R. L., Valentine, F., Nelson, E. L., Sista, P. R., Dusek, A., and Kilby, J. M. (2003) *AIDS* 17, 691-698

123. Derdeyn, C. A., Decker, J. M., Sfakianos, J. N., Wu, X., O'Brien, W. A., Ratner, L., Kappes, J. C., Shaw, G. M., and Hunter, E. (2000) *J Virol* 74, 8358-8367

124. Melby, T., Despirito, M., Demasi, R., Heilek-Snyder, G., Greenberg, M. L., and Graham, N. (2006) *J Infect Dis* 194, 238-246

www.manaraa.com

125. Reeves, J. D., Gallo, S. A., Ahmad, N., Miamidian, J. L., Harvey, P. E., Sharron, M., Pohlmann, S., Sfakianos, J. N., Derdeyn, C. A., Blumenthal, R., Hunter, E., and Doms, R. W. (2002) *Proc Natl Acad Sci U S A* 99, 16249-16254

126. Tilton, J. C., and Doms, R. W. (2010) *Antiviral Res* 85, 91-100

127. Skre, H. (1974) *Clin Genet* 6, 98-118

128. Jessen, K. R., and Mirsky, R. (2005) *Nat Rev Neurosci* 6, 671-682

129. Sherman, D. L., and Brophy, P. J. (2005) *Nat Rev Neurosci* 6, 683-690

130. Edgar, J. M., and Garbern, J. (2004) *J Neurosci Res* 76, 593-598

131. Esper, R. M., Pankonin, M. S., and Loeb, J. A. (2006) *Brain Res Rev* 51, 161-175

132. Poliak, S., and Peles, E. (2003) *Nat Rev Neurosci* 4, 968-980

133. Suter, U., and Scherer, S. S. (2003) *Nat Rev Neurosci* 4, 714-726

134. Niemann, A., Berger, P., and Suter, U. (2006) *Neuromolecular Med* 8, 217-242

135. Berger, P., Niemann, A., and Suter, U. (2006) *Glia* 54, 243-257

136. Roa, B. B., Dyck, P. J., Marks, H. G., Chance, P. F., and Lupski, J. R. (1993) *Nat Genet* 5, 269-273

137. Raeymaekers, P., Timmerman, V., De Jonghe, P., Swerts, L., Gheuens, J., Martin, J. J., Muylle, L., De Winter, G., Vandenberghe, A., and Van Broeckhoven, C. (1989) *Am J Hum Genet* 45, 953-958

www.manaraa.com

138. Vance, J. M., Nicholson, G. A., Yamaoka, L. H., Stajich, J., Stewart, C. S., Speer, M. C., Hung, W. Y., Roses, A. D., Barker, D., and Pericak-Vance, M. A. (1989) *Exp Neurol* 104, 186-189

139. Lupski, J. R., de Oca-Luna, R. M., Slaugenhaupt, S., Pentao, L., Guzzetta, V., Trask, B. J., Saucedo-Cardenas, O., Barker, D. F., Killian, J. M., Garcia, C. A., Chakravarti, A., and Patel, P. I. (1991) *Cell* 66, 219-232

140. Raeymaekers, P., Timmerman, V., Nelis, E., De Jonghe, P., Hoogendijk, J. E., Baas, F., Barker, D. F., Martin, J. J., De Visser, M., Bolhuis, P. A., and et al. (1991) *Neuromuscul Disord* 1, 93-97

141. Valentijn, L. J., Baas, F., Wolterman, R. A., Hoogendijk, J. E., van den Bosch, N. H., Zorn, I., Gabreels-Festen, A. W., de Visser, M., and Bolhuis, P. A. (1992) *Nat Genet* 2, 288-291

142. Nelis, E., Van Broeckhoven, C., De Jonghe, P., Lofgren, A., Vandenberghe, A., Latour, P., Le Guern, E., Brice, A., Mostacciuolo, M. L., Schiavon, F., Palau, F., Bort, S., Upadhyaya, M., Rocchi, M., Archidiacono, N., Mandich, P., Bellone, E., Silander, K., Savontaus, M. L., Navon, R., Goldberg-Stern, H., Estivill, X., Volpini, V., Friedl, W., Gal, A., and et al. (1996) *Eur J Hum Genet* 4, 25-33

143. Hoogendijk, J. E., Hensels, G. W., Zorn, I., Valentijn, L., Janssen, E. A., de Visser, M., Barker, D. F., Ongerboer de Visser, B. W., Baas, F., and Bolhuis, P. A. (1991) *Hum Genet* 88, 215-218

www.manaraa.com

144. Chance, P. F., Alderson, M. K., Leppig, K. A., Lensch, M. W., Matsunami, N., Smith, B., Swanson, P. D., Odelberg, S. J., Disteche, C. M., and Bird, T. D. (1993) *Cell* 72, 143-151

145. Suter, U., Moskow, J. J., Welcher, A. A., Snipes, G. J., Kosaras, B., Sidman, R. L., Buchberg, A. M., and Shooter, E. M. (1992) *Proc Natl Acad Sci U S A* 89, 4382-4386

146. Ionasescu, V. V., Searby, C. C., Ionasescu, R., Chatkupt, S., Patel, N., and Koenigsberger, R. (1997) *Muscle Nerve* 20, 97-99

147. Suter, U., Welcher, A. A., Ozcelik, T., Snipes, G. J., Kosaras, B., Francke, U., Billings-Gagliardi, S., Sidman, R. L., and Shooter, E. M. (1992) *Nature* 356, 241-244

