Title
Studies of the HIV-1 protease through chemical synthesis

Author
Walker, Sharon Michele

Publication Date
1995

Peer reviewed|Thesis/dissertation

Studies of the HIV-1 protease through chemical synthesis

by

Sharon Michele Walker

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Pharmaceutical Chemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

Date **University Librarian**

Degree Conferred: . . . . . .

In memory of Lloyd Rehl (1954-1994), a friend lost to AIDS.

There are many people I would like to acknowledge for their unwavering support before and during my time in graduate school. I honestly believe I could not have either strarted or finished this work without their enouragement.

H.J. Hoffman at Lexington Christian Academy for establishing the foundation and encouraging my interest in chemistry.

Peter Trumper for his guidance, moral support and, especially, mentoring while at Bowdoin College. Thank you so much for believing in me and for being such a great teacher. I'll always be greatful to you for introducing me to the world of organic chemistry.

As for my UCSF connections, I'd like to thank the following people for their friendship and helping me maintain my sanity: Peter and Evelyne James, my fellow partners at BESIX Incorporated Bob Stull and Darren Wong, Carolyn Koo, Cindy and Roger Levinson, Barbara Frank, Nancy Phillips, Arnie Falick, Twink Stern, Sue Hawkes. A special thanks to my UCSF advisor, Tack Kuntz, for supporting me while I was in residence at The Scripps Research Institute.

A big heartfelt thanks to Steve Kent for giving me the opportunity to continue my thesis work in his lab at The Scripps Research Institute and for all his encouragement. It has been a great learning experience in so many ways. To the Kent lab crew: Manuel Baca, Lynne Canne, Phil Dawson, Michael Fitzgerald, Tilman Hackeng and Wuyuan Lu, Tom Muir for their helpful disscussions/suggestions and moral support. To Bernard Orsat,

Katharina and Jules Orsat, Mike Cappi, Amy and William Park, Milan Bruncko, Jarmila Brunckova, Lisa Wang--I feel so lucky to have had the opportunity to get to know you all. I can't begin to thank you all enough for your support!

A special thanks to Chuck McEwen at DuPont for taking an interest in my scientific career and for his constant encouragement.

Studies of the HIV-1 protease through chemical synthesis

Sharon M. Walker

The protease encoded by the human immunodeficiency virus type 1 (HIV-1 PR) has become an important target for the design of anti-AIDS therapeutics due to its essential role in viral maturation. Newly developed tools for the chemical synthesis of proteins have made it possible to construct a variety of HIV-1 PR analogues in a facile and straightforward manner. Autolysis-resistant forms of the HIV-1 PR monomer were chemically synthesized by both stepwise solid-phase peptide synthesis (SPPS) and the native chemical ligation of unprotected peptide segments. Synthesis of the HIV-1 PR monomer by native chemical ligation provided the enzyme consistently in high purity and yield. This method was also used for the rapid, modular construction of drug-resistant forms of the HIV-1 PR.

Drug-resistant forms of the HIV-1 PR, identified as arising from the most advanced PR inhibitor Ro31-8959, were constructed. Kinetic parameters were measured for the enzyme variants, along with inhibition constants (Ki) against two representative examples of canonical classes of inhibitors. In addition, the properties of a pair of inhibitors developed by comparative studies against both HIV-1 and feline immunodeficiency virus (FIV) proteases were evaluated.

Rational changes in the HIV-1 PR were made to explore the structural basis of substrate specificity. Geometrically-constrained amino acids were

used to systematically alter the S1 and S1' binding pockets of the enzyme. The objective was to design changes in the enzyme with predictable consequences. 4-Cis-(**X**)-proline derivatives, where **X** is either guanidino or guanidinomethyl, were synthesized and incorporated in positions 81 and 181 of the HIV-1 PR homodimer. These chemically synthesized analogues of the HIV-1 PR were examined kinetically and for substrate specificity. It was found that these enzyme analogues accommodated the mutations and, unexpectedly, recognized the same substrates.

The results of this work help establish a foundation of knowledge concerning the structural basis of the specificity and catalytic activity of HIV-1 PR. This will aid in refining the design of HIV-1 PR inhibitors for the development of improved AIDS therapeutics.

Table of Contents

List of Tables

List of Figures

Chapter I

Introduction

The first cases of Acquired Immunodeficiency Syndrome (AIDS), a crippling disease of the immune system, were reported in the early 1980's to the Centers for Disease Control (CDC). In the decade that has passed, we have continued to watch the epidemic grow to the point where the total number of AIDS cases reported worldwide to the World Health Organiation (WHO) Global Programme on AIDS totals greater than 1 million. However, due to under-diagnosis, incomplete reporting and delays in reporting, it is estimated that the number may actually be closer to 4.5 million[1]. In addition, the number of people infected with the human immunodeficiency virus (HIV), the virus that causes AIDS, is estimated to be greater than 18 million. In the US alone, the total number of AIDS cases reported to the CDC through the end of 1994 totalled over 440,000[2]. By the end of the century, WHO estimates that between 30-40 million men, women and children worldwide will have been infected with HIV. Alarmingly, the HIV epidemic continues to grow at an estimated rate of 6000 new infections each day. These numbers are staggering, especially then thinking about the vast numbers of people who will be affected not only by the disease itself, but by the loss of friends and family to the disease.

AIDS has become a serious threat to our global health and ecomomy. The disease continues to be one of the leading causes of death for adults between 25-45, the age group on which our society is most dependent for productivity. Although AIDS continues to disproportionately affect homosexual men and intravenous drug users, the disease is not solely confined to these groups. AIDS is increasingly affecting the whole of the population, regardless of age, gender, race, religion, sexual orientation or socioeconomic background. Despite huge efforts by scientists to understand

the mechanism of HIV, there is currently no effective treatment or cure for AIDS.

1.1 The human immunodeficiency virus as the causative agent of AIDS

The general consensus among the scientific community is that HIV is the causative agent of AIDS. However, there are some who believe that infection with HIV has little or nothing to do with AIDS. Increasing evidence continues to appear which provides sufficient proof that HIV infection leads to AIDS. In a recent issue of the journal Nature, a study appeared which looked at the death rates among British haemophiliacs with and without HIV infection. Darby et al.[3] showed that the death rate among those infected with HIV is approximately 10 times greater than those not infected. In addition, they estimate that 85% of the deaths among infected patients were caused by HIV. The data in this study significantly strengthens the argument that there is link between HIV infection and AIDS.

HIV is transmitted by sexual contact, blood or blood products, or from an infected woman to her unborn or newborn child. T4 lymphocytes are the primary target of HIV infection, as well as macrophages and monocytes. A distinctive characteristic of AIDS is the reduction in the number of T lymphocytes known as CD4s, which orchestrate the immune system. Those infected tend to die from the opportunistic infections acquired because the immune system is impaired. Also, a long latency period exists between infection (i.e., integration of the virus into the host cell's genome) and the onset of disease.

HIV is a retrovirus of the *Lentiviridae* family, whose genome is encoded in RNA. Originally referred to as Human T-cell leukemia virus III (HTLV-III) or lymphadenopathy-associated virus (LAV), it is now designated HIV[4,5]. There are two distinct subtypes of virus, HIV-1[6-8] (which is the most widespread and of greater consequence) and HIV-2[9] and their genomes show about 50% sequence identity at the amino acid level[10]. HIV-2, however, is more closely related to the simian immunodeficiency virus (SIVmac)[11,12], which has beens shown to cause an AIDS-like disease in captive Asian macaques[13]. HIV-1 exhibits the prototypal 5'-*gag-pol-env*-3' genomic organization typical of other retroviruses and, additionally, encodes a number of proteins with regulatory or unknown function[6,7,14-16] (Figure 1.1). There exist many different viral strains within the catagories of HIV-1 and -2. This is likely due to the fact that the reverse transcriptases of retroviruses, enzymes responsible for transcribing viral RNA to DNA, are highly prone to base incorporation error with an error rate of about one base/1,700[17,18] and have no DNA repair activity. Thus, it is not surprising that a great deal of diversity exists between viral isolates of both HIV-1 and -2. In addition, this diversity could also provide a mechanism for the emergence of drug-resistant variants of HIV.

1.2 The viral life cycle

The life cycle of HIV[19] (schematically represented in Figure 1.2) can be broken down into two phases: pre-and post-integration stages. In the pre-integration stage, a cell becomes infected with HIV when the glycoprotein portion of the viral envelope protein gp120 binds to its cellular receptor protein CD4. The viral and cellular membranes fuse and expel the virion

Figure 1.1. Genomic Organization of the Murine Leukemia and the Human and Simian Immunodeficiency Viruses[16].

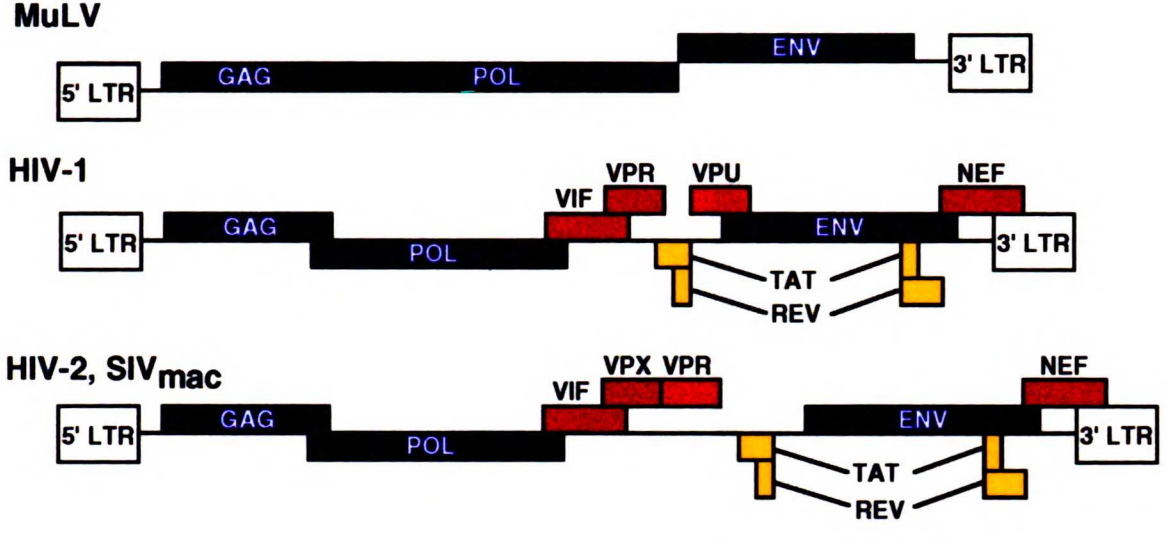
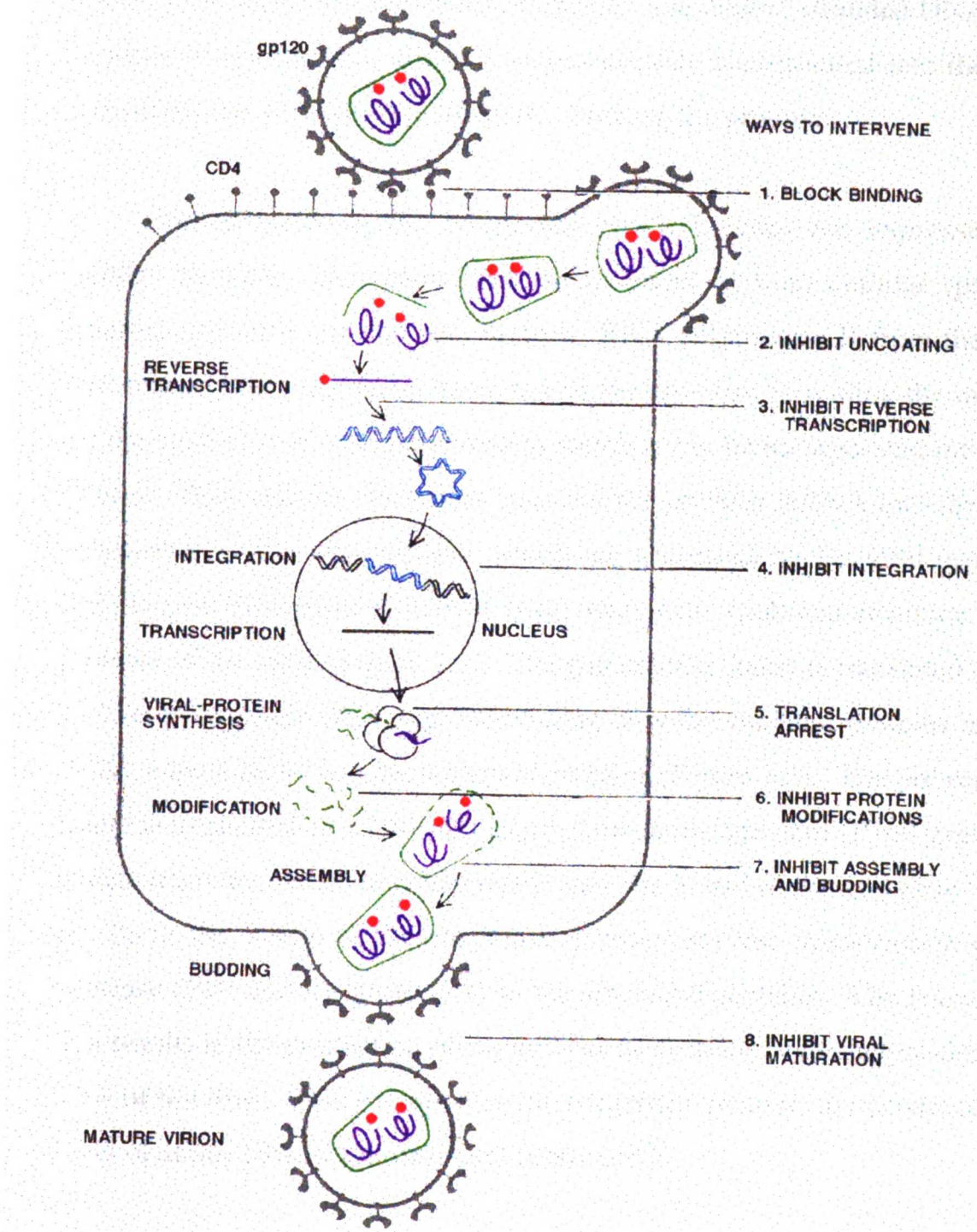


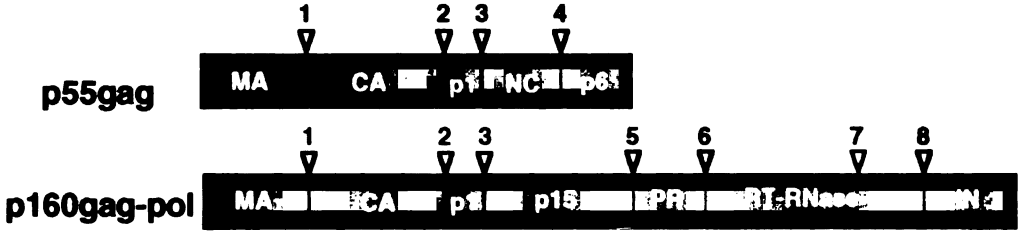
Figure 1.2 The life cycle of HIV within a human T4 lymphocyte (adapted from Yarchoan et al.[19]).


gp120
CD4
WAYS TO INTERVENE
1. BLOCK BINDING
2. INHIBIT UNCOATING
REVERSE TRANSCRIPTION
3. INHIBIT REVERSE TRANSCRIPTION
INTEGRATION
4. INHIBIT INTEGRATION
TRANSCRIPTION
NUCLEUS
VIRAL-PROTEIN SYNTHESIS
5. TRANSLATION ARREST
MODIFICATION
6. INHIBIT PROTEIN MODIFICATIONS
ASSEMBLY
7. INHIBIT ASSEMBLY AND BUDDING
BUDDING
8. INHIBIT VIRAL MATURATION
MATURE VIRION

into the cell. Uncoating of the virion releases its contents (viral RNA, nucleocapsid proteins, reverse transcriptase, protease and integrase) into the host cell's cytoplasm. Next, viral RNA is transcribed into double-stranded DNA by the action of reverse transcriptase. The double-stranded DNA then migrates to the nucleus, where it is permanently incorporated into the host genome by the viral enzyme integrase, forming the provirus.