148. Snipes, G. J., Suter, U., Welcher, A. A., and Shooter, E. M. (1992) *J Cell Biol* 117, 225-238

149. Naef, R., and Suter, U. (1998) *Microsc Res Tech* 41, 359-371

150. Vallat, J. M. (2003) *J Neuropathol Exp Neurol* 62, 699-714

151. Hayasaka, K., Himoro, M., Sato, W., Takada, G., Uyemura, K., Shimizu, N., Bird, T. D., Conneally, P. M., and Chance, P. F. (1993) *Nat Genet* 5, 31-34

152. Hayasaka, K., Himoro, M., Sawaishi, Y., Nanao, K., Takahashi, T., Takada, G., Nicholson, G. A., Ouvrier, R. A., and Tachi, N. (1993) *Nat Genet* 5, 266-268

153. Hayasaka, K., Takada, G., and Ionasescu, V. V. (1993) *Hum Mol*

www.manaraa.com

Genet 2, 1369-1372

154. Hayasaka, K., Himoro, M., Wang, Y., Takata, M., Minoshima, S., Shimizu, N., Miura, M., Uyemura, K., and Takada, G. (1993) *Genomics* 17, 755-758

155. Shy, M. E., Jani, A., Krajewski, K., Grandis, M., Lewis, R. A., Li, J., Shy, R. R., Balsamo, J., Lilien, J., Garbern, J. Y., and Kamholz, J. (2004) *Brain* 127, 371-384

156. Nelis, E., Haites, N., and Van Broeckhoven, C. (1999) *Hum Mutat* 13, 11-28

157. Shy, M. E. (2006) *J Neurol Sci* 242, 55-66

158. Konde, V., and Eichberg, J. (2006) *J Neurosci Res* 83, 957-964

159. Shames, I., Fraser, A., Colby, J., Orfali, W., and Snipes, G. J. (2003) *J Neuropathol Exp Neurol* 62, 751-764

160. Lemke, G., Lamar, E., and Patterson, J. (1988) *Neuron* 1, 73-83

161. Eylar, E. H., Uyemura, K., Brostoff, S. W., Kitamura, K., Ishaque, A., and Greenfield, S. (1979) *Neurochem Res* 4, 289-293

162. Wong, M. H., and Filbin, M. T. (1994) *J Cell Biol* 126, 1089-1097

163. Filbin, M. T., Walsh, F. S., Trapp, B. D., Pizzey, J. A., and Tennekoon, G. I. (1990) *Nature* 344, 871-872

164. Xu, W., Shy, M., Kamholz, J., Elferink, L., Xu, G., Lilien, J., and Balsamo, J. (2001) *J Cell Biol* 155, 439-446.

165. Filbin, M. T., Walsh, F. S., Trapp, B. D., Pizzey, J. A., and

www.manaraa.com

Tennekoon, G. I. (1990) *Nature* 344, 871-872

166. Shapiro, L., Doyle, J. P., Hensley, P., Colman, D. R., and Hendrickson, W. A. (1996) *Neuron* 17, 435-449

167. Giese, K. P., Martini, R., Lemke, G., Soriano, P., and Schachner, M. (1992) *Cell* 71, 565-576

168. Wrabetz, L., Feltri, M. L., Quattrini, A., Imperiale, D., Previtali, S., D'Antonio, M., Martini, R., Yin, X., Trapp, B. D., Zhou, L., Chiu, S. Y., and Messing, A. (2000) *J Cell Biol* 148, 1021-1034

169. Filbin, M. T., Zhang, K., Li, W., and Gao, Y. (1999) *Ann N Y Acad Sci* 883, 160-167

170. Street, V. A., Goldy, J. D., Golden, A. S., Tempel, B. L., Bird, T. D., and Chance, P. F. (2002) *Am J Hum Genet* 70, 244-250

171. Street, V. A., Bennett, C. L., Goldy, J. D., Shirk, A. J., Kleopa, K. A., Tempel, B. L., Lipe, H. P., Scherer, S. S., Bird, T. D., and Chance, P. F. (2003) *Neurology* 60, 22-26

172. Bennett, C. L., Shirk, A. J., Huynh, H. M., Street, V. A., Nelis, E., Van Maldergem, L., De Jonghe, P., Jordanova, A., Guergueltcheva, V., Tournev, I., Van Den Bergh, P., Seeman, P., Mazanec, R., Prochazka, T., Kremensky, I., Haberlova, J., Weiss, M. D., Timmerman, V., Bird, T. D., and Chance, P. F. (2004) *Ann Neurol* 55, 713-720

173. Saifi, G. M., Szigeti, K., Wiszniewski, W., Shy, M. E., Krajewski, K.,

www.manaraa.com

Hausmanowa-Petrusewicz, I., Kochanski, A., Reeser, S., Mancias, P., Butler, I., and Lupski, J. R. (2005) *Hum Mutat* 25, 372-383

174. Warner, L. E., Mancias, P., Butler, I. J., McDonald, C. M., Keppen, L., Koob, K. G., and Lupski, J. R. (1998) *Nat Genet* 18, 382-384

175. Mirsky, R., and Jessen, K. R. (1999) *Brain Pathol* 9, 293-311

176. Zorick, T. S., Syroid, D. E., Brown, A., Gridley, T., and Lemke, G. (1999) *Development* 126, 1397-1406

177. Nagarajan, R., Svaren, J., Le, N., Araki, T., Watson, M., and Milbrandt, J. (2001) *Neuron* 30, 355-368

178. Berger, P., Young, P., and Suter, U. (2002) *Neurogenetics* 4, 1-15

179. Leblanc, S. E., Srinivasan, R., Ferri, C., Mager, G. M., Gillian-Daniel, A. L., Wrabetz, L., and Svaren, J. (2005) *J Neurochem* 93, 737-748

180. Julian, B. A., Phillips, J. A., 3rd, Orlando, P. J., Wyatt, R. J., and Butler, M. G. (1987) *Semin Nephrol* 7, 306-310

181. Liu, Q., Xie, F., Siedlak, S. L., Nunomura, A., Honda, K., Moreira, P. I., Zhua, X., Smith, M. A., and Perry, G. (2004) *Cell Mol Life Sci* 61, 3057-3075

182. Bergoffen, J., Scherer, S. S., Wang, S., Scott, M. O., Bone, L. J., Paul, D. L., Chen, K., Lensch, M. W., Chance, P. F., and Fischbeck, K. H. (1993) *Science* 262, 2039-2042

183. Vickrey, J. F., Logsdon, B. C., Proteasa, G., Palmer, S., Winters, M. A., Merigan, T. C., and Kovari, L. C. (2003) *Protein Expr Purif* 28,

www.manaraa.com

165-172

184. Laemmli, U. K. (1970) *Nature* 227, 680-685

185. Liu, Y., Manna, A., Li, R., Martin, W. E., Murphy, R. C., Cheung, A. L., and Zhang, G. (2001) *Proc Natl Acad Sci U S A* 98, 6877-6882

186. Kobe, B., Center, R. J., Kemp, B. E., and Poumbourios, P. (1999) *Proc Natl Acad Sci U S A* 96, 4319-4324

187. Ke, A., and Wolberger, C. (2003) *Protein Sci* 12, 306-312

188. Sharff, A. J., Rodseth, L. E., Spurlino, J. C., and Quiocho, F. A. (1992) *Biochemistry* 31, 10657-10663

189. Spurlino, J. C., Lu, G. Y., and Quiocho, F. A. (1991) *J Biol Chem* 266, 5202-5219

190. Quiocho, F. A., Spurlino, J. C., and Rodseth, L. E. (1997) *Structure* 5, 997-1015

191. Bond, J. P., Saavedra, R. A., and Kirschner, D. A. (2001) *Protein Expr Purif* 23, 398-410

192. Luft, J. R., Collins, R. J., Fehrman, N. A., Lauricella, A. M., Veatch, C. K., and DeTitta, G. T. (2003) *J Struct Biol* 142, 170-179

193. (1994) *Acta Crystallogr D Biol Crystallogr* 50, 760-763

194. Logsdon, B. C., Vickrey, J. F., Martin, P., Proteasa, G., Koepke, J. I., Terlecky, S. R., Wawrzak, Z., Winters, M. A., Merigan, T. C., and Kovari, L. C. (2004) *J Virol* 78, 3123-3132

195. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta*

www.manaraa.com

Crystallogr D Biol Crystallogr 53, 240-255

196. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) *Acta Crystallogr D Biol Crystallogr* 60, 2184-2195

197. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132

198. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr D Biol Crystallogr* 49, 129-147

199. Vaguine, A. A., Richelle, J., and Wodak, S. J. (1999) *Acta Crystallogr D Biol Crystallogr* 55, 191-205

200. Krissinel, E., and Henrick, K. (2004) *Acta Crystallogr D Biol Crystallogr* 60, 2256-2268

201. Krissinel, E., and Henrick, K. (2007) *J Mol Biol* 372, 774-797

202. Anjuere, F., Monsigny, M., Lelievre, Y., and Mayer, R. (1993) *Biochem J* 291 (Pt 3), 869-873

203. Geohegan, K. F., Spencer, R. W., Danley, D. E., Contillo, L. G., Jr., and Andrews, G. C. (1990) *FEBS Lett* 262, 119-122

204. Stoner, E. J., Cooper, A. J., Dickman, D. A., Kolaczkowski, L., Lallaman, J. E., Liu, J. H., Oliver-Shaffer, P. A., Patel, K. M., Paterson, J. B., Plata, D. J., Riley, D. A., Sham, H. L., Stengel, P. J., and Tien, J. H. J. (2000) *Organic Process Research & Development* 4, 264-269

www.manaraa.com

205. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2002) *Structure* 10, 369-381

206. Martin, P., Vickrey, J. F., Proteasa, G., Jimenez, Y. L., Wawrzak, Z., Winters, M. A., Merigan, T. C., and Kovari, L. C. (2005) *Structure* 13, 1887-1895

207. Lindahl, E., Hess, B., and van der Spoel, D. (2001) *Journal of Molecular Modeling* 7, 306-317

208. Hermans, J., Jr., Lohr, D., and Ferro, D. (1969) *Nature* 224, 175-177

209. Peng, C., Ho, B. K., Chang, T. W., and Chang, N. T. (1989) *J Virol* 63, 2550-2556

210. Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A., Scolnick, E. M., and Sigal, I. S. (1988) *Proc Natl Acad Sci U S A* 85, 4686-4690

211. Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J. R., Whittle, P. J., Danley, D. E., Geoghegan, K. F., and et al. (1989) *Nature* 342, 299-302

212. Navia, M. A., Fitzgerald, P. M., McKeever, B. M., Leu, C. T., Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L., and Springer, J. P. (1989) *Nature* 337, 615-620

213. Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., and Kent, S. B. (1989) *Science* 245, 616-621

www.manaraa.com

214. Spinelli, S., Liu, Q. Z., Alzari, P. M., Hirel, P. H., and Poljak, R. J. (1991) *Biochimie* 73, 1391-1396

215. Kumar, M., Kannan, K. K., Hosur, M. V., Bhavesh, N. S., Chatterjee, A., Mittal, R., and Hosur, R. V. (2002) *Biochem Biophys Res Commun* 294, 395-401