Post-integration, the integrated proviral genes are expressed by cellular enzymes and require the involvement of both cellular (protein kinases) and viral (tat and rev) factors. RNA polymerase II from the host then makes a primary RNA transcript of the provirus, providing the mRNA for the synthesis of viral polyproteins which are to be incorporated into new virions. Host cellular ribosomes translate the proviral mRNA into the viral polyproteins--gag, and gag/pol--which are processed by the viral protease (PR) to give the mature structural (gag), enzymatic (pol) and envelope (env) proteins of the virus (Figure 1.1). The gag reading frame is translated into a 55 kDa polyprotein (P55 gag), containing only the stuctural proteins matrix (MA), capsid (CA) and nucleocapsid proteins (Figure 1.3). The pol reading frame is expressed as a 160 kDa gag/pol fusion polyprotein (P160 gag-pol), which arises by means of a frameshift near the 3' end of the gag gene[20], and results in the fusion of the enzymatic proteins reverse transcriptase (RT), protease (PR) and integrase (IN) to the structural proteins. The frameshift that results in the production of gag/pol fusion proteins occurs approximately 10% of the time. This ensures that the enzymatic proteins are produced at a level of about 10% of the structural proteins[20-22].

Figure 1.3 HIV-1 protease cleavage sites within HIV-1 gag and gag-pol polyproteins. The mature proteins released by proteolytic cleavage are the following: matrix (MA); capsid (CA); nucleocapsid (NC); p6; protease (PR); reverse transcriptase p51 (RT51); reverse transcriptase p66 (RT--contains RNase H domain); integrase (IN).



		P4	P3	P2	P1	**	P1'	P2'	P3'	P4'
1	MA/CA	Ser	Gln	Asn	Tyr		Pro	Ile	Val	Gln
2	CA/p1	Ala	Arg	Val	Leu		Ala	Glu	Ala	Met
3	p1/NC	Ala	Thr	Ile	Met		Met	Gln	Arg	Gly
4	NC/p6	Pro	Gly	Asn	Phe		Leu	Gln	Ser	Arg
5	/PR	Ser	Phe	Asn	Phe		Pro	Gln	Ile	Thr
6	PR/RT	Thr	Leu	Asn	Phe		Pro	Ile	Ser	Pro
7	RT51/RNase H	Ala	Glu	Thr	Phe		Tyr	Val	Asp	Gly
8	RT/IN	Arg	Lys	Ile	Leu		Phe	Leu	Asp	Gly

Essential post-translational modifications of the polyproteins, such as the glycosylation of env and myristoylation of the N-termini of p17 in the gag and gag/pol polyproteins, are performed by cellular enzymes. Myristoylation of the viral polyproteins is essential for their proper assembly into virion particles. After all post-translational modifications have occurred, the polyproteins and genomic RNA are driven to the plasma membrane. The packaging or "budding" of the immature viral particle is the next step in this process. Proteolytic processing of the gag and gag/pol polyproteins by the virally encoded protease results in the formation of a mature, infectious virion via the release and activation of the structural proteins and enzymes. The virion is now ready to begin a new cycle of infection.

1.3 Interfering with the viral life cycle as a route to drug therapy

As noted in Figure 1.2, the HIV life cycle consists many steps, any one of which may be exploited for chemotherapeutic intervention[23,24]. The virus-specific enzymes, reverse transcrpitase/RNase H, integrase and protease, are particularly attractive candidates since it has been well established by medicinal chemists that enzymes are good targets for drug design[25]. Some representative examples are the use of methotrexate as an inhibitor of the enzyme dihydrofolate reductase for cancer therapy; captopril, which acts as an antihypertensive by inhibiting angiotensin-converting enzyme[26]; and familiar over-the-counter products such as aspirin, ibuprofen and naproxen, which serve as anti-inflammatory drugs via inhibition of prostaglandin H2 synthase-1[27,28]. It is important to emphasize that any antiviral chemotherapeutic intervention must be highly specific to the viral

target and not interfere with other cellular processes. In the case of HIV, the inherent differences between the viral and host enzymes must be exploited in designing inhibitors. Nucleoside inhibitors to reverse transcrpitase have had some degree of clinical success in the treatment of AIDS. AZT, ddI, ddC and d4T are the only anti-HIV drugs currently approved for use in the United States[29]. However, these first generation drugs have high levels of toxicity and the development of resistance to these drugs rapidly diminishes their effectiveness as antivirals[30,31]. In addition, these drugs serve as only a "temporary fix" in that they do not provide a cure for AIDS, but only slow the progression of the disease. Therefore, other more effective alternatives must be explored.

1.4 HIV-1 protease--a target for drug design

One of the most well studied of the viral enzymes at this time is the HIV-1 protease (HIV-1 PR). It has been demonstrated that the protease plays critical role in the viral lifecycle in that it is absolutely necessary for the production and replication of infectious virions. Studies performed in a number of laboratories have shown that budded, immature viral particles which contain a catalytically inactive or an inhibited form of the protease cannot mature into an infectious form[32-36]. It is also believed that the protease may be involved in the hydrolysis of the nucleocaspid proteins which surround the viral DNA early in the viral lifecycle[26,37]. Inhibition of the HIV-1 PR at this early stage in the viral life-cycle may prevent proviral integration. Therefore, if it were possible to **specifically** inhibit the viral protease, it would be possible to prevent production of infectious virus. However, before beginning to conceive of creating novel chemotherapeutics

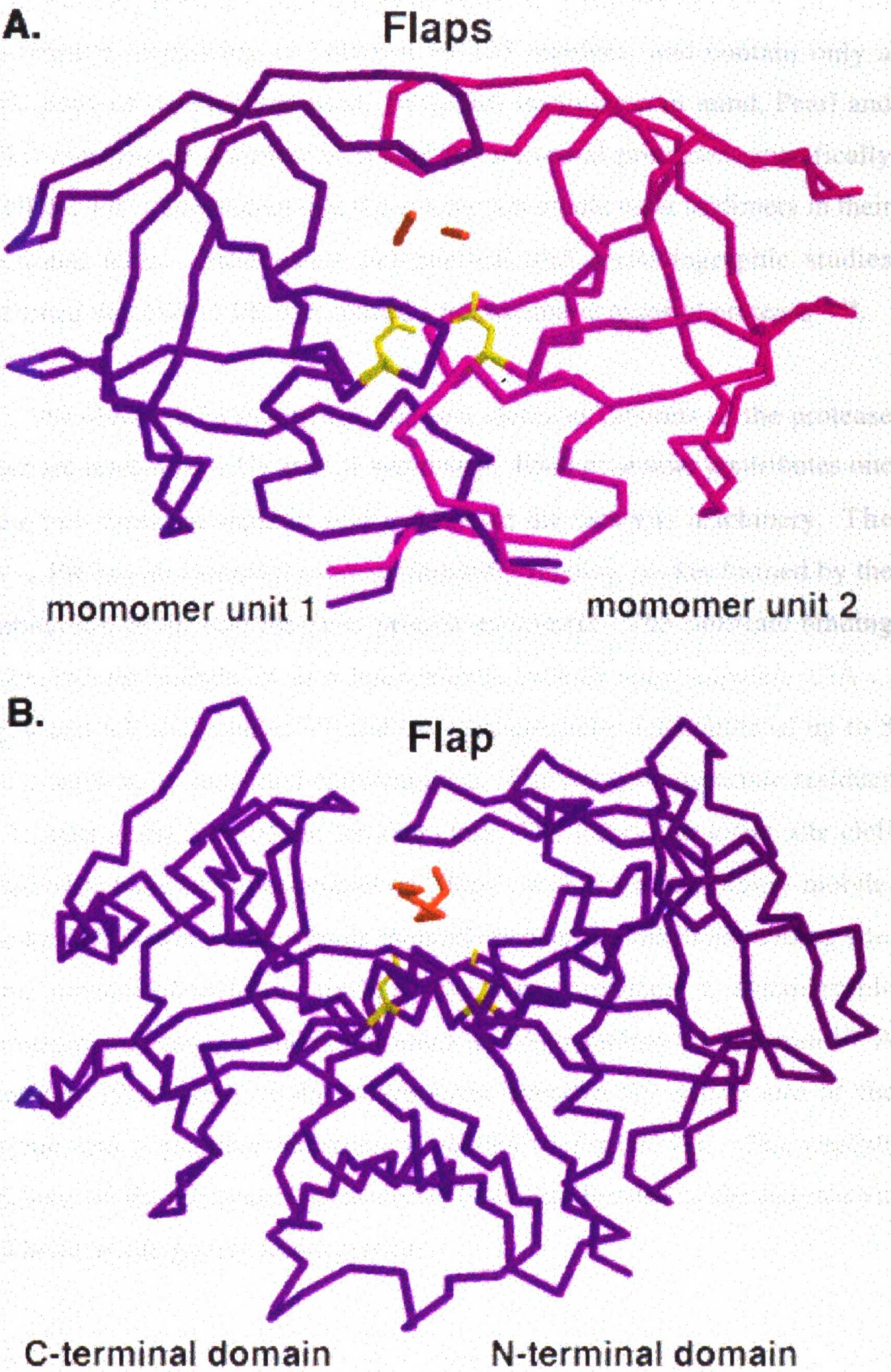
for AIDS by the inhibition of HIV-1 PR, it is necessary to have a reasonable understanding of the enzyme's biochemistry, mechanism and structure. Many important advances have been made toward this end, providing some much needed insights into the structure and function of HIV-1 PR .

1.5 The HIV-1 PR is a dimeric aspartyl protease

A great deal of information has been inferred about the mechanism and structure of the HIV-1 PR based on what was already known from studies of related enzymes, the eukaryotic aspartyl proteases. This extensively characterized family of enzymes includes such members as cathepsin D, pepsin, and renin. HIV-1 PR and other retroviral proteases were suggested to be aspartyl proteases based on sequence conservation of residues in the active site. In addition, Toh and coworkers observed that the characteristic Asp-Thr(Ser)-Gly active site triad present in eukaryotic apartyl proteases was also present in retroviral proteases and concluded that these enzymes must be members of the aspartyl protease family[38].

There are, however, some significant structural differences which exist between the retroviral and "classical" aspartyl proteases. The cell-derived, pepsin-like proteases are typically single-chain, two domain monomers of about 325 amino acids containing two active site aspartates as shown in Figure 1.4[39]. Tang observed that the aspartyl protease pepsin was composed of two homologous halves and proposed that it was derived by a process of gene duplication and fusion from a dimeric ancestral protein. As other members of the pepsin-like aspartyl protease family were characterized, it was also noted that they were of similar structure.

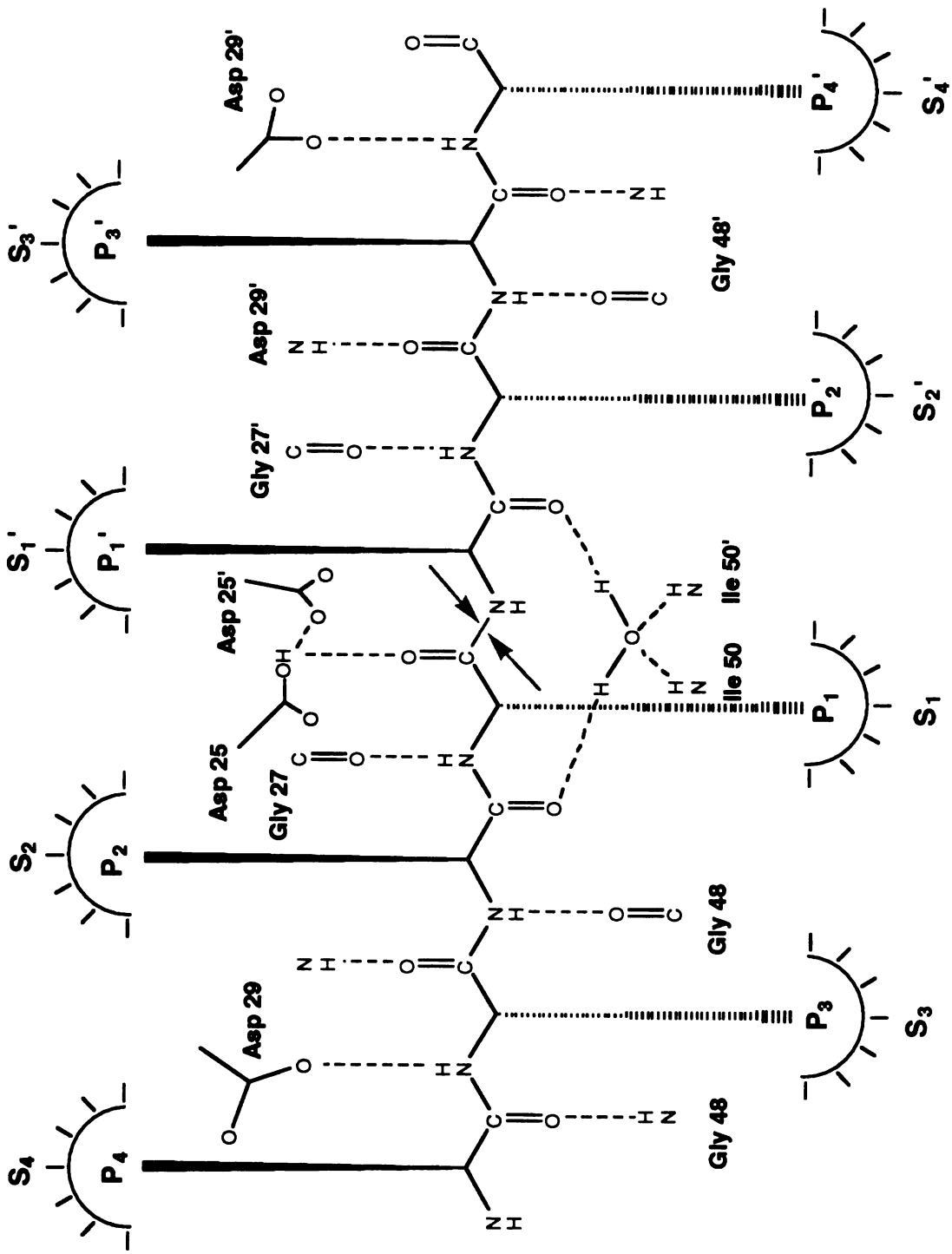
Figure 1.4 Comparison of Cα backbone structures of HIV-1 PR and the cell-encoded aspartic protease, endothiapepsin. (A). Structure of the HIV-1 PR in the presence of the substrate-based inhibitor MVT-101[44]. The figure shows the two monomeric subunits and flaps which form the roof of the active site. (B). The structure of endothiapepsin shows the two domains and a single flap[98]. The catalytic aspartate residues are displayed in yellow in both structures. The structural homology between these two types of aspartic proteases is evident by comparing one of the monomer subunits of HIV-1 PR with the N-terminal domain of endothiapepsin. The figure was prepared using the program RasMol v.2.6 (Roger Sayle, Biomolecular Structure, Glaxo Research and Development, Middlesex, UK.).


A.
Flaps
momomer unit 1
momomer unit 2
B.
Flap
C-terminal domain
N-terminal domain

However, the retroviral proteases are about half the size of their cellular counterparts, consisting of between 99-125 residues, and contain only a single copy of the catalytic triad. With this informaton in mind, Pearl and Taylor constructed a structural model for retroviral proteases, specifically the HIV-1 PR[40], suggesting that these enzymes would exist as dimers in their functional form. Subsequent biochemical and crystallographic studies confirmed that HIV-1 PR was indeed a homodimeric aspartyl protease[41-44].

As observed in Figure 1.4, the two identical subunits of the protease dimer are related by a C_2 axis of symmetry. Each monomer contributes one of the two essential aspartic acid residues to the catalytic machinery. The HIV-1 PR has an extended substrate/inhibitor binding pocket formed by the combination of the two identical protein monomers. The substrate binding pocket can be thought of as a hydrophobic tube of approximately 25Å in length and 8Å in diameter[45,46] and can accommodate a substrate of up to 8 amino acids in an extended conformation. The catalytic aspartate residues are located at the base of the binding cleft. The top of the active site cleft forms what have been designated as "flaps", which are two highly-mobile, glycine-rich β-strands that situate themselves over the substrate binding site. Upon substrate/inhibitor binding, the flaps undergo a considerable conformational change, moving as much as 7Å to enclose the substrate[47]. A schematic illustration of the interactions between the active site of the enzyme and a putative substrate is shown in Figure 1.5. The peptide backbone of the substrate is extensively hydrogen bonded to the enzyme via backbone amide groups in the protein.

Figure 1.5 Hydrogen bonding interactions between HIV-1 PR and substrate, as deduced from cocrystal structures of HIV-1 PR bound to substrate-based inhibitors. The scissile bond is designated with arrows. Symbols Sn-Sn' denote the enzyme subsites which accommodate the substrate side chains. (Figure taken from Wlodawer and Erickson[69]).