216. Ishima, R., Torchia, D. A., Lynch, S. M., Gronenborn, A. M., and Louis, J. M. (2003) *J Biol Chem* 278, 43311-43319

217. Liu, F., Kovalevsky, A. Y., Louis, J. M., Boross, P. I., Wang, Y. F., Harrison, R. W., and Weber, I. T. (2006) *J Mol Biol* 358, 1191-1199

218. Heaslet, H., Rosenfeld, R., Giffin, M., Lin, Y. C., Tam, K., Torbett, B. E., Elder, J. H., McRee, D. E., and Stout, C. D. (2007) *Acta Crystallogr D Biol Crystallogr* 63, 866-875

219. Louis, J. M., Ishima, R., Torchia, D. A., and Weber, I. T. (2007) *Adv Pharmacol* 55, 261-298

220. Ishima, R., and Louis, J. M. (2008) *Proteins* 70, 1408-1415

221. Ganser-Pornillos, B. K., Yeager, M., and Sundquist, W. I. (2008) *Curr Opin Struct Biol* 18, 203-217

222. Palmer, S., Shafer, R. W., and Merigan, T. C. (1999) *AIDS* 13, 661-667

223. Hornak, V., Okur, A., Rizzo, R. C., and Simmerling, C. (2006) *Proc Natl Acad Sci U S A* 103, 915-920

224. Lexa, K. W., Damm, K. L., Quintero, J. J., Gestwicki, J. E., and

www.manaraa.com

Carlson, H. A. (2009) *Proteins* 74, 872-880

225. Layten, M., Hornak, V., and Simmerling, C. (2006) *J Am Chem Soc* 128, 13360-13361

226. Doyon, L., Croteau, G., Thibeault, D., Poulin, F., Pilote, L., and Lamarre, D. (1996) *J Virol* 70, 3763-3769

227. Kelly, B. N., Kyere, S., Kinde, I., Tang, C., Howard, B. R., Robinson, H., Sundquist, W. I., Summers, M. F., and Hill, C. P. (2007) *J Mol Biol* 373, 355-366

228. Croteau, G., Doyon, L., Thibeault, D., McKercher, G., Pilote, L., and Lamarre, D. (1997) *J Virol* 71, 1089-1096

229. Martinez-Picado, J., Savara, A. V., Sutton, L., and D'Aquila, R. T. (1999) *J Virol* 73, 3744-3752

230. Darke, P. L., Nutt, R. F., Brady, S. F., Garsky, V. M., Ciccarone, T. M., Leu, C. T., Lumma, P. K., Freidinger, R. M., Veber, D. F., and Sigal, I. S. (1988) *Biochem Biophys Res Commun* 156, 297-303

231. Tozser, J., Gustchina, A., Weber, I. T., Blaha, I., Wondrak, E. M., and Oroszlan, S. (1991) *FEBS Lett* 279, 356-360

232. Tozser, J., Weber, I. T., Gustchina, A., Blaha, I., Copeland, T. D., Louis, J. M., and Oroszlan, S. (1992) *Biochemistry* 31, 4793-4800

233. Griffiths, J. T., Phylip, L. H., Konvalinka, J., Strop, P., Gustchina, A., Wlodawer, A., Davenport, R. J., Briggs, R., Dunn, B. M., and Kay, J. (1992) *Biochemistry* 31, 5193-5200

www.manaraa.com

234. Cameron, C. E., Grinde, B., Jacques, P., Jentoft, J., Leis, J., Wlodawer, A., and Weber, I. T. (1993) *J Biol Chem* 268, 11711-11720

235. Poorman, R. A., Tomasselli, A. G., Heinrikson, R. L., and Kezdy, F. J. (1991) *J Biol Chem* 266, 14554-14561

236. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2000) *J Mol Biol* 301, 1207-1220

237. Pettit, S. C., Moody, M. D., Wehbie, R. S., Kaplan, A. H., Nantermet, P. V., Klein, C. A., and Swanstrom, R. (1994) *J Virol* 68, 8017-8027

238. Altman, M. D., Nalivaika, E. A., Prabu-Jeyabalan, M., Schiffer, C. A., and Tidor, B. (2008) *Proteins* 70, 678-694

239. Partaledis, J. A., Yamaguchi, K., Tisdale, M., Blair, E. E., Falcione, C., Maschera, B., Myers, R. E., Pazhanisamy, S., Futer, O., Cullinan, A. B., and et al. (1995) *J Virol* 69, 5228-5235

240. Craig, J. C., Duncan, I. B., Hockley, D., Grief, C., Roberts, N. A., and Mills, J. S. (1991) *Antiviral Res* 16, 295-305

241. Kempf, D. J., Marsh, K. C., Denissen, J. F., McDonald, E., Vasavanonda, S., Flentge, C. A., Green, B. E., Fino, L., Park, C. H., Kong, X. P., and et al. (1995) *Proc Natl Acad Sci U S A* 92, 2484-2488

242. Koh, Y., Nakata, H., Maeda, K., Ogata, H., Bilcer, G., Devasamudram, T., Kincaid, J. F., Boross, P., Wang, Y. F., Tie, Y., Volarath, P., Gaddis, L., Harrison, R. W., Weber, I. T., Ghosh, A. K.,

www.manaraa.com

and Mitsuya, H. (2003) *Antimicrob Agents Chemother* 47, 3123-3129

243. Patick, A. K., Mo, H., Markowitz, M., Appelt, K., Wu, B., Musick, L., Kalish, V., Kaldor, S., Reich, S., Ho, D., and Webber, S. (1996) *Antimicrob Agents Chemother* 40, 292-297

244. Robinson, B. S., Riccardi, K. A., Gong, Y. F., Guo, Q., Stock, D. A., Blair, W. S., Terry, B. J., Deminie, C. A., Djang, F., Colonno, R. J., and Lin, P. F. (2000) *Antimicrob Agents Chemother* 44, 2093-2099

245. Sham, H. L., Kempf, D. J., Molla, A., Marsh, K. C., Kumar, G. N., Chen, C. M., Kati, W., Stewart, K., Lal, R., Hsu, A., Betebenner, D., Korneyeva, M., Vasavanonda, S., McDonald, E., Saldivar, A., Wideburg, N., Chen, X., Niu, P., Park, C., Jayanti, V., Grabowski, B., Granneman, G. R., Sun, E., Japour, A. J., Leonard, J. M., Plattner, J. J., and Norbeck, D. W. (1998) *Antimicrob Agents Chemother* 42, 3218-3224

246. Vacca, J. P., Dorsey, B. D., Schleif, W. A., Levin, R. B., McDaniel, S. L., Darke, P. L., Zugay, J., Quintero, J. C., Blahy, O. M., Roth, E., and et al. (1994) *Proc Natl Acad Sci U S A* 91, 4096-4100

247. Turner, S. R., Strohbach, J. W., Tommasi, R. A., Aristoff, P. A., Johnson, P. D., Skulnick, H. I., Dolak, L. A., Seest, E. P., Tomich, P. K., Bohanon, M. J., Horng, M. M., Lynn, J. C., Chong, K. T., Hinshaw, R. R., Watenpaugh, K. D., Janakiraman, M. N., and Thaisrivongs, S.

www.manaraa.com

(1998) *J Med Chem* 41, 3467-3476

248. Kempf, D. J., King, M. S., Bernstein, B., Cernohous, P., Bauer, E., Moseley, J., Gu, K., Hsu, A., Brun, S., and Sun, E. (2004) *J Infect Dis* 189, 51-60

249. Liu, T. F., and Shafer, R. W. (2006) *Clin Infect Dis* 42, 1608-1618

250. Friend, J., Parkin, N., Liegler, T., Martin, J. N., and Deeks, S. G. (2004) *AIDS* 18, 1965-1966

251. Conradie, F., Sanne, I., Venter, W., and Eron, J. (2004) *AIDS* 18, 1084-1085

252. Nijhuis, M., Wensing, A. M., Bierman, W. F., de Jong, D., Kagan, R., Fun, A., Jaspers, C. A., Schurink, K. A., van Agtmael, M. A., and Boucher, C. A. (2009) *J Infect Dis* 200, 698-709

253. Gifford, R. J., Liu, T. F., Rhee, S. Y., Kiuchi, M., Hue, S., Pillay, D., and Shafer, R. W. (2009) *Bioinformatics* 25, 1197-1198

254. Mo, H., King, M. S., King, K., Molla, A., Brun, S., and Kempf, D. J. (2005) *J Virol* 79, 3329-3338

255. Li, J., Bai, Y., Ianakova, E., Grandis, M., Uchwat, F., Trostinskaia, A., Krajewski, K. M., Garbern, J., Kupsky, W. J., and Shy, M. E. (2006) *J Comp Neurol* 498, 252-265

256. Warner, L. E., Hilz, M. J., Appel, S. H., Killian, J. M., Kolodry, E. H., Karpati, G., Carpenter, S., Watters, G. V., Wheeler, C., Witt, D., Bodell, A., Nelis, E., Van Broeckhoven, C., and Lupski, J. R. (1996)

www.manaraa.com

Neuron 17, 451-460

257. Gabreels-Festen, A. A., Hoogendijk, J. E., Meijerink, P. H., Gabreels, F. J., Bolhuis, P. A., van Beersum, S., Kulkens, T., Nelis, E., Jennekens, F. G., de Visser, M., van Engelen, B. G., Van Broeckhoven, C., and Mariman, E. C. (1996) *Neurology* 47, 761-765

258. Komiyama, A., Ohnishi, A., Izawa, K., Yamamori, S., Ohashi, H., and Hasegawa, O. (1997) *J Neurol Sci* 149, 103-109

259. Marrosu, M. G., Vaccargiu, S., Marrosu, G., Vannelli, A., Cianchetti, C., and Muntoni, F. (1998) *Neurology* 50, 1397-1401

260. Meijerink, P. H., Hoogendijk, J. E., Gabreels-Festen, A. A., Zorn, I., Veldman, H., Baas, F., de Visser, M., and Bolhuis, P. A. (1996) *Ann Neurol* 40, 672-675

261. Ikegami, T., Nicholson, G., Ikeda, H., Ishida, A., Johnston, H., Wise, G., Ouvrier, R., and Hayasaka, K. (1996) *Biochem Biophys Res Commun* 222, 107-110

262. Grandis, M., Vigo, T., Passalacqua, M., Jain, M., Scazzola, S., La Padula, V., Brucal, M., Benvenuto, F., Nobbio, L., Cadoni, A., Mancardi, G. L., Kamholz, J., Shy, M. E., and Schenone, A. (2008) *Hum Mol Genet* 17, 1877-1889

263. Briani, C., Adami, F., Cavallaro, T., Taioli, F., Ferrari, S., and Fabrizi, G. (2008) *Muscle Nerve* 38, 921-923

264. Kelly, J. W. (1998) *Curr Opin Struct Biol* 8, 101-106

www.manaraa.com

265. Dobson, C. M. (1999) *Trends Biochem Sci* 24, 329-332

266. Bellotti, V., Mangione, P., and Stoppini, M. (1999) *Cell Mol Life Sci* 55, 977-991

267. Rochet, J. C., and Lansbury, P. T., Jr. (2000) *Curr Opin Struct Biol* 10, 60-68

268. Wright, P. E., and Dyson, H. J. (1999) *J Mol Biol* 293, 321-331

269. Uversky, V. N. (2002) *Eur J Biochem* 269, 2-12

270. Uversky, V. N. (2002) *Protein Sci* 11, 739-756

271. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) *Biochemistry* 41, 6573-6582