S4
P4
S3
P3
S2
P2
S1
P1
S1'
P1'
S2'
P2'
S3'
P3'
S4'
P4'
Asp 29
Gly 48
Asp 25
Gly 27
Asp 25'
Gly 27'
Asp 29'
Gly 48'
Ile 50
Ile 50'

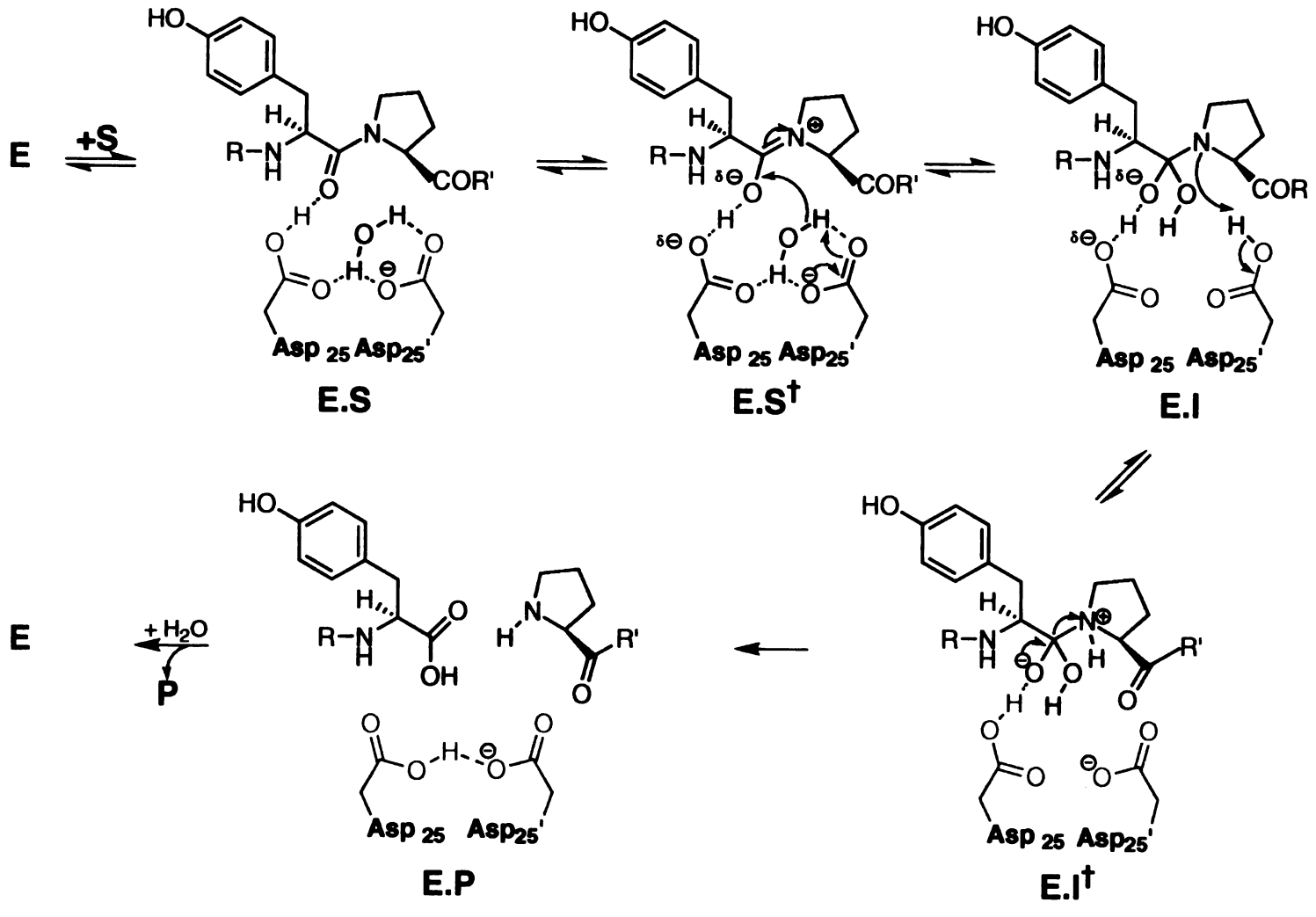
1.6 Chemical mechanism of the HIV-1 PR

Our present understanding of the chemical mechanism of HIV-1 PR comes by analogy with the pepsin-like proteases and is supported by the scant available kinetic and structural data[48-51]. Figure 1.6 schematically illustrates the chemical mechanism as adapted from Meek and coworkers[51]. As indicated in Figure 1.6, one essential feature in the chemical mechanism of the HIV-1 PR is that there is no enzyme-substrate covalent intermediate as with the cysteine and serine proteases. The two active site aspartates act in concert playing general acid-general base roles, where one is in an unprotonated form (pK_1 = 3.4-3.7) and the other protonated (pK_2 = 5.5-6.5). These values are in accord with the low pH optimum for protease activity, which ranges between pH = 4.5-6. Water serves as the attacking nucleophile, and attack of the amide carbonyl at the scissile bond of the substrate results in the formation of an enzyme-bound, non-covalent, tetrahedral amide hydrate intermediate. Subsequent collapse of the hydrate intermediate and peptide bond cleavage results in product release.

1.7 HIV-1 PR substrate specificity

Much of what is understood regarding the substrate specificity of the HIV-1 PR is based on the analysis of cleavage sites in the viral protein substrates. The HIV-1 PR cleaves eight peptide bonds in the P55 gag and P160 gag-pol polyproteins[52] (Figure 1.3), suggesting that the HIV-1 PR is a highly specific protease. Based on this analysis, substrates of the HIV-1 PR have been divided into two or three classes[53-55]. However, few consider more than two substrate classes, those where the amino acids in the P1-P1'[§]

Figure 1.6 Proposed chemical mechanism for the HIV-1 protease-catalyzed peptidolysis of a -Tyr * Pro- containing peptide substrate. This mechanism is based on the cumulative work of Meek and coworkers[48-51].


E
+S
HO
R–N
H
COR'
Asp 25 Asp25'
E.S
E.S†
E.I
E.I†
E.P
+ H2O
P
OH
R'
COR

positions are Tyr/Phe*Pro (Class 1--cleavage sites 1, 5, and 6 in Figure 1.3) or Hydrophobic*Hydrophobic (Class 2--cleavage sites 2-4, 7, and 8 in Figure 1.3) [55]. In Class 1 substrates, glutamate is highly favored in the P2' position. In addition, this class of substrates contains an unusual cleavage site for most endopeptidases, since cleavage occurs N-terminal of a proline residue. This feature distinguishes the HIV-1 PR and other retroviral proteases from most known endopeptidases. Class 2 substrates typically contain large, unbranced amino acids at the P1 position and have a preference for β-branched residues (i.e., isoleucine, valine and threonine) and glutamate/glutamine at P2 and P2', respectively. Oligopeptide substrates derived from natural cleavage sites, which consist of six or more amino acids, have been shown to be suitable substrates for the HIV-1 PR.

In addition, nonviral proteins have also been evaluated as substrates of the HIV-1 PR in an attempt to extend the data set of possible cleavage sites beyond those observed in the viral substrates[56-64]. The overall conclusions which can be drawn from the sum of these studies are that the specificity of HIV-1 PR is determined by the cumulative interactions of the substrate/inhibitor with the enzyme subsites. (The enzyme interacts with amino acid side chains from P4 through P4'[65] in its substrate. Refer to Figure 1.5.) Thus, it appears that more has been learned about which amino acids are not allowed in certain Pn or Pn' postions than being able to accurately predict a which bond in a protein will be succeptible to cleavage by the HIV-1 PR.

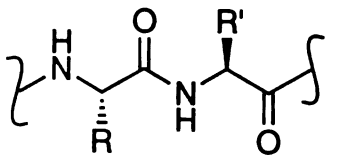
§Pn, Pn' notation of Schecter and Berger[65] designates the position of residues relative to the scissile bond in the peptide substrate. The unprimed notation (P1, P2...Pn) designates residues on the N-terminal side of the substrate and the primed notation (P1', P2', Pn') designates residues on the C-terminal side. The corresponding subsites in the protein are designated Sn.

1.8 Substrate- vs. structure-based design of HIV-1 PR inhibitors

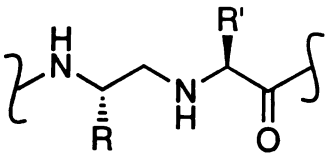
Armed with the knowledge of HIV-1 PR specificity and the fact that this enzyme is an aspartyl protease, it became possible to investigate the design of HIV-1 PR inhibitors as potential therapeutics. Numerous reviews have been written on the subject of HIV-1 PR inhibitors[26,66-71], and a great deal of the literature published on the HIV-1 PR deals with the study of its inhibition. This section will deal with inhibitors designed via substrate- or structure-based design methods.

1.8.1 Substrate-based inhibitors--important lessons learned from a predecessor.

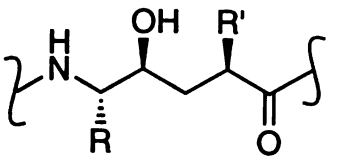
With regard to the development of assays and the design of inhibitors of HIV-1 PR, it was of great advantage that much was already known about another aspartyl protease which plays a critical role in blood pressure regulation, human renin. Many academic and industrial laboratories had been investigating the design of inhibitors of human renin for use as possible antihypertensive agents[72]. Several different classes of renin inhibitors have been developed, some of which also turned out to be inhibitors of HIV-1 PR[73]. What arose from this body of information was the development of substrate-based inhibitors, where nonhydrolyzable dipeptide isosteres which mimic the tetrahedral transition state were substituted at postitions P1-P1' into short peptide sequences corresponding to known proteolytic cleavage sites of HIV-1 PR. Some of these "transition state analogues" turned out to be good inhibitors of the HIV-1 PR, with Ki values usually in the low micromolar-nanomolar range. Figure 1.7 shows examples of the various



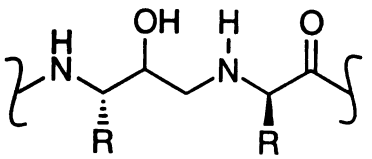
Peptide Bond



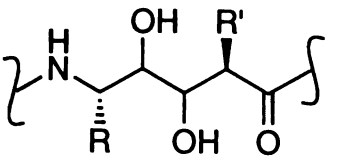
Reduced Amide



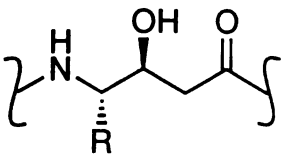
Hydroxyethylene



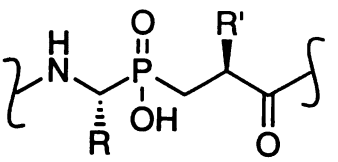
Hydroxyethylamine



Dihydroxyethylene



Statine



Phosphinate

Figure 1.7 Nonhydrolyzable dipeptide isostere replacements for the scissile bond in a protease substrate.

types of dipeptide isosteres which have been incorporated into HIV-1 PR inhibitors. Upon comparison of the inhibitory effects of different isosteres, it appeared that the hydroxyl-containing isosteres demonstrated a higher level of inhibition than those without. This observation has been attributed to the fact that the hydroxyl-containing isosteres better mimicked the proposed substrate tetrahedral intermediate of the cleavage reaction. The hydroxyethylamine inhibitor is one such example of a hydroxyl-containing isostere. However, this particular isostere is unique in that it contains an extra backbone atom (a methylene group--see Figure 1.7). When this dipeptide isostere is substitiuted at the scissile bond of a peptide substrate, very potent inhibitors in the subnanomolar range are usually produced.

An example of a potent hydroxyethylamine-containing inhibitor is the compund Ro 31-8959[74], developed by Roche Products Ltd (Figure 1.8). This compound's design was based on Class 1 substrates (Phe/Tyr-Pro), with the notion that this class of inhibitors would have greater selectivity for the viral protease over mammalian endopeptidases. This does indeed appear to be the case, for while the compound exhibits inhibitory activity against HIV-1 and -2, with Ki's of about 10^{-10}M, at 10^{-5}M it has little effect on the human apsartyl proteases renin, pepsin, gastrin, cathepsin D and cathepsin E[74,75]. Ro 31-8959 is primarily peptidomimetic in nature, although some changes were made in the overall structure (with the addition of either large, hydrophobic groups and/or unnatural amino acid side chains) in order to optimize binding at each susbsite in the inhibitor. Interestingly, this compound was in advanced clinical trials[76] and has just been approved by the FDA for use as a therapeutic.

1.8.2 Structure-based inhibitor design

However, as observed with the development of inhibitors of human renin, it became necessary to get away from the peptidic scaffold inherent in these compounds in order to obtain the needed bioavailablilty and pharmacokinetic properties necessary for good therapeutics. Thus, other viable approaches to the creation of inhibitors needed to be found. A greater divergence from peptide structure could likely result by the design of inhibitors based directly on knowledge of the three-dimensional structure of an enzyme. Access to information provided by the X-ray crystallographic structures of HIV-1 PR has been of great use in the design of inhibitors. The structure-based design of HIV-1 PR inhibitors has been a rapdily expanding field, as evidenced by the sheer quantity of available crystal structures of the PR. More than 300 structures of the HIV-1 PR (both with and without inhibtors) have been generated since 1989[66,69]. Knowledge of the tertiary structure of the HIV-1 PR has allowed for the development of what have come to be termed "structure-based" inhibitors. Examples of the inhibitors which can be considered in this catagory are pseudosymmetric and symmetric inhibitors[77,78], and those designed with shape complementarity to the active site[45,46].

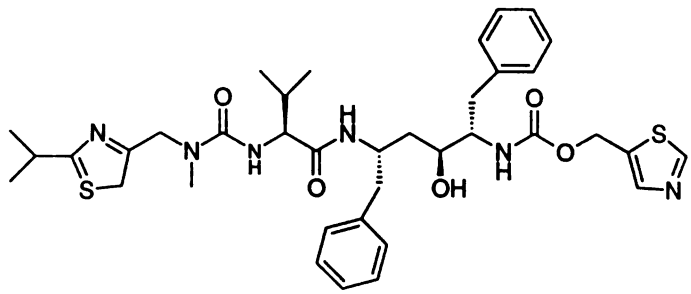
Symmetry-based inhibitors were designed based on the knowledge that the HIV-1 PR functions as a C2 symmetric homodimer. Of particular interest is the inhibitor ABT-538[77], derived using standard medicinal chemical approaches from A-77003[79], which is a highly potent inhibitor that was designed to have an improved pharmacokinetic profile when compared to its predecessor (Figure 1.8). Another symmetric inhibitor of particular

interest is that developed by the DuPont-Merck group. The inhibitor DMP-323, shown in Figure 1.9, was designed on mechanistic information available based on knowledge of HIV-1 PR-inhibitor crystal structures[78]. One fundamental feature is the cyclic urea carbonyl oxygen that mimics the hydrogen-bonding features of a key structural water molecule, designated as water 301. This tetrahedrally-coordinated water molecule links the bound inhibitor to the flexible glycine-rich beta strands or "flaps" of the HIV PR dimer, accepting two hydrogen bonds from the backbone amide hydrogens of HIV-PR residues Ile 50 and Ile 50' and donates two hydrogen binds to carbonyl oxygens of the inhibitor, inducing the fit of the flaps over the inhibitor. In designing this inhibitor, the group at DuPont-Merck speculated that incorporation of a mimic for water 301 within the inhibitor would imart a particular specificity for the HIV-1 PR over the cell-encoded aspartyl proteases since this water molecule has only been observed in the crystal structure of the HIV proteases.

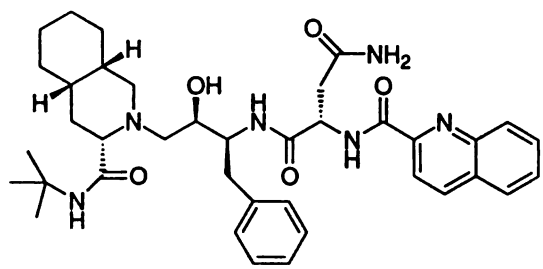
DesJarlais and coworkers used the the computer program DOCK[80] to characterize the shape of the HIV-1 PR's substrate binding pocket. From this analysis, potential inhibitors were selected based on their shape and electrostatic complementarity in the active site. This analysis resulted in the identification of haloperidol, an antipsychotic drug, as a potential "hit" (Figure 1.9). Indeed, haloperidol did inhibit the HIV-1 PR, but weakly with a Ki in the micromolar range. With the identification of haloperidol as a lead compound, analogues were synthesized in order to achieve a greater level of inhbition. One such analogue, designated UCSF-8 (Figure 1.9) possessed an 8-fold increase in affinity for the protease[46]. Although the level of inhibition achieved with these compounds was still not sufficient to make

Figure 1.8 Inhibitors of the HIV-1 PR presently in clinical trials.

RITONAVIR
ABT-538
A-84538



SAQUINAVIR
RO 31-8959



MK-639
L-735,524

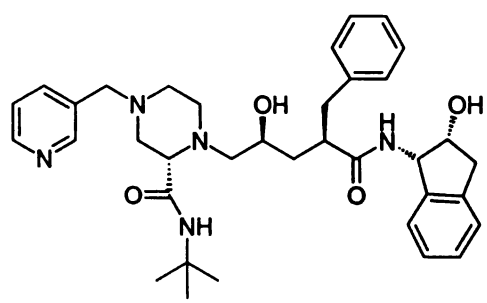
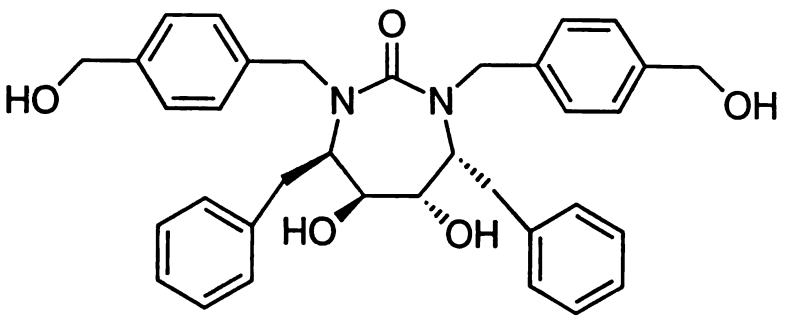
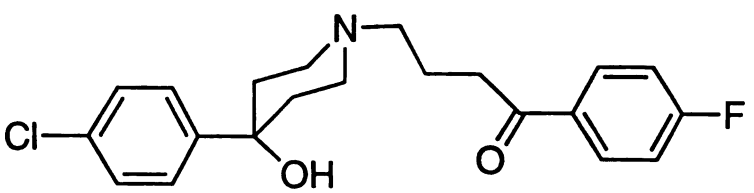


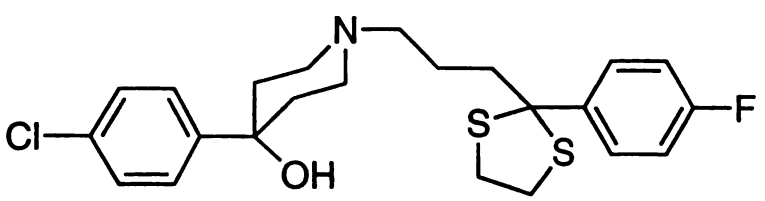
Figure 1.9 Examples of structure-based inhibitors of HIV-1 PR.



DMP-323



Haloperidol



UCSF-8

a useful AIDS drug, it is a good beginning in the area of "rational drug design" efforts.

Many of the HIV-1 PR inhibitiors which have been developed in the last few years, using either substrate- or structure-based methods, have shown potent anti-viral activity and are presently being tested in clinical trials. At the present time, Abbott Laboratories[77], Hoffmann-La Roche[74] and Merck & Co.[81] each have protease inhibitors in phase III clinical trials to test the efficacy of these compounds[29]. The structures of these compounds are presented in Figure 1.8.

1.9 A chemical synthesis approach to the study of HIV-1 PR

Alternative approaches were needed in order to acquire adequate HIV-1 PR for biochemical studies. The enzyme is present in such small quantities in the virus that it is impractical to try to isolate it directly from virus particles. In addition, it is quite difficult and potentially dangerous to isolate the enzyme directly from virus particles[82]. A straightforward method of obtaining a safe and sufficient source of PR was necessary in order have access to milligram quantities for biochemical and structural characterization. Recombinant DNA technology has been used sucessfully to obtain the HIV-1 PR in sufficient quantities for biochemical and structural studies. The small size of the HIV-1 PR monomer polypeptide chain has also made this enzyme amenable to total chemical synthesis. In using a chemical synthesis approach, biochemical and mechanistic studies of the HIV-1 PR which require mutagenesis can be performed without limitation to the number or types of changes which can be made in the protein[83-86].

The field of solid-phase peptide chemistry has been continually evolving, as indicated by the many important advances which have been made in the areas of stepwise solid-phase peptide synthesis (SPPS). These advances were necessitated by the desire to chemically synthesize proteins of biological interest and has subsequently made it possible to construct native and uniquely modified proteins. Synthetic methods that proved effective for the synthesis of small peptides could not readily be adapted to the assembly of proteins. Incomplete reaction at each cycle in the chain assembly results in the accumulation of low-level byproducts. The complex mixture obtained upon resin cleavage made isolating the product of interest difficult at best [84]. Improvements in the chemistry of the chain assembly and the analytical techniques now available, for example, the use of mass spectrometry[87], have made it possible to overcome the problems associated with the difficulty in synthesizing long peptides in high purity and yields[88,89].

1.9.1 Total chemical synthesis of HIV-1 PR--the early days

Chemical synthesis has been of great use in the study of HIV-1 PR. The first chemical synthesis of HIV-1 PR was described by Schneider and Kent[90], where highly-optimized stepwise Boc chemistry was used to obtain milligram quantities of enzyme[88]. This material was ultimately used to obtain the first correct crystal structures of the HIV-1 PR in the apoenzyme[43] and inhibited forms[47], confirming the structural model proposed by Pearl and Taylor that the enzyme is indeed a homodimeric aspartyl protease. This structural information laid the foundation for much of the progress in the structure-based drug design of HIV-1 PR inhibitors and the determination of the chemical mechanism. Subsequent syntheses of the HIV-1 PR

incorporated the use of α-aminobutyric acid (Aba), an isosteric replacement for cysteine. The replacement of Cys 67 and 95 with Aba showed that these residues were not essential for activity. In addition, this also demonstrated that the substitution of unnatural amino acids in proteins can be useful preventing potential handling problems (e.g., via oxidation).

It has also been possible to take advantage of the ability of chemical synthesis to allow for the incorporation of fixed structural elements into proteins. This was demonstrated by the work of Baca and coworkers[91], where a constrained, non-peptidic type II' beta turn mimic (BTD) was incorporated, replacing Gly 16, 17 in each subunit of the homodimeric HIV-1 PR. These residues are the central two residues involved a type I' beta turn in the native enzyme. The BTD-containing HIV-1 PR analogue possessed an identical activity profile as the native enzyme, but also displayed increased resistance to thermal inactivation. This work demonstrated that the precise geometry of the beta turn was not critical for enzymatic activity, and replacement of these residues with a rigid beta turn mimetic can give the protein enhanced stability. This type of mutagenesis experiment can be extended to other proteins in order to make them more robust under varying conditions--for reactions in organic solvents, higher temperatures, in the presence of denaturants, etc.

Advances in the methods of SPPS which made it possible to consistently and reproducably produce L-[$Aba^{67,95,167,195}$]HIV-1 PR in high purity has also provided for the synthesis of its enantiomeric form using all D-amino acids[92]. The protein enantiomer, as expected, displayed reciprocal chiral specificity in its biochemical interactions. This type of

experiment can only be performed using chemcical synthesis techniques since ribosomal biosynthetic machinery will not allow for the incorporation of D-amino acids[93].

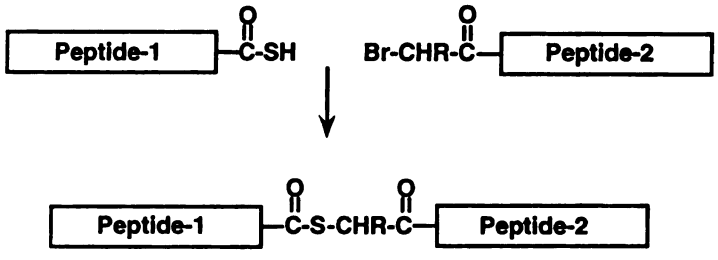
1.9.2 Backbone engineering and chemical ligation--a whole new look for proteins

Infinite access to the world of proteins has been gained through the development of chemical ligation techiniques which allow for the coupling of large unprotected peptide segments. Chemical ligation involves the chemoselective reaction of unique, mutually reactive functionalities, one on each segment, forming a stable bond even in the presence of the range of functional groups usually found in peptides. Significantly, the ligation chemistry used is not limited to the formation of a peptide bond as observed in Figure 1.10.

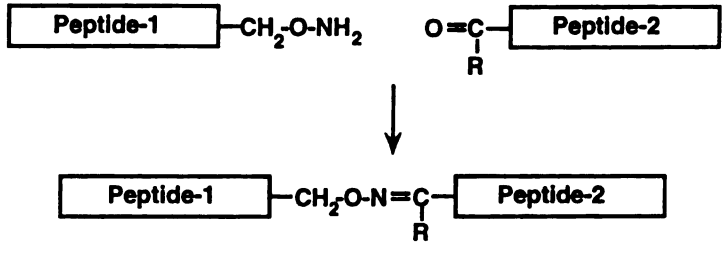
A variety of ligation chemistries have been developed and used to construct many of proteins of biological interest[94-100]. The advent of these novel ligation techniques have made it possible to perform studies on the HIV-1 PR that allowed for the elucidation of certain aspects of the enzyme's mechanism which could not have been done by conventional means. In addition, the development of techniques for the chemical ligation of unprotected peptide segments has provided for the evolution of the next generation of HIV-1 PR analogues--backbone-engineered HIV-1 PR. Chemical ligation techniques were used in the preparation large quantities of all L- and D-HIV-1 PR for crystal structure determination and for mechanistic studies[97,101-103]. An example of the general application of these newly developed chemical ligation techniques is in the total chemical

Figure 1.10 Examples of the chemoselective reactions used for the chemical ligation of unprotected peptide segments.

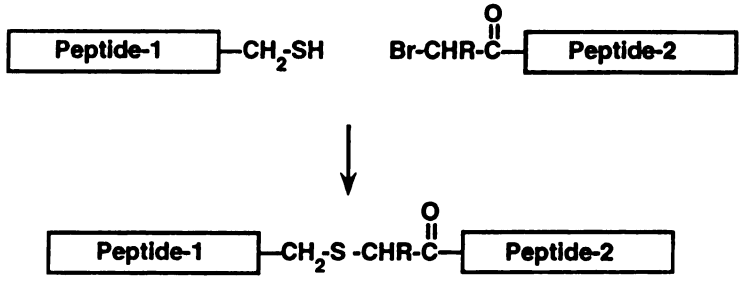
I. Thioester formation



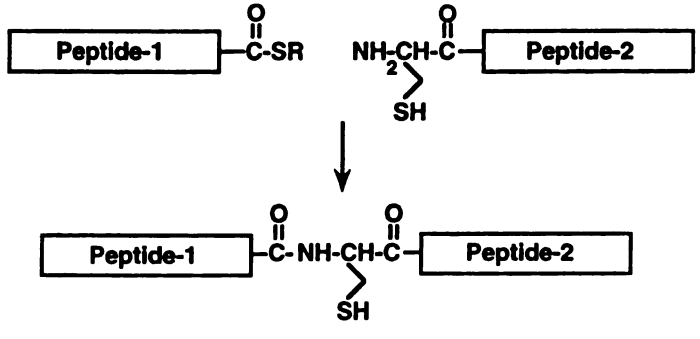
III. Oxime formation



II. Thioether formation



IV. Native bond formation



synthesis of a tethered dimer form of the HIV-1 PR[94]. The 22kDa enzyme displayed full enzymatic activity and is the largest functional protein ever prepared by total chemical synthesis. Chemical access to a tethered dimer form of the HIV-1 PR will make it possible to introduce backbone modifications and unnatural elements of structure asymetrically in only one subunit of the enzyme. These examples truly demonstrate the versatility of the chemical synthesis approach in creating "tailor-made" proteins.

1.10 Aims of thesis studies

The body of work to be described deals with the total chemical synthesis and subsequent characterization of HIV-1 PR and mutants. Firstly, a modified form of the protease which would be stable to autolysis was chemically synthesized using two different methods, total stepwise methods and native chemical ligation. Secondly, the technique of native chemical ligation was used to synthesize HIV-1 PR for the rapid generation of drug-resistant forms. Finally, the investigation of whether it would be possible to alter the substrate specificity of the HIV-1 PR via incorporation of an unnatural amino acid in the enzyme's substrate binding pocket was explored.

References

(1). "In Point of Fact," World Health Organization Press Office, 1995.

(2). *Centers for Disease Control and Prevention. HIV/AIDS Surveillance Report* **1994**, *6*, 5.

(3). Darby, S. C.; Ewart, D. W.; Giangrande, P. L. F.; Dolin, P. J.; Spooner, R. J. D.; Rizza, C. R. *Nature* **1995**, *377*, 79-82.

(4). Coffin, J.; Haase, A.; Levy, J. A.; Montagnier, L.; Oroszlan, S.; Teich, N.; Temin, H.; Toyoshima, K.; Varmus, H.; Vogt, P.; Weiss, R. *Science* **1986**, *232*, 697.

(5). Gallo, R. C.; Montagnier, L. *Sci Am* **1988**, *259*, 41-48.

(6). Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K. J.; Starcich, B.; Josephs, S. F.; Doran, E. R.; Rafalski, J. A.; Whitehorn, E. A.; Baumeister, K.; Ivanoff, L.; Petteway, S. R.; Pearson, M. L.; Lautenberger, L. A.; Papas, T. S.; Ghrayeb, J.; Chang, N. T.; Gallo, R. C.; Wong-Stahl, F. *Nature* **1985**, *313*, 277-284.

(7). Sanchez-Pescador, R.; Power, M. D.; Barr, P. J.; Steimer, K. S.; Stempien, M. M.; Brown-Shimer, S. L.; Gee, W. W.; Renard, A.; Randolph, A.; Levy, J. A.; Dina, D.; Luciw, P. A. *Science* **1985**, *227*, 484-492.

(8). 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.; Montagnier, L. *Science* **1983**, *220*, 868-871.

(9). Clavel, F.; Guétard, D.; Brun-Vézinet, F.; Chamaret, S.; Rey, M. A.; Santos-Ferreira, M. O.; Laurent, A. G.; Dauguet, C.; Katlama, C.; Rouzioux, C.; Klatzmann, D.; Champalimaud, J. L.; Montagnier, L. *Science* **1986**, *233*, 343-346.

(10). Guyader, M.; Emerman, M.; Sonigo, P.; Clavel, F.; Montagnier, L.; Alizon, M. *Nature* **1987**, *326*, 662-669.

(11). Kanki, P. J.; McLane, M. F.; King, N. W.; Letvin, N. L.; Hunt, R. D.; Sehgal, P. K.; Daniel, M. D.; Desrosiers, R. C.; Essex, M. *Science* **1985**, *228*, 1199-1201.

(12). Daniel, M. D.; Letvin, N. L.; King, N. W.; Kannagi, M.; Sehgal, P. K.; Hunt, R. D.; Kanki, P. J.; Essex, M.; Desrosiers, R. C. *Science* **1985**, *228*, 1201-1204.

(13). Letvin, N. L.; Daniel, M. D.; Sehgal, P. K.; Desrosiers, R. C.; Hunt, R. D.; Waldron, L. M.; MacKey, J. J.; Schmidt, D. K.; Chalifoux, L. V.; King, N. W. *Science* **1985**, *230*, 71-73.

(14). Muesing, M. A.; Smith, D. H.; Cabradilla, C. D.; Benton, C. V.; Lasky, L. A.; Capon, D. J. *Nature* **1985**, *313*, 450-458.

(15). Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M. *Cell* **1985**, *40*, 9-17.

(16). Trono, D. *Cell* **1995**, *82*, 189-192.

(17). Preston, B. D.; Poiesz, B. J.; Loeb, L. A. *Science* **1988**, *242*, 1168-1171.

(18). Roberts, J. D.; Bebenek, K.; Kunkel, T. A. *Science* **1988**, *242*, 1171-1173.

(19). Yarchoan, R.; Mitsuya, H.; Broder, S. *Science* **1988**, *259*, 110-119.

(20). Jacks, T.; Power, M. D.; Masiarz, F. R.; Luciw, P. A.; Barr, P. J.; Varmus, H. E. *Nature* **1988**, *331*, 280-283.

(21). Jacks, T.; Varmus, H. E. *Science* **1985**, *230*, 1237-1242.

(22). Jacks, T.; Townsley, K.; Varmus, H. E.; Majors, J. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 4298-4302.

(23). Mitsuya, H.; Broder, S. *Nature* **1987**, *325*, 773-778.

(24). Mitsuya, H.; Yarchoan, R.; Broder, S. *Science* **1990**, *249*, 1533-1544.

(25). Mitsuya, H. *J. Enzym. Inhib.* **1992**, *6*, 1-8.

(26). Tomasselli, A. G.; Howe, W. J.; Sawyer, T. K.; Wlodawer, A.; Heinrikson, R. L. *Chimica Oggi* **1991**, *9*, 6-27.

(27). Munroe, D. G.; Lau, C. Y. *Chemistry and Biology* **1995**, *2*, 343-350.

(28). Loll, P. J.; Picot, D.; Garavito, R. M. *Nature Structural Biology* **1995**, *2*, 637-643.

(29). Steele, F. *Nature Medicine* **1995**, *1*, 285-286.

(30). Larder, B. A.; Darby, G.; Richman, D. D. *Science* **1989**, *243*, 1731-1734.