272. Dunker, A. K., Brown, C. J., and Obradovic, Z. (2002) *Adv Protein Chem* 62, 25-49

273. Tompa, P. (2002) *Trends Biochem Sci* 27, 527-533

274. Dyson, H. J., and Wright, P. E. (2002) *Curr Opin Struct Biol* 12, 54-60

275. Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997) *J Mol Biol* 273, 729-739

276. Comenzo, R. L. (2007) *J Natl Compr Canc Netw* 5, 179-187

277. Gertz, M. A. (2004) *Am J Clin Pathol* 121, 787-789

278. Miyazaki, D., Yazaki, M., Gono, T., Kametani, F., Tsuchiya, A., Matsuda, M., Takenaka, Y., Hosh, Y., 2nd, and Ikeda, S. (2008) *Amyloid* 15, 125-128

www.manaraa.com

279. Ionescu-Zanetti, C., Khurana, R., Gillespie, J. R., Petrick, J. S., Trabachino, L. C., Minert, L. J., Carter, S. A., and Fink, A. L. (1999) *Proc Natl Acad Sci U S A* 96, 13175-13179

280. Astbury, W. T., and Dickinson, S. (1935) *Biochem J* 29, 2351-2360 2351

281. Lai, C., Brow, M. A., Nave, K. A., Noronha, A. B., Quarles, R. H., Bloom, F. E., Milner, R. J., and Sutcliffe, J. G. (1987) *Proc Natl Acad Sci U S A* 84, 4337-4341

282. Bork, P., Holm, L., and Sander, C. (1994) *J Mol Biol* 242, 309-320

283. Lesk, A. M., and Chothia, C. (1982) *J Mol Biol* 160, 325-342

284. Jones, E. Y., Davis, S. J., Williams, A. F., Harlos, K., and Stuart, D. I. (1992) *Nature* 360, 232-239

285. Wang, J. H., Yan, Y. W., Garrett, T. P., Liu, J. H., Rodgers, D. W., Garlick, R. L., Tarr, G. E., Husain, Y., Reinherz, E. L., and Harrison, S. C. (1990) *Nature* 348, 411-418

286. Ryu, S. E., Kwong, P. D., Truneh, A., Porter, T. G., Arthos, J., Rosenberg, M., Dai, X. P., Xuong, N. H., Axel, R., Sweet, R. W., and et al. (1990) *Nature* 348, 419-426

287. Leahy, D. J., Axel, R., and Hendrickson, W. A. (1992) *Cell* 68, 1145-1162

288. Jones, E. Y., Harlos, K., Bottomley, M. J., Robinson, R. C., Driscoll, P. C., Edwards, R. M., Clements, J. M., Dudgeon, T. J., and Stuart, D.

www.manaraa.com

I. (1995) *Nature* 373, 539-544

289. Derewenda, Z. S. (2010) *Acta Crystallogr D Biol Crystallogr* 66, 604-615

290. Corsini, L., Hothorn, M., Scheffzek, K., Sattler, M., and Stier, G. (2008) *Protein Sci* 17, 2070-2079

291. Stoll, V. S., Manohar, A. V., Gillon, W., MacFarlane, E. L., Hynes, R. C., and Pai, E. F. (1998) *Protein Sci* 7, 1147-1155

292. Ware, S., Donahue, J. P., Hawiger, J., and Anderson, W. F. (1999) *Protein Sci* 8, 2663-2671

293. Zhang, Z., Devarajan, P., Dorfman, A. L., and Morrow, J. S. (1998) *J Biol Chem* 273, 18681-18684

294. Tang, L., Guo, B., Javed, A., Choi, J. Y., Hiebert, S., Lian, J. B., van Wijnen, A. J., Stein, J. L., Stein, G. S., and Zhou, G. W. (1999) *J Biol Chem* 274, 33580-33586

295. Kuge, M., Fujii, Y., Shimizu, T., Hirose, F., Matsukage, A., and Hakoshima, T. (1997) *Protein Sci* 6, 1783-1786

296. Nauli, S., Farr, S., Lee, Y. J., Kim, H. Y., Faham, S., and Bowie, J. U. (2007) *Protein Sci* 16, 2542-2551

297. Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G., Thian, F. S., Kobilka, T. S., Choi, H. J., Kuhn, P., Weis, W. I., Kobilka, B. K., and Stevens, R. C. (2007) *Science* 318, 1258-1265

298. Jaakola, V. P., Griffith, M. T., Hanson, M. A., Cherezov, V., Chien, E.

www.manaraa.com

Y., Lane, J. R., Ijzerman, A. P., and Stevens, R. C. (2008) *Science* 322, 1211-1217

299. Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G., Thian, F. S., Kobilka, T. S., Choi, H. J., Yao, X. J., Weis, W. I., Stevens, R. C., and Kobilka, B. K. (2007) *Science* 318, 1266-1273

300. Braun, P., Hu, Y., Shen, B., Halleck, A., Koundinya, M., Harlow, E., and LaBaer, J. (2002) *Proc Natl Acad Sci U S A* 99, 2654-2659