(31). Larder, B. A. In *Reverse Transcriptases*; A. M. Shalka and S. P. Goff, Ed.; Cold Spring Harbor Laboratory Press: New York, 1993; pp 205-222.

(32). Seelmeier, S.; Schmidt, H.; Turk, V.; von der Helm, K. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 6612-6.

(33). Mous, J.; Heimer, E. P.; Le Grice, S. F. *J. Virol.* **1988**, *62*, 1433-6.

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

(35). Ashorn, P.; McQuade, T. J.; Thaisrivongs, S.; Tomasselli, A. G.; Tarpley, W. G.; Moss, B. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 7472-6.

(36). McQuade, T. J.; Tomasselli, A. G.; Liu, L.; Karacostas, V.; Moss, B.; Sawyer, T. K.; Heinrikson, R. L.; Tarpley, W. G. *Science* **1990**, *247*, 454-456.

(37). Roberts, M. M.; Oroszlan, S. *Biochemical and Biophysical Research Communications* **1989**, *160*, 486-494.

(38). Toh, H.; Ono, M.; Saigo, K.; Miyata, T. *Nature* **1985**, *315*, 691.

(39). Cooper, J.; Foundling, S.; Hemmings, A.; Blundell, T.; Jones, D. M.; Hallett, A.; Szelke, M. *Eur. J. Biochem.* **1987**, *169*, 215-221.

(40). Pearl, L. H.; Taylor, W. R. *Nature* **1987**, *329*, 351-354.

(41). Meek, T. D.; Dayton, B. D.; Metcalf, B. W.; Dreyer, G. B.; Strickler, J. E.; Gorniak, J. G.; Rosenberg, M.; Moore, M. L.; Magaard, V. W.; Debouck, C. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 1841-1845.

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

(43). Miller, M.; Jaskólski, M.; Rao, J. K.; Leis, J.; Wlodawer, A. *Nature* **1989**, *337*, 576-579.

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

(45). DesJarlais, R. L.; Seibel, G. L.; Kuntz, I. D.; Furth, P. S.; Alvarez, J. C.; Ortiz de Montellano, P. R.; DeCamp, D. L.; Babe, L. M.; Craik, C. S. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 6644-6648.

(46). Rosé, J. R.; Craik, C. S. *Am. J. Respir. Crit Care Med.* **1994**, *150*, S176-S182.

(47). Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. *Science* **1989**, *246*, 1149-1152.

(48). Rodriguez, E. J.; Angeles, T. S.; Meek, T. D. *Biochemistry* **1993**, *32*, 12380-12385.

(49). Hyland, L. J.; T.A., T.; Meek, T. D. *Biochemistry* **1991**, *30*, 8454-8463.

(50). Hyland, L. J.; T.A., T.; Roberts, G. D.; Carr, S. A.; Magaard, V. W.; Bryan, H. L.; Fakhoury, S. A.; Moore, M. L.; Minnich, M. D.; Culp, J. S. *Biochemistry* **1991**, *30*, 8441-8453.

(51). Meek, T. D.; Rodriguez, E. J.; Angeles, T. S. *Methods in Enzymology* **1994**, *241*, 127-156.

(52). Billich, S.; Knoop, M.; Hansen, J.; Strop, P.; Sedlacek, J.; Mertz, R.; Moelling, K. *J. Biol. Chem.* **1988**, *263*, 17905-17908.

(53). Pettit, S. C.; Simsic, J.; Loeb, D. D.; Everitt, L.; Hutchison III, C. A.; Swanstrom, R. *Journal of Biological Chemistry* **1991**, *266*, 14539-14547.

(54). Henderson, L. E.; Copeland, T. D.; Sowder, R. C.; Schultz, A. M.; Oroszlan, S. In *Human Retroviruses, Cancer, and AIDS: Approaches to Prevention and Therapy*; D. Bolognesi, Ed.; Alan R. Liss, Inc.: New York, 1988; pp 135-147.

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

(56). Tomasselli, A. G.; Hui, J. O.; Sawyer, T. K.; Staples, D. J.; Bannow, C. A.; Reardon, I. M.; Chaudhary, V. K.; Fryling, C. M.; Pastan, I.; Fitzgerald, D. J. *J. Virol.* **1990**, *64*, 3157-61.

(57). Tomasselli, A. G.; Hui, J. O.; Sawyer, T. K.; Staples, D. J.; FitzGerald, D. J.; Chaudhary, V. K.; Pastan, I.; Heinrikson, R. L. *J. Biol. Chem.* **1990**, *265*, 408-13.

(58). Tomasselli, A. G.; Howe, W. J.; Hui, J. O.; Sawyer, T. K.; Reardon, I. M.; DeCamp, D. L.; Craik, C. S.; Heinrikson, R. L. *Proteins* **1991**, *10*, 1-9.

(59). Tomasselli, A. G.; Hui, J. O.; Adams, L.; Chosay, J.; Lowery, D.; Greenberg, B.; Yem, A.; Deibel, M. R.; Zurcher-Neely, H.; Heinrikson, R. L. *J. Biol. Chem.* **1991**, *266*, 14548-53.

(60). Tomasselli, A. G.; Hui, J. O.; Sawyer, T. K.; Thaisrivongs, S.; Hester, J. B.; Heinrikson, R. L. *Adv. Exp. Med. Biol.* **1991**, *306*, 469-82.

(61). Hui, J. O.; Tomasselli, A. G.; Zurcher-Neely, H. A.; Heinrikson, R. L. *J. Biol. Chem.* **1990**, *265*, 21386-9.

(62). Riviere, Y.; Blank, V.; Kourilsky, P.; Israel, A. *Nature* **1991**, *350*, 625-626.

(63). Shoeman, R. L.; Kesselmier, C.; Mothes, E.; Honer, B.; Traub, P. *FEBS Lett.* **1991**, *278*, 199-203.

(64). Oswald, M.; von der Helm, K. *FEBS Lett.* **1991**, *292*, 298-300.

(65). Schecter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157-162.

(66). Wlodawer, A. *Pharmacotherapy* **1994**, *14*, 9S-20S.

(67). Huff, J. R. *J. Med. Chem.* **1991**, *34*, 2305-2314.

(68). Debouck, C. *AIDS Res. Hum. Retroviruses* **1992**, *8*, 153-164.

(69). Wlodawer, A.; Erickson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543-585.

(70). Meek, T. D. *J. Enzym. Inhib.* **1992**, *6*, 65-98.

(71). Darke, P. L.; Huff, J. R. *Adv. Pharmacol.* **1994**, *25*, 399-454.

(72). Greenlee, W. J. *Med. Res. Rev.* **1990**, *10*, 173-236.

(73). Richards, A. D.; Roberts, R.; Dunn, B. M.; Graves, M. C.; Kay, J. *FEBS Lett.* **1989**, *247*, 113-7.

(74). Roberts, N. A.; Martin, J. A.; Kinchington, D.; Broadhurst, A. V.; Craig, J. C.; Duncan, I. B.; Galpin, S. A.; Handa, B. K.; Kay, J.; Krohn, A.; Lambert, R. W.; Merrett, J. H.; Mills, J. S.; Parkes, K. E. B.; Redshaw, S.; Ritchie, A. J.; Taylor, D. L.; Thomas, G. J.; Machin, P. J. *Science* **1990**, *248*, 358-361.

(75). Roberts, N. A.; Craig, J. C.; Duncan, I. B. *Biochem. Soc. Trans.* **1992**, *20*, 513-6.

(76). Vella, S. *AIDS* **1994**, *8*, S25-S29.

(77). 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.; Wideburg, N. E.; Saldivar, A.; Ruiz, L.; Kati, W.; Sham, H. L.; Robins, T.; Stewart, K. D.; Hsu, A.; Plattner, J. J.; Leonard, J. M.; Norbeck, D. W. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2484-2488.

(78). Lam, P. Y.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; Chang, C. H.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R.; Erickson-Viitanen, S. *Science* **1994**, *263*, 380-384.

(79). Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Wideburg, N. *Science* **1990**, *249*, 527-533.

(80). Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.* **1982**, *161*, 269-288.

(81). 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.; et, a. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 4096-4100.

(82). Lillehoj, E. P.; Salazar, F. H.; Mervis, R. J.; Raum, M. G.; Chan, H. W.; Ahmad, N.; Venkatesan, S. *J. Virol.* **1988**, *62*, 3053-3058.

(83). Hilvert, D. *Chemistry and Biology* **1994**, *1*, 201-203.

(84). Muir, T. W.; Kent, S. B. H. *Curr. Opin. Biotech.* **1993**, *4*, 420-427.

(85). Muir, T. M. *Structure* **1995**, *3*, 649-652.

(86). Wallace, C. J. A. *Current Opinion in Biotechnology* **1995**, *6*, 403-410.

(87). Schnölzer, M.; Jones, A.; Alewood, P. F.; Kent, S. B. H. *Anal. Biochem.* **1992**, *204*, 335-343.

(88). Kent, S. B. H. *Ann. Rev. Biochem.* **1988**, *57*, 957-989.

(89). Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180-193.

(90). Schneider, J.; Kent, S. B. H. *Cell* **1988**, *54*, 363-368.

(91). Baca, M.; Alewood, P. F.; Kent, S. B. H. *Protein Sci.* **1993**, *2*, 1085-1091.

(92). Milton, R. C.; Milton, S. C.; Kent, S. B. H. *Science* **1992**, *256*, 1445-8.

(93). Ellman, J. A.; Mendel, D.; Schultz, P. G. *Science* **1992**, *255*, 197-200.

(94). Baca, M.; Muir, T. W.; Schnölzer, M.; Kent, S. B. H. *J. Am. Chem. Soc.* **1994**, *117*, 1881-1887.

(95). Canne, L. E.; Ferré-D'Amaré, A. R.; Burley, S. K.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 2998-3007.

(96). Muir, T. W.; Williams, M. J.; Ginsberg, M. H.; Kent, S. B. H. *Biochemistry* **1994**, *33*, 7701-7708.

(97). Schnölzer, M.; Kent, S. B. H. *Science* **1992**, *256*, 221-225.

(98). Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776-779.

(99). Williams, M. J.; Muir, T. W.; Ginsberg, M. H.; Kent, S. B. H. *J. Am. Chem. Soc.* **1994**, *116*, 10797-10798.

(100). Gaertner, H. F.; Rose, K.; Cotton, R.; Timms, D.; Camble, R.; Offord, R. E. *Bioconjugate Chem.* **1992**, *3*, 262-268.

(101). Baca, M.; Kent, S. B. H. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11638-11642.

(102). Miller, M. "D-HIV-1 protease structure," 1993.

(103). Milton, R. C.; Milton, S. C. F.; Schnölzer, M.; Kent, S. B. H. In *Techniques in Protein Chemistry IV*; R. H. Angeletti, Ed.; Academic Press: San Diego, 1993; pp 257-267.

Chapter II

Total Chemical Synthesis of HIV-1 Protease: Comparison of Synthetic Methods

2.1 Introduction

The protease from the human immunodeficiency virus type 1 (HIV-1 PR) is an important therapeutic target for the treatment of acquired immunodeficiency syndrome (AIDS) and, as such, has become one of the most intensively studied enzymes [1-5]. The action of HIV-1 PR is essential for viral maturation, as shown by the prevention of the formation of infectious virions in the presence of inhibitors of HIV-1 PR[6-10] or by mutations which render the protease inactive [11,12]. Access to reagent quantities of this enzyme is necessary on a continuing basis in order to carry out essential biochemical and structural studies. The predominant source of HIV-1 PR has been via recombinant DNA technology, yet the relatively small size of the protease monomer polypeptide (99 amino acids in length) also makes it an accessible target by total chemical synthesis [13,14].

Chemical synthesis has proven to be a good method for producing sufficient quantities of the HIV-1 PR and has facilitated its study by providing crystallizable material for use in the original structural studies of the enzyme[14-16]. A chemical synthesis approach also allows for a wider range of changes to be made in the covalent structure of a protein[17-19] than is generally acessible by recombinant methods such as site-directed muatagenesis[20,21]. However, the large size of most proteins generally makes their synthesis by stepwise solid-phase peptide synthesis (SPPS) difficult, mainly because of the complexity of the final crude product. This arises from slight imperfections in the chemistry of chain assembly and the consequent accumulation of resin-bound byproducts which copurify with the target sequence, leading to complex product mixtures. Because of the

geometric shortcomings of stepwise SPPS, this problem is particularly severe for long polypeptides. However, in practice, one can readily take advantage of the ability of optimized stepwise solid phase methods to make peptides of up to 60 amino acids in high purity and yield[22,23]. These peptides may be linked together unambiguously using recently developed chemical ligation methods[24-27], where large, unprotected peptides are linked together by chemoselective reactions.

Chemical ligation takes peptide synthesis into the next generation by providing access to even larger, more complex protein analogues[28-31]. The inherent advantages of the chemical ligation approach are in the practicality of preparing unprotected peptide segments up to approximately 60 amino acids and in the clean, reproducible chemoselective ligation reactions, which yield products that are more easily purified. Recently, another chemical ligation approach has been introduced which allows for direct synthetic access to proteins of native backbone structure. This technique, appropriately termed native chemical ligation[24], has the potential to provide expanded total chemical synthetic access to the world of proteins[17].

We were interested in comparing and contrasting stepwise SPPS vs. chemical ligation methods for the total chemical synthesis of the HIV-1 PR. Also, we wanted to produce an autolysis-resistant form of the HIV-1 PR by chemical synthesis for our ongoing biochemical studies of the molecule. Instability of the enzyme upon storage can interfere with proper functional characterization[32]. In addition, autolysis breakdown products can hamper the growth of good crystals and interfere with obtaining accurate data in NMR studies and other physical measurements. Following the extensive

work of Mildner et al.[33] and Rosé[34], where mutations introduced at P1 or P2' residues (notation according to Schecter and Burger[35]) of the major autolysis sites in the HIV-1 PR, we have chemically synthesized autolysis-resistant analogues of the HIV-1 PR by two methods, via stepwise SPPS and the newly developed method of native chemical ligation. Here we compare contrast the results obtained with the two different synthetic strategies.

2.2 Results

2.2.1 Stepwise solid phase synthesis of the HIV-1 PR monomer.

The target monomer of the homodimeric HIV-1 protease consists of 99 amino acid residues. The target sequence synthesized from the SF2 isolate of the virus[36] which has been modified for resistance against autolysis is shown in Figure 2.1 and the general strategy for its synthesis is outlined in Figure 2.2. L-α-Amino-n-butyric acid (Aba) was used in place of the native cysteine residues at positions 67 and 95 in the sequence[16]. The synthesis was performed by stepwise solid-phase methods using machine-assisted in situ neutralization/2-(1-H-benzotriazol-1-yl)-1,1,3,3,-tetramethyluronium hexafluorophosphate (HBTU) activation protoclols for tert-butoxycarbonyl (Boc) chemistry according to published procedures[23]. The peptide was assembled on Boc-Phe-OCH_2-Pam resin, and the assembly of the peptide chain was performed in a stepwise fashion to give the fully protected protease monomer. The efficiency of the chain assembly at each step in the synthesis was determined by quantitative ninhydrin assay for residual free amine[37]. Based on this assay, the average yield for the chain assembly was approximately 99%. Deprotection and cleavage of the peptide resin in

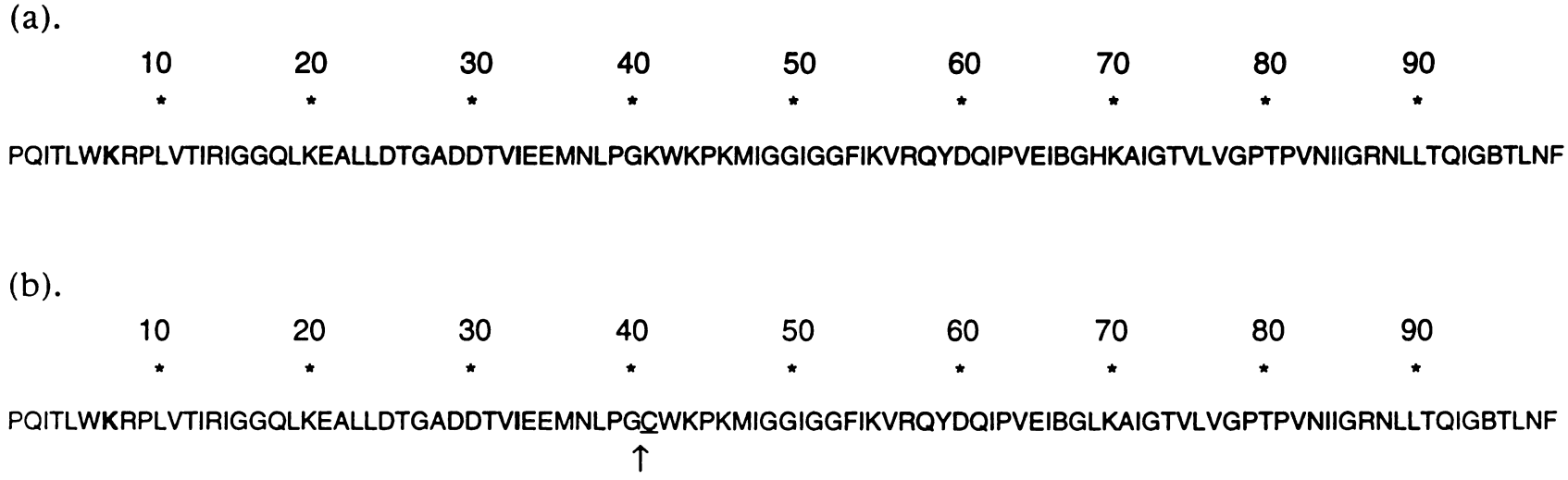
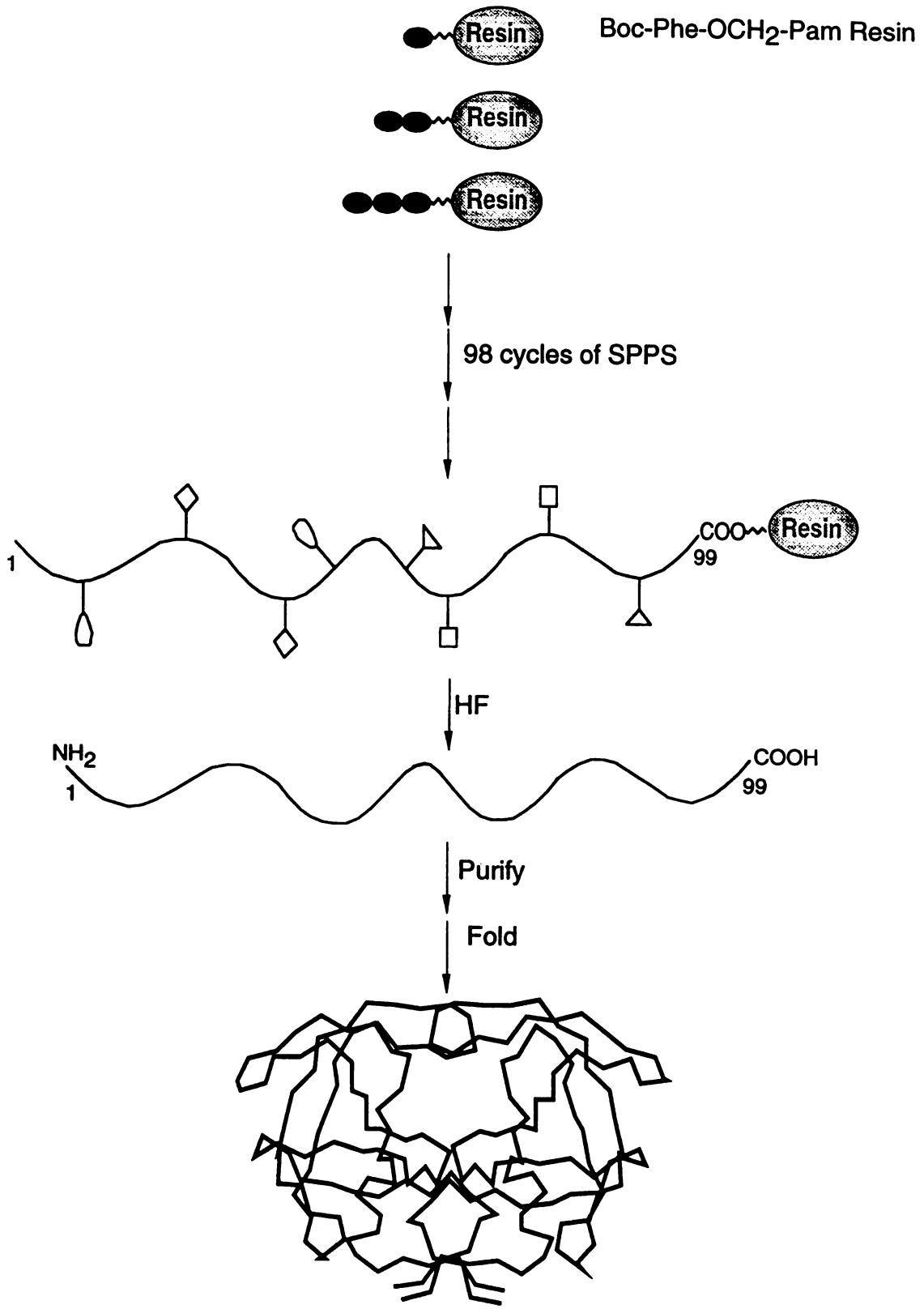


Figure 2.1 Sequence of the autolysis-resistant HIV-1 PR monomer synthetic target via (a). stepwise and (b). native chemical ligation. Changes in the sequence made for resistance are highlighted in bold. B is L-α-amino-n-butyric acid (Aba), which has been substituted for $Cys^{67,95}$ in the HIV-1 PR monomer. The arrow indicates the site of ligation and the K41C mutation is underlined.

Figure 2.2 Schematic of the total chemical synthesis of HIV-1 PR via stepwise SPPS.

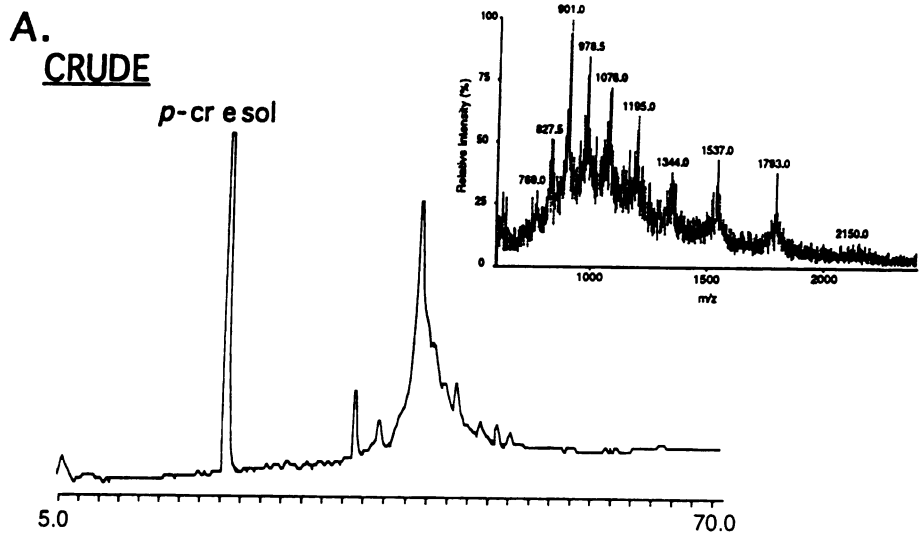

Resin
Resin
Resin
Boc-Phe-OCH2-Pam Resin
98 cycles of SPPS
1
COO
Resin
99
HF
NH2
1
COOH
99
Purify
Fold

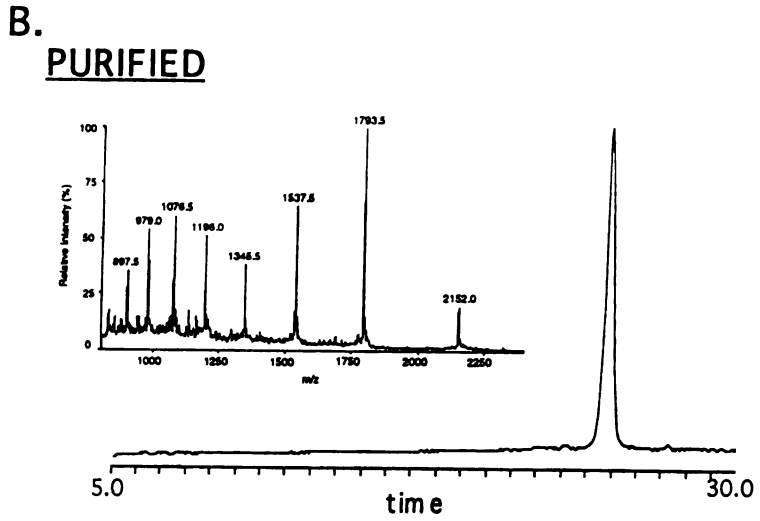
anhydrous HF afforded crude HIV-1 PR. The crude product was characterized by analytical reversed-phase HPLC and by ESMS (Figure 2.3) and checked for the levels of deletions, terminations and modified pepides present in the crude product. Upon analysis of the ESMS spectrum, it appeared that most of the byproducts could be attributed to termination and deletion products. The accumulation of deletion products of mass -113 Da from the product can most likely be accounted for due to the abundance of Leu/Ile residues present in the primary sequence of HIV-1 PR (residue mass 113 Da). Of the residues, 23 out of 99 in the primary sequence of the enzyme are either Leu or Ile residues.

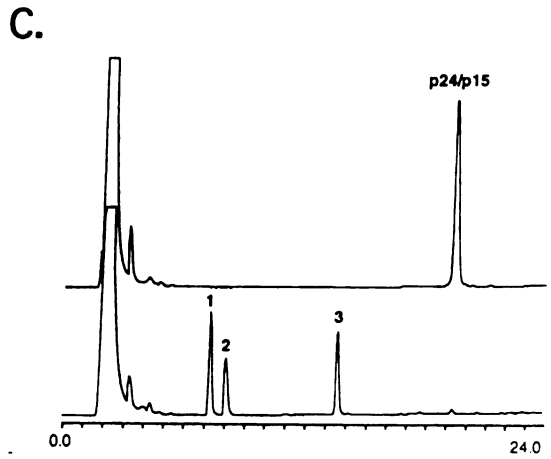
HIV-1 PR monomer from stepwise SPPS was purified by preparative HPLC, and the 99-residue polypeptide chain was characterized by analytical HPLC and ionspray MS (Figure 2.3). The observed molecular weight for the HIV-1 PR monomer was 10,755 Da and is within experimental error of the calculated MW [10,755 Da (average isotope composition)]. After folding by dialysis from a 6M guanidine hydrochloride (GuHCl) solution of the synthetic enzyme into assay buffer, HPLC assays were used to confirm the enzymatic activity. Treatment of synthetic peptides spanning known cleavage sites in the HIV-1 gag polyprotein resulted in cleavage at sites known to be processed by the viral enzyme in the course of viral maturation[38].

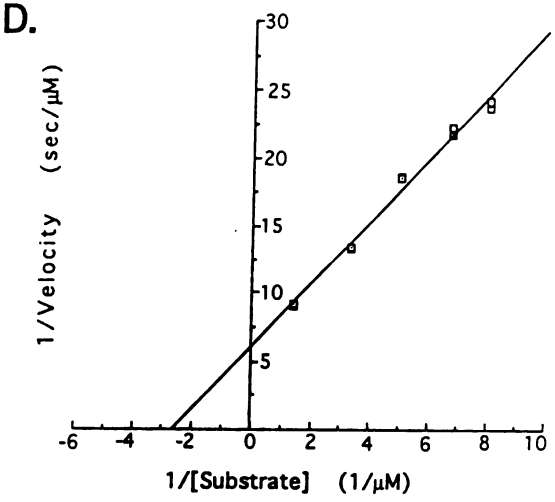
The kinetic parameters of autolysis-resistant [Aba$^{67, 95, 167, 195}$] HIV-1 PR were analyzed based on the cleavage of a peptide substrate (RNTATIM*MQRGNFR-amide) derived from the second capsid/nucleocapsid (p1/NC) junction of the viral gag polyprotein. Kinetic

Figure 2.3 Synthesis and characterization of autolysis-resistant HIV-1 PR from stepwise SPPS. (A). Crude 99 residue product obtained after cleavage from peptide resin analyzed by analytical HPLC (using a linear gradient of 0-70%B over 70 minutes) and electrospray MS (inset). (B). Purified autolysis-resistant HIV-1 PR was analyzed by analytical HPLC. The protein peak was collected in a single fraction and analyzed by electrospray MS (inset). The experimentally determined mass of the synthetic enzyme (10, 755 ± 2 Da) was in agreement with the calculated mass of 10,755 Da. (C). The specificity of the synthetic HIV-1 PR enzyme was determined by the cleavage of a synthetic peptide analogue of the capsid/nucleocapsid (CA/NC) processing site in the HIV-1 gag polypeptide. The CA/NC processing site peptide GHKARVL*AEAMSQVTNSATIM*MQRGNFRNQRK (200μM) was incubated with synthetic HIV-1 PR at pH=5.5, 37°C for 3 hours. Analytical HPLC and electrospray MS analysis of the reaction mixture showed 3 cleavage products which correspond to the expected cleavage at the Met*Met and Leu*Ala bonds. The cleavage products were characterized as MQRGNFRNQRK (peak 1), GHKARVL (peak 2), and AEAMSQVTNSATIM (peak 3). (D). Kinetics. Steady state kinetic parameters were determined for cleavage of the synthetic peptide substrate RNTATIM*MQRGNFR-amide. Intitial velocities were measured as a function of substrate concentration. The data are presented graphically in a double-reciprocal plot. Kinetic parameters were calculated by nonlinear regression analysis as k_{cat} = 13 sec^{-1} and K_m = 424μM. These values are similar to those obtained with native enzyme[39,40].


A.
CRUDE
p-cresol
5.0
70.0
Relative Intensity (%)
m/z
901.0
978.5
1076.0
1196.0
827.5
788.0
1344.0
1537.0
1783.0
2150.0


B.
PURIFIED
5.0
30.0
time
Relative Intensity (%)
m/z
1793.5
1537.5
1076.5
979.0
1196.0
897.5
1345.5
2152.0


C.
p24/p15
1
2
3
0.0
24.0


D.
1/Velocity (sec/μM)
1/[Substrate] (1/μM)

parameters determined for the cleavage of p1/NC substrate are presented in Table 2.1.

2.2.2 Synthesis of the HIV-1 PR monomer via native chemical ligation.

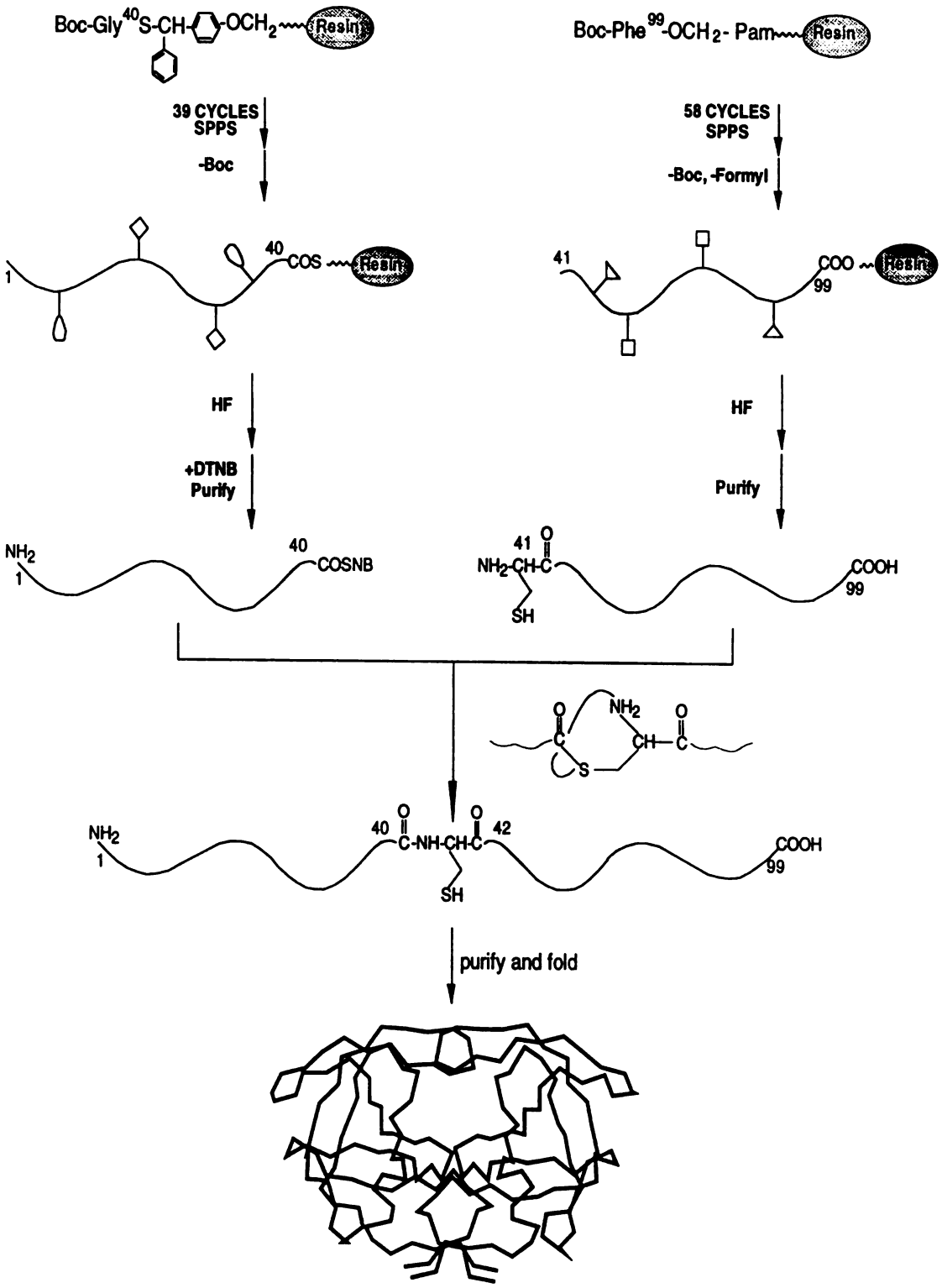
The 99-residue target sequence was synthesized in two peptide segments, as shown in Figure 2.4. The two segments, comprising HIV-1 PR residues 1-40COSNB and 41-99 (K41C), were ligated via native chemical ligation[24]. This chemical ligation technique allows for the synthesis of proteins with native backbone structure via the chemoselective reaction of two unprotected peptide segments. At the present time, native chemical ligation is limited to joining two segments at a Xxx-Cys site. The HIV-1 PR does contain two cysteine residues in its native sequence. However, they are contained in the latter half of the molecule and are inconveniently located for segment condensation. Thus, we chose to mutate one of the residues (K41C) near the middle of the molecule, directly adjacent to a glycine residue, in order to optimize both the ligation site and give comparably-sized segments. The carboxy-terminal segment, corresponding to residues 41-99 in the HIV-1 PR monomer, was synthesized on Boc-Phe-OCH_2-Pam resin. The amino-terminal segment, corresponding to residues 1-40, was synthesized on a Gly-thioester support, which upon cleavage yields the peptide α-thiocarboxylate. This peptide is then treated with DTNB (Ellman's Reagent) to generate an active α-thioester (COSNB). Thiophenol was used in the reaction as a reducing agent in order to keep the N-terminal Cys residue in a reduced and reactive form, without interfering with the reaction. It appeared that the addition of thiophenol also aided in the reaction via transthioesterification of any 1-40 COSNB which had not ligated