301. Hammarstrom, M., Hellgren, N., van Den Berg, S., Berglund, H., and Hard, T. (2002) *Protein Sci* 11, 313-321

302. Shih, Y. P., Kung, W. M., Chen, J. C., Yeh, C. H., Wang, A. H., and Wang, T. F. (2002) *Protein Sci* 11, 1714-1719

303. Edwards, A. M., Arrowsmith, C. H., Christendat, D., Dharamsi, A., Friesen, J. D., Greenblatt, J. F., and Vedadi, M. (2000) *Nat Struct Biol* 7 Suppl, 970-972

304. Stevens, R. C. (2000) *Structure* 8, R177-185

305. Kapust, R. B., and Waugh, D. S. (1999) *Protein Sci* 8, 1668-1674

306. Sachdev, D., and Chirgwin, J. M. (2000) *Methods Enzymol* 326, 312-321

307. Grisshammer, R., Little, J., and Aharony, D. (1994) *Receptors Channels* 2, 295-302

308. Chen, G. Q., and Gouaux, J. E. (1996) *Protein Sci* 5, 456-467

309. Kanamori, M., Kamata, H., Yagisawa, H., and Hirata, H. (1999) *J*

www.manaraa.com

Biochem 125, 454-459

310. White, J. F., Trinh, L. B., Shiloach, J., and Grisshammer, R. (2004) *FEBS Lett* 564, 289-293

311. Derewenda, Z. S. (2004) *Structure* 12, 529-535

312. Derewenda, Z. S., and Vekilov, P. G. (2006) *Acta Crystallogr D Biol Crystallogr* 62, 116-124

313. Cooper, D. R., Boczek, T., Grelewska, K., Pinkowska, M., Sikorska, M., Zawadzki, M., and Derewenda, Z. (2007) *Acta Crystallogr D Biol Crystallogr* 63, 636-645

314. Moon, A. F., Mueller, G. A., Zhong, X., and Pedersen, L. C. (2010) *Protein Sci* 19, 901-913

315. Saul, F. A., Vulliez-le Normand, B., Lema, F., and Bentley, G. A. (1997) *Proteins* 27, 1-8

316. Wiltzius, J. J., Sievers, S. A., Sawaya, M. R., and Eisenberg, D. (2009) *Protein Sci* 18, 1521-1530

317. Mendillo, M. L., Putnam, C. D., and Kolodner, R. D. (2007) *J Biol Chem* 282, 16345-16354

318. Kim, M. W., Chelliah, Y., Kim, S. W., Otwinowski, Z., and Bezprozvanny, I. (2009) *Structure* 17, 1205-1212

319. Kawano, S., Yamano, K., Naoe, M., Momose, T., Terao, K., Nishikawa, S., Watanabe, N., and Endo, T. (2009) *Proc Natl Acad Sci U S A* 106, 14403-14407

www.manaraa.com

320. Gilbreth, R. N., Esaki, K., Koide, A., Sidhu, S. S., and Koide, S. (2008) *J Mol Biol* 381, 407-418

321. Potter, J. A., Randall, R. E., and Taylor, G. L. (2008) *BMC Struct Biol* 8, 11

322. Pioszak, A. A., Parker, N. R., Suino-Powell, K., and Xu, H. E. (2008) *J Biol Chem* 283, 32900-32912

323. Monne, M., Han, L., Schwend, T., Burendahl, S., and Jovine, L. (2008) *Nature* 456, 653-657

324. Chao, J. A., Prasad, G. S., White, S. A., Stout, C. D., and Williamson, J. R. (2003) *J Mol Biol* 326, 999-1004

325. Center, R. J., Kobe, B., Wilson, K. A., Teh, T., Howlett, G. J., Kemp, B. E., and Poumbourios, P. (1998) *Protein Sci* 7, 1612-1619

326. Laporte, S. L., Forsyth, C. M., Cunningham, B. C., Miercke, L. J., Akhavan, D., and Stroud, R. M. (2005) *Proc Natl Acad Sci U S A* 102, 1889-1894

327. Song, J. J., Liu, J., Tolia, N. H., Schneiderman, J., Smith, S. K., Martienssen, R. A., Hannon, G. J., and Joshua-Tor, L. (2003) *Nat Struct Biol* 10, 1026-1032

328. Pioszak, A. A., and Xu, H. E. (2008) *Proc Natl Acad Sci U S A* 105, 5034-5039

329. Kukimoto, M., Nureki, O., Shirouzu, M., Katada, T., Hirabayashi, Y., Sugiya, H., Furuyama, S., Yokoyama, S., and Hara-Yokoyama, M.

www.manaraa.com

(2000) *J Biochem* 127, 181-184

330. Bethea, H. N., Xu, D., Liu, J., and Pedersen, L. C. (2008) *Proc Natl Acad Sci U S A* 105, 18724-18729

331. Ullah, H., Scappini, E. L., Moon, A. F., Williams, L. V., Armstrong, D. L., and Pedersen, L. C. (2008) *Protein Sci* 17, 1771-1780

332. Huang, D. T., Hunt, H. W., Zhuang, M., Ohi, M. D., Holton, J. M., and Schulman, B. A. (2007) *Nature* 445, 394-398

www.manaraa.com

ABSTRACT

CRYSTALLOGRAPHIC, MOLECULAR DYNAMICS, AND ENZYMATIC STUDIES OF MULTI-DRUG RESISTANT HIV-1 PROTEASE AND IMPLICATIONS FOR STRUCTURE BASED DRUG DESIGN (PROJECT 1); CRYSTALLOGRAPHIC STUDIES OF HUMAN MYELIN PROTEIN ZERO (PROJECT 2)

by

ZHIGANG LIU

May 2011

Advisor: Dr. Ladislau C. Kovari

Major: Biochemistry and Molecular Biology

Degree: Doctor of Philosophy

PROJECT 1

Under drug selective pressure, emerging mutations render HIV-1 protease drug resistant, leading to antiretroviral therapy failure. The multidrug-resistant isolate 769 (MDR769) of HIV-1 protease, with resistance mutations at residues 10, 36, 46, 54, 62, 63, 71, 82, 84, 90, was selected for the present study to investigate the drug resistance issue.