Table 2.1 Kinetic Constants for Autolysis-Resistant HIV-1 PR Produced by Stepwise SPPS or Native Chemical Ligation

Proteases	K_m (μM)	k_{cat} (sec^{-1})	k_{cat}/K_m ($\mu M^{-1}sec^{-1}$)
Stepwise	424	13	0.031
Ligated	141	4.4	0.031

The enzyme concentations used in both assays was approximately 11 nM. Kinetic parameters were determined for the cleavage of the synthetic peptide substrate RNTATIM*MQRGNFR-amide, where * indicates the site of cleavage. The substrate concentration varied between 50-900mM in the assays. Stepwise PR was assayed in a buffer containing 100mM sodium acetate, 0.5 M NaCl, 10% glycerol and 0.5 mgs/ml BSA at pH=5.5, 37°C. DTT (2mM) and EDTA (1mM) were added to the above buffer in assays of the native ligated PR.

Figure 2.4 Schematic of the total chemical synthesis of HIV-1 PR by native chemical ligation.


Boc-Gly 40 S-CH- -OCH2 Resin
39 CYCLES SPPS
-Boc
Boc-Phe 99 -OCH2- Pam Resin
58 CYCLES SPPS
-Boc, -Formyl
40 COS Resin
1
41
COO Resin
99
HF
+DTNB Purify
HF
Purify
NH2 1
40 COSNB
41 O
NH2-CH-C
SH
COOH 99
O NH2 O
C CH C
S
NH2 1
40 O
C-NH-CH-C 42
SH
COOH 99
purify and fold

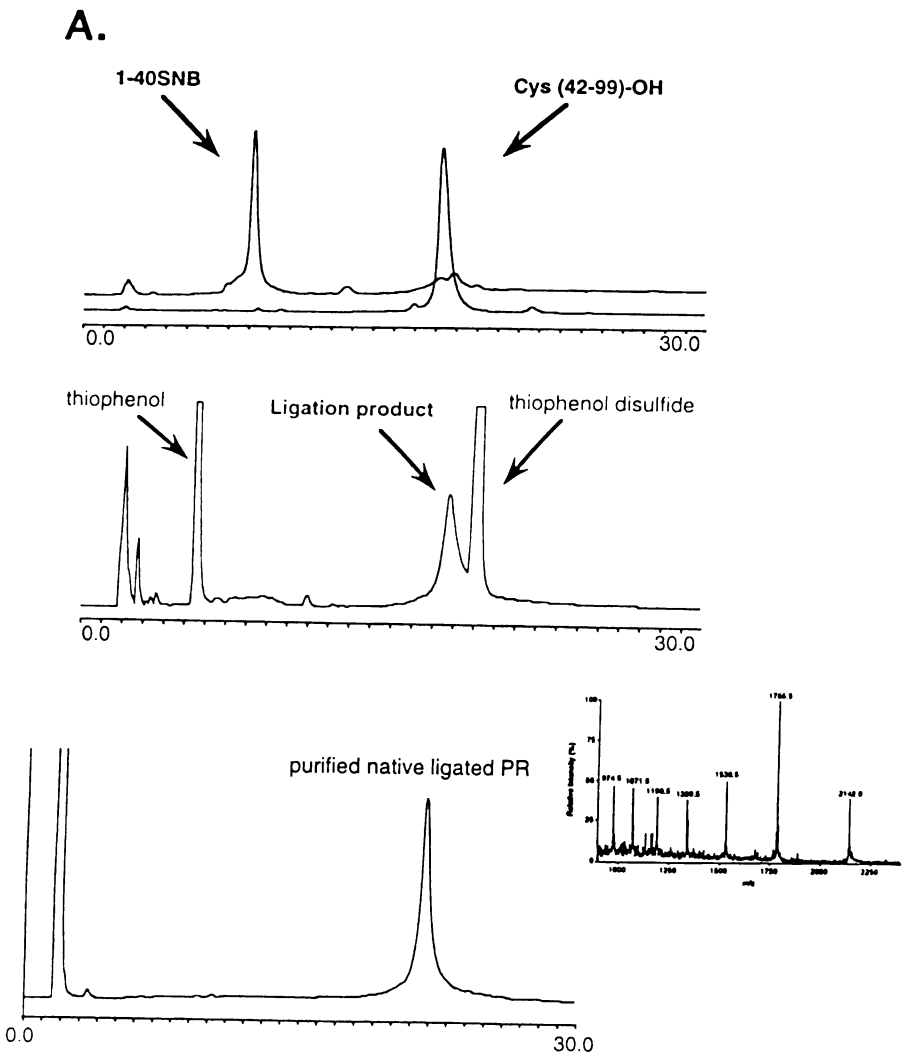
immediately and reduced the level of hydrolysis typically seen with SNB peptide esters since they are highly activated thioesters. The peptide thiophenol ester is more stable to hydrolysis, but reactive enough to facilitate the ligation and generate product in near quantitative yields (P. Dawson, unpublished results).

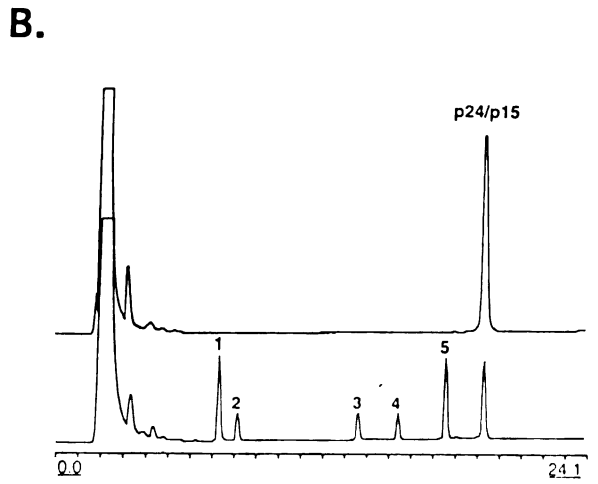
The ligation reaction of the two segments was performed in aqueous buffer at pH=6.5, and the course of the reaction was followed by analytical reverse-phase HPLC, as shown in Figure 2.5. The ligation reaction went in high yield (near quantitative, >95%), cleanly, and showed little or no side product formation. The resulting native ligated monomer was readily purified by gel-filtration and semipreparative HPLC. Analytical HPLC showed the 99 residue polypeptide to be of high purity (>97%). Subsequent characterization by electrospray MS gave a molecular weight of 10,705 Da. [calculated:10,706 Da (average isotope composition)]. Kinetic parameters for this enzyme were determined in the same manner as per the HIV-1 PR synthesized by direct stepwise methods and are shown in Table 2.1. The kinetic parameters were similar to those of the native enzyme, despite the K41C mutation. The observed substrate specificity on synthetic peptide analogues of the gag polyprotein cleavage sites were identical to that of the viral enzyme (Figure 2.5).

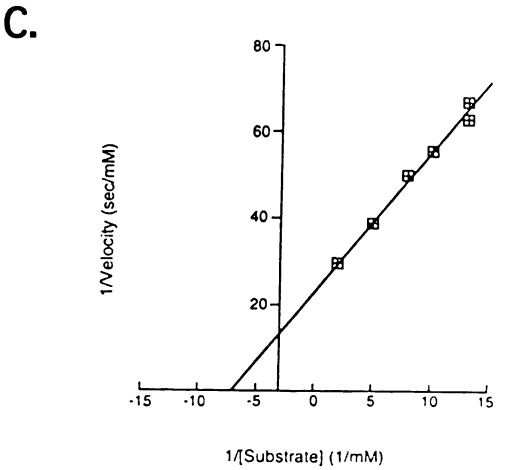
2.3 Discussion

The two HIV-1 PR analogues synthesized by either stepwise methods or native chemical ligation possessed similar substrate specificites and kinetic properties. The values obtained and summarized in Table 2.1 are

Figure 2.5 Native chemical ligation of autolysis-resistant HIV-1 PR. (A). Ligation. The upper HPLC trace shows the purified peptide starting materials at t=0. The middle trace shows the extent of the ligation reaction at t > 24 hours. An aliquot of the purified 99 amino acid monomer analyzed by HPLC and ESMS is shown in the lower trace. (B). Specificity. Specific cleavage of a synthetic peptide analogue of the capsid/nucleocapsid (CA/NC) processing site in the HIV-1 gag translation product is shown. CA/NC peptide GHKARVL*AEAMSQVTNSATIM*MQRGNFRNQRK (200μM) was incubated with native ligated PR at pH = 5.5, 37°C for 3 hours. Reaction products were separated by analytical HPLC and identified by ESMS. Products were characterized as MQRGNFRNQRK (peak 1), GHKARVL (peak 2), AEAMSQVTNSATIM (peak 3), AEAMSQVTNSATIMMQRGNFRNQRK (peak 4) and GHKARVLAEAMSQVTNSATIM (peak 5). (C). Kinetics. Steady state kinetic parameters were determined for the cleavage of the synthetic peptide substrate RNTATIM*MQRGNFR-amide. Initial velocities were measured as a function of substrate concentration. The data are represented graphically in a double-reciprocal plot. Kinetic parameters for native-ligated HIV-1 PR calculated by non-linear regression analysis were kcat = 4.4 sec^{-1} and Km = 141μM. These values were similar to those obtained with native enzyme[39,40].


A.
1-40SNB
Cys (42-99)-OH
0.0
30.0
thiophenol
Ligation product
thiophenol disulfide
0.0
30.0
purified native ligated PR
0.0
30.0


B.
p24/p15
1
2
3
4
5
0.0
24.1


C.
80
60
40
20
1/Velocity (sec/mM)
-15
-10
-5
0
5
10
15
1/[Substrate] (1/mM)

similar to those obtained with the native enzyme[39,40]. Based on the information available regarding the substrate specificity of the HIV-1 PR, Mildner et al. set out to make substitutions in the HIV-1 PR which would block cleavage at major sites of autolysis. The rationale was to replace amino acids in positions which flanked the scissile bond with those that are normally excluded at these sites. By introducing β-branched amino acids at P1 or Lys at P2' in the three major autolysis sites in the PR, hydrolysis was found to be greatly reduced at the corresponding P1-P1' positions. Therefore, three changes were introduced (Q7K/L33I/L63I) in the HIV-1 PR strain under study in their laboratory (pBH10), which stabilized the enzyme against autolysis. In the isolate used in our laboratory (SF2 strain), only two changes were made in the sequence(Q7K/L33I). The enzyme was found to be quite stable, even after months of storage in folding buffer at 4°C. No problems with autolysis were observed in our hands, even at the high concentrations of HIV-1 PR (>5mg/ml) needed for NMR (R. Smith, personal communication).

A stabilized form of the HIV-1 PR has been produced by two different methods employing total chemical synthesis which proves excellent for biophysical analysis. The yields of purified protease were far superior for the HIV-1 PR molecule synthesized by native chemical ligation (33% yield) as compared to that prepared via direct stepwise methods (4.7% yield). The stepwise synthesis of a protein greater than 60 residues is, indeed, a challenging endeavor. Since the number of byproducts which accumulate over the many steps in the synthesis that are needed increases as the polypeptide chain is built, the indentification and purification of the final product becomes an incredibly labor-intensive process. In addition, the

general stepwise synthesis of large proteins is not always straightforward and is usually achieved only in the most favorable of cases. The approach is also not a practical one if the desire is to generate multiple variants for structure-function studies. However, due to the great success of SPPS in the synthesis of peptide segments of about 60 amino acids in high purity and yield, chemical ligation methods tend to be a more generalized and reliable approach to the construction of proteins of biological interest.

This study has verified that a chemical ligation approach to the production of the HIV-1 PR has many inherent advantages over direct stepwise SPPS. Ease of purification, high purity and yields of the protein product are just a few of the advantages of ligation method for the production of HIV-1 PR. In addition, the synthetic convenience of ligation approach allows for a "cassette" approach to constructing proteins. It is feasible to think of purified synthetic peptide fragments as "molecular cassettes", which can be mixed and matched for the rapid and versatile generation of protein analogues. This method could be of great use for investigating the effects of different mutations on the HIV-1 PR's activity. For example, such a system could be especially useful for analyzing the kinetic effects of mutations that render the HIV-1 PR resistant to various inhibitors which have been developed as potential AIDS therapeutics. Such a study has been done and will be described in more detail in Chapter 3.

2.4 Experimental section

2.4.1 Materials and methods.

Boc-amino acids and 2-(1-H-benzotriazol-1-yl)-1,1,3,3,-tetramethyluronium hexafluorophosphate (HBTU) were obtained from Nova Biochem (San Diego, CA.). Boc-Phe-OCH_2Pam-resin and diisopropylethylamine were purchased from Applied Biosysyems, Inc. (Foster City, CA.). Aminomethylcopoly(styrene-divinylbenzene) resin was prepared by L. E. Canne (TSRI). N,N-Dimethylformamide (Synthesis-grade) was obtained from Mallinckrodt Chemical Co. (Paris, KY) or Fisher. AR-grade methylene chloride was obtained from Fisher. TFA was purchased from Halocarbon (River Edge, N. J.) and HF obtained from Matheson Gas. HPLC-grade acetonitrile was obtained from either EM Science or Fisher. Vydac HPLC columns were purchased form Western Analytical Co. (Temecula, CA.). All other reagents were AR grade or better, and purchased from either Aldrich Chemical or Fisher.

Machine-assisted solid-phase peptide syntheses were performed on a custom-modified Applied Biosystems 430A peptide/protein synthesizer. This instrument was modified by installation of bypass valves to accelerate flows and to prevent clogging of valve blocks[23]. Gel-filtration was performed on a Pharmacia FPLC System using a Superdex 75 HR 10/30 column and developed in 50% acetic acid. Preparative HPLC was performed on a Waters Delta-prep 4000 HPLC system. Analytical and semipreparative HPLC were performed on either a Rainin HPXL dual pump system with detection on a Dynamax UV detector or on an integrated

Hewlett-Packard 1050 system. Preparative HPLC was run on a Vydac C18 column (15-20μ, 5.0 x 25 cm) at a flow rate of 30mL/min; semipreparative HPLC was run on a Vydac C4 column (10μ, 1.0 x 25 cm) at a flow rate of 3mL/min: and analytical HPLC was run on a Vydac C4 column (5μ, 0.46 x 15 cm) at 1mL/min. Chromatographic separations were achieved using linear gradients of 0.1% TFA in water (solvent A) versus 90% acetonitrile/ 10% water, 0.09% TFA (solvent B). Mass analyses of all peptide segments and proteins were performed a Sciex API-III triple quadrupole electrospray mass spectrometer as previously described[41]. Observed masses were derived from the experimental m/z values for all observed protonation states of a molecular species using the program MacSpec (Sciex). Calculated masses were based on average isotope composition and were derived using the program MacProMass (Sunil Vemuri and Terry Lee, Beckman Research Institute, Duarte, CA.). Ninhydrin yields and substitution values were calculated using the program Synthesis Editor 9 (J. J. Kent, K.P. Instruments).