Ten additional mutations were introduced into to MDR769 HIV-1 protease to study the structural influences introduced by these mutations. We obtained crystal structures of four variants (I10V, A82F, A82S and A82T) of MDR769 HIV-1 protease. All these mutations failed to further open the flaps and further expand the active site cavity of MDR769 HIV-1 protease, which was

www.manaraa.com

characterized by wide open flaps and expanded active site cacity. The conserved flaps and active site cavity, despite the introduction of additional mutations, indicated that the MDR769 HIV-1 protease represented the end stage form of HIV-1 protease. In addition, these crystal structures provided the first structure based evidence for the mutation induced conformational changes in the 80s loops of the HIV-1 protease apo-enzyme, although the flap and active site cavity were not changed dramatically. The alternate conformations of Pro81 (proline switch) in the I10V mutant and the side chain of Phe82 with flipped-out conformation in A82F mutant showed distorted S1/S1' binding pockets that caused loss of contacts and unstable binding of the inhibitors. Similarly, the mutants A82S and A82T showed distortion in the S1/S1' binding pockets due to local changes in the electrostatics caused by the mutation from non-polar to polar residues.

Molecular mechanics studies performed to understand the wide-open nature of the MDR769 HIV-1 protease flaps showed that the MDR protease exhibited a state of conformational rigidity with respect to the flap closure compared to that of the wild type protease. This suggested that the accumulation of mutations changed the structure of the MDR HIV-1 protease and resulted in a cumulative steric hindrance during the flap closure. Our studies showed that modeling a substrate (Gag – Capsid) into the active site cavity of the MDR protease did not result in flap closure. Since flap closure was crucial in protease inhibitor binding, the conformational rigidity of MDR

www.manaraa.com

protease might represent a novel mechanism for multidrug-resistance of the MDR protease. In addition, our molecular dynamics simulation revealed the realignment of the substrate peptide in the MDR769 HIV-1 protease, making it less accessible to the Asp25 and Asp125 amino acid residues in the active site. This finding indirectly indicated the reduced catalytic activity of MDR769 HIV-1 protease in substrate cleavage compared to that of the WT HIV-1 protease.

The IC50 values of the FDA approved HIV-1 protease inhibitors and the library of reduced CA/p2 peptide analogs were measured. The results indicated that the reduced peptide analogs bind to the MDR769 HIV-1 protease and WT HIV-1 protease with equal affinity, while the FDA approved inhibitors showed reduced binding affinity to the MDR769 HIV-1 protease compared to that of WT HIV-1 protease. The enzyme measurements demonstrated that lopinavir was the least resistant HIV-1 protease inhibitor and that reduced peptide P1'F was comparable to FDA approved inhibitors from the aspect of the inhibitory activity against HIV-1 protease.

Based on these studies, a library of potential drug candidates against HIV-1 protease was proposed to overcome the drug resistance issue. These drug candidates are being synthesized in the lab and will be evaluated in the future.

www.manaraa.com

PROJECT 2

The wild type extracellular domain of human myelin protein zero (hP0ex) fused with maltose binding protein (MBP) was crystallized to investigate the molecular mechanism of Charcot-Marie-Tooth disease subtype 1B. Based on the wild type structure of the extracellular domain of the human myelin protein zero, five clinically important mutants were further structurally investigated in details. The molecular pathology was proposed for these mutants individually. The relationship between amyloidosis disease and CMT1B is being further explored. From a technical point of view,the WT hP0ex MBP fusion structure was another example of crystallographic studies facilited by the presence of a purification and crystallization protein tag, the maltose binding protein.

www.manaraa.com

AUTOBIOGRAPHICAL STATEMENT

Education

- Ph.D. Major in Biochemistry (Sep 2005 to Dec. 2010)
  Department of Biochemistry and Molecular Biology, School of Medicine, Wayne State University, Detroit, MI 48201 USA

- B.M. Major in Clinical Medicine (Sep 2000 to July 2005)
  School of Medicine, Southeast Univerisy, Nanjing, Jiangsu 210009 China

Research and Work Experience

- **December 2005–March 2006, May 2006–Present** School of Medicine, Wayne State University, Worked in Dr. Kovari's lab focusing on structure-function study of HIV-1 protease and myelin protein zero
- **March 2006 – May 2006** School of Medicine, Wayne State University Worked in Dr. Rita Mukhopadhyay's lab focusing on the characterization of aquaporin in *Leishmania* species
- **July 2004 – July 2005** School of Medicine, Southeast University, Worked in Jiang Bei People's Hospital as a resident physician
- **February 2004 – July 2004** School of Medicine, Southeast University, Worked in Dr. Dou Jun's lab focusing on HCV and cancer gene vaccine
- **September 2001 – February 2004** School of Medicine, Southeast University, Worked in Dr. Li Xinrong's lab focusing on signal transduction of insulin receptor

Special Training Areas

- Four weeks hands on training in protein mass spectrometry – MALDI TOF in the Department of Pharmacology, Wayne State University
- Two weeks extensive training in macromolecular crystallography with special synchrotron data collection training at Illinois Institute of Technology – Sponsored by American Crystallographic Association (ACA)
- One year intensive training in Synthetic, Organic and Medicinal Chemistry Lab in the Eugene Applebaum College of Pharmacy and Health Sciences, Wayne State University

www.manaraa.com