2.4.2 Solid-phase peptide synthesis

All peptides were synthesized in a stepwise fashion according to the *in situ* neutralization/HBTU activation protocols for machine-assisted Boc solid-phase chemistry[23]. Side chain protection was as follows: Arg(*p*-toluenesulfonyl), Asn(xanthyl), Asp(*O*-cyclohexyl), Glu(*O*-cyclohexyl), His(Dnp), Lys(2-Cl-Z), Ser(benzyl), Trp(formyl), Thr(benzyl), and Tyr(2-Br-Z). Gln and Met (and, when synthesizing thioacid peptides Trp) were used without side chain protection. Coupling yields were determined by the quantitative ninhydrin assay on peptide-resin sapmles taken under machine

control during the chain assembly[23,37]. The HIV-1 PR monomer and its C-terminal fragment (41-99) were prepared on a Boc-Phe-4-(carboxamidomethyl)benzyl ester-copoly(styrene-1%-divinylbenzene) resin (Boc-Phe-OCH_2-Pam-resin). The N-terminal fragment corresponding to HIV-1 PR (1-40) was constructed on a 4-[α-(Boc-Gly-S)benzyl]-phenoxyacetamidomethyl resin, which yields a peptide with a C-terminal Gly-αCOSH post cleavage[42-44]. Upon completion of the chain assembly, peptides were deprotected as necessary and cleaved from the resin by treatment with liquid HF containing 10% *p*-cresol for 1 hour at 0 °C. After evaporation of HF under reduced pressure, crude peptide products were precipitated in ice-cold anhydrous diethyl ether, dissolved in 50% aqueous acetic acid, and either purified immediately or after prior lyophilization.

2.4.3 Purification and characterization of peptide segments

Crude lyophilized peptides were dissolved in either acidic aqueous buffers or 50% aqueous acetic acid and purified by preparative or semipreparative HPLC. Purified peptides were stored as lyophilized powders at -20°C. The purified peptide segments were characterized by electrospray MS, and all possessed observed masses within experimental error of the calculated masses.

2.4.4 Stepwise synthesis of the autolysis-resistant HIV-1 PR monomer

The synthesis was carried out on a 0.2 mmole scale using Boc-Phe-OCH_2-Pam resin. Fully-protected peptide-resin (1.720 grams) was obtained after completion of the stepwise chain assembly (88% of theoretical yield

after correction for ninhydrin sampling.). A portion of peptide-resin (300mgs) was subjected to deprotection and cleavage. Protecting groups were removed as follows: the dinitrophenyl protecting group of histidine was removed by two 30 minute treatments with 20% mercaptoethanol/10% DIEA/DMF, the amino-terminal Boc group was removed by treatment with neat TFA (two 1 minute treatments) and the formyl protecting groups of tryptophan were removed by two 30 minute treatments with aqueous ethanolamine. The partially unprotected peptide resin was further deprotected and cleaved from the resin with HF, containing 10% p-cresol, for 1 hour at 0°C. After removal of the HF under reduced pressure, the crude protein mixture was precipitated in ice-cold anhydrous diethyl ether, dissolved in 50% aqueous acetic acid and immediately purified by preparative HPLC (45-60% B in 60 minutes). This afforded 10 mgs (4.7% of theoretical yield) of purified, autolysis-resistant HIV-1 PR, which was characterized by electrospray MS to give a mass of 10,755 ± 2 Da [10,755 Da (average isotope composition)].

2.4.5 Synthesis of the autolysis-resistant HIV-1 PR monomer via native chemical ligation

The synthesis of HIV-1 PR (41-99) was performed on a 0.15 mmole scale using Boc-Phe-OCH_2-Pam resin. Peptide-resin (529 mgs) was subjected to deprotection and cleavage as earlier described for the HIV-1 PR monomer produced by direct stepwise methods. The crude peptide was then immediately subjected to preparative HPLC (40-55%B in 60 minutes) and afforded 33mgs (8.9% yield) of purified HIV-1 PR (41-99) for use in the ligation.

Synthesis of HIV-1 PR (1-40)αCOSH was performed on a 0.2 mmol scale on a 4-[α-(Boc-Gly-S)benzyl]-phenoxyacetamidomethyl resin. Peptide-resin (300mgs) was deprotected and cleaved to give 160 mgs of crude peptide (81% yield). The crude synthetic fragment HIV-1 PR(1-40)αCOSH (160mgs, 36.2 μmol) was converted to the 5-thio-2-nitrobenzoic acid ester (-COSNB) by treatment with approximately 2 equivalents (35 mgs) of 5, 5'-dithio-bis(2-nitrobenzoic acid) [DTNB, Ellman's reagent] in 6M GuHCl, 0.2M sodium acetate, pH=5.5 for 20 minutes. The crude reaction mixture was then subjected to preparative HPLC (25-55%B in 60 minutes) for purification and subsequently lyophilized to yield 15 mgs, 11.2 μmol (9% yield) of purified (1-40)αCOSNB.

Ligation to give the native-ligated monomer of HIV-1 PR was performed by reacting purified HIV-1 PR fragments (1-40)αCOSNB [4.3mgs, 0.93μmol] and (41-99) [3.0 mgs, 0.47μmol] in 6M GuHCl, 0.2M phosphate, pH=6.5 (0.5 ml) in the presence of 1-2% thiophenol at room temperature overnight. The progress of the reaction was monitored by reversed-phase analytical HPLC and electrospray MS. The product was purified by gel filtration and semipreparative HPLC (42-55%B in 30 minutes) to yield 1.7 mgs, (0.16 μmol, 33% yield based on 41-99) as a white solid after lyophilization. The purified ligated PR was charaterized by electrospray MS to give a mass of 10,706 ± 2 Da [calculated 10,706 Da (average isotope composition)].

2.4.6 Enzyme folding

Both HIV-1 PR monomers produced, either via stepwise methods or native chemical ligation, were folded against decreasing concentrations of guanidine hydrochloride (6M-0M) into dialysis buffer (25mM sodium acetate, pH=5.0 /20% glycerol for the stepwise 99-mer and 25mM sodium acetate, pH=5.0, 2mM DTT, 1mM EDTA/20% glycerol for native-ligated HIV-1 PR, both containing appropriate amounts (0.80 equivalents) of the competitive inhibitor, MVT-101) over a 2-day period. DTT and EDTA were added to the assay buffer for the native-ligated enzyme to prevent oxidation, as observed by MS analysis (+16 Da. to the MW of product). Following dialyisis, the sample was centrifuged to remove any precipitate and stored at 4°C. Concentrations of folded protein were determined by measuring the peak area obtained in an analytical HPLC run and comparing to the area obtained from a sample of denatured enzyme of known concentration. (The concentration of the denatured enzyme was determined by absorbance at 280 nm, using a molar extinction coefficient of 25,500 M^{-1} cm^{-1}[145]). The yield of folded native-ligated HIV-1 PR was 25% and that from stepwise synthesis was 10%.

2.4.7 Enzyme assays

Substrate specificity of the two enzymes was assessed by the cleavage of a synthetic peptide substrate (GHKARVL*AEAMSQVTNSATIM*MQRGNFRNQRK, where * designates HIV-1 PR processing sites) spanning the capsid/nucleocapsid cleavage site in the viral gag and gag-pol polyproteins.

Kinetic parameters were determined for the cleavage of the peptide substrate RNTATIM*MQRGNFR in the following assay buffers: for HIV-1 PR produced via stepwise methods, pH 5.5 buffer containing 100mM sodium acetate, 0.5M NaCl, 10% glycerol and bovine serum albumin (BSA) at 0.5 mg/ml. For the native-ligated HIV-1 PR, 2mM DTT and 1mM EDTA were added to the above assay buffer. DMSO was used to solubilize the synthetic substate. Because DMSO has been found to weakly inhibit the enzyme[40], its final assay concentration was kept standardized at 2%. Cleavage products were quantitated by HPLC with peak detection by UV absorbance (214 nm). Initial rates were determined by single time points of reactions in which consumption of substrate was <20% of the starting concentration (12-15 minute reaction time). Kinetic parameters were obtained by fitting data to the Michaelis-Menten equation using a nonlinear regression computer program[46].

References

(1). Debouck, C. *AIDS Res. Hum. Retroviruses* **1992**, *8*, 153-164.

(2). Meek, T. D.; Dreyer, G. B. *Ann. N. Y. Acad. Sci.* **1990**, *616*, 41-53.

(3). Tomasselli, A. G.; Howe, W. J.; Sawyer, T. K.; Wlodawer, A.; Heinrikson, R. L. *Chimica Oggi* **1991**, *9*, 6-27.

(4). Wlodawer, A. *Pharmacotherapy* **1994**, *14*, 9S-20S.

(5). Huff, J. R. *J. Med. Chem.* **1991**, *34*, 2305-2314.

(6). von der Helm, K.; Gurtler, L.; Eberle, J.; Deinhardt, F. *FEBS Lett.* **1989**, *247*, 349-52.

(7). Seelmeier, S.; Schmidt, H.; Turk, V.; von der Helm, K. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 6612-6.

(8). Katoh, I.; Yasunaga, T.; Ikawa, Y.; Yoshinaka, Y. *Nature* **1987**, *329*, 654-656.

(9). Ashorn, P.; McQuade, T. J.; Thaisrivongs, S.; Tomasselli, A. G.; Tarpley, W. G.; Moss, B. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 7472-6.

(10). McQuade, T. J.; Tomasselli, A. G.; Liu, L.; Karacostas, V.; Moss, B.; Sawyer, T. K.; Heinrikson, R. L.; Tarpley, W. G. *Science* **1990**, *247*, 454-456.

(11). Meek, T. D.; Dayton, B. D.; Metcalf, B. W.; Dreyer, G. B.; Strickler, J. E.; Gorniak, J. G.; Rosenberg, M.; Moore, M. L.; Magaard, V. W.; Debouck, C. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 1841-1845.

(12). Mous, J.; Heimer, E. P.; Le Grice, S. F. *J. Virol.* **1988**, *62*, 1433-6.

(13). Nutt, R. F.; Brady, S. F.; Darke, P. L.; Ciccarone, T. M.; Colton, C. D.; Nutt, E. M.; Rodkey, J. A.; Bennett, C. D.; Waxman, L. H.; Sigal, I. S. *Proc. Natl. Acad. Sci USA* **1988**, *85*, 7129-7133.

(14). Schneider, J.; Kent, S. B. H. *Cell* **1988**, *54*, 363-368.

(15). Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. *Science* **1989**, *246*, 1149-1152.

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

(17). Hilvert, D. *Chemistry and Biology* **1994**, *1*, 201-203.

(18). Muir, T. W.; Kent, S. B. H. *Curr. Opin. Biotech.* **1993**, *4*, 420-427.

(19). Baca, M.; Kent, S. B. H. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11638-11642.

(20). Cornish, V. W.; Mendel, D.; Schultz, P. G. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 621-633.

(21). Cornish, V. W.; Schultz, P. G. *Curr. Opin. Struct. Biol.* **1994**, *4*, 601-607.

(22). Kent, S. B. H. *Ann. Rev. Biochem.* **1988**, *57*, 957-989.

(23). Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180-193.

(24). Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776-779.

(25). Schnölzer, M.; Kent, S. B. H. *Science* **1992**, *256*, 221-225.

(26). Muir, T. M. *Structure* **1995**, *3*, 649-652.

(27). Wallace, C. J. A. *Current Opinion in Biotechnology* **1995**, *6*, 403-410.

(28). Baca, M.; Muir, T. W.; Schnölzer, M.; Kent, S. B. H. *J. Am. Chem. Soc.* **1994**, *117*, 1881-1887.

(29). Canne, L. E.; Ferré-D'Amaré, A. R.; Burley, S. K.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 2998-3007.

(30). Muir, T. W.; Williams, M. J.; Ginsberg, M. H.; Kent, S. B. H. *Biochemistry* **1994**, *33*, 7701-7708.

(31). Rose, K. *J. Am. Chem. Soc.* **1994**, *116*, 30-33.

(32). Strickler, J. E.; Gorniak, J.; Dayton, B.; Meek, T.; Moore, M.; Magaard, V.; Malinowski, J.; Debouck, C. *Proteins* **1989**, *6*, 139-154.

(33). Mildner, A. M.; Rothrock, D. J.; Leone, J. W.; Bannow, C. A.; Lull, J. M.; Reardon, I. M.; Sarcich, J. L.; Howe, W. J.; Tomich, C. C.; Smith, C. W.; Heinrikson, R. L.; Tomasselli, A. G. *Biochemistry* **1994**, *33*, 9405-9413.

(34). Rosé, J. R.; Salto, R.; Craik, C. S. *J. Biol. Chem.* **1993**, *268*, 11939-11945.

(35). Schecter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157-162.

(36). Sanchez-Pescador, R.; Power, M. D.; Barr, P. J.; Steimer, K. S.; Stempien, M. M.; Brown-Shimer, S. L.; Gee, W. W.; Renard, A.; Randolph, A.; Levy, J. A.; Dina, D.; Luciw, P. A. *Science* **1985**, *227*, 484-492.

(37). Sarin, V. K.; Kent, S. B. H.; Tam, J. P.; Merrifield, R. B. *Analytical Biochemistry* **1981**, *117*, 147-157.

(38). Oroszlan, S. In *Viral Proteinases as Targets For Chemotherapy*; H. Kräusslich; S. Oroszlan and E. Wimmer, Ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989; pp 87-100.

(39). Richards, A. D.; Phylip, L. H.; Farmerie, W. G.; Scarborough, P. E.; Alvarez, A.; Dunn, B. M.; Hirel, P. H.; Konvalinka, J.; Strop, P.; Pavlickova, L. *J. Biol. Chem.* **1990**, *265*, 7733-7736.

(40). Bergman, D. A.; Alewood, D.; Alewood, P. F.; Andrews, J. L.; Brinkworth, R. I.; Englebretsen, D. R.; Kent, S. B. H. *Letters in Peptide Science* **1995**, in press.

(41). Schnölzer, M.; Jones, A.; Alewood, P. F.; Kent, S. B. H. *Anal. Biochem.* **1992**, *204*, 335-343.

(42). Yamashiro, D.; Li, C. H. *Int. J. Pept. Protein Res.* **1988**, *31*, 322-334.

(43). Canne, L. E.; Walker, S. M.; Kent, S. B. H. *Tetrahedron Letters* **1995**, *36*, 1217-1220.

(44). Blake, J. *Int. J. Pept. Protein Res.* **1981**, *17*, 273-274.

(45). Gill, S. C.; von Hippel, P. H. *Anal. Biochem.* **1989**, *182*, 319-326.

(46). Duggleby, R. G. *Comput. Biol. Med.* **1984**, *14*, 447-455.

Chapter III

Analysis of Drug Resistant Mutants of the HIV-1 PR Prepared by Total Chemical Synthesis

3.1 Introduction

The genome of the human immunodeficiency virus type 1 (HIV-1), the causative agent of acquired immunodeficiency syndrome (AIDS)[1-3], encodes a protease (PR) which acts to process the viral gag and gag-pol polyprotein precursors into the structural proteins and enzymes of the viral particle[4-6]. HIV-1 PR activity is essential for the formation of infectious viral particles[7]. Engineered viruses that express inactive protease give rise to immature virions which are incapable of initiating new cycles of infection[7,8]. The essential nature of HIV-1 PR in the viral life cycle makes it an important target for chemotherapeutic intervention and, as such, there has been an enormous amount of research activity aimed at developing inhibitors of the HIV-1 PR as anti-AIDS drugs[9-15]. Several protease inhibitors, derived from substrate-based or structure-based approaches, are now in clinical trials and have shown considerable success in suppressing HIV replication[14-16].

However, as seen with inhibitors of reverse transcriptase (RT), the spectre of resistance has also emerged during in human trials with the protease inhibitors[17-21]. The development of resistance to protease inhibitors may severely limit their potential effectiveness as antivirals[21] and is thus of great concern[22]. The mechanisms by which protease-resistant strains of HIV-1 emerge needs to be considered in the design of protease inhibitors and, as such, a current major research objective in this area is to understand the molecular basis of resistance to protease inhibitors[23-26]. In this chapter, we describe kinetic and inhibition studies on inhibitor-resistant forms of the HIV-1 PR identified as arising from the most advanced protease inhibitor, Ro31-8959[14,27-29]. These "mutant" forms of the HIV-1 PR were prepared by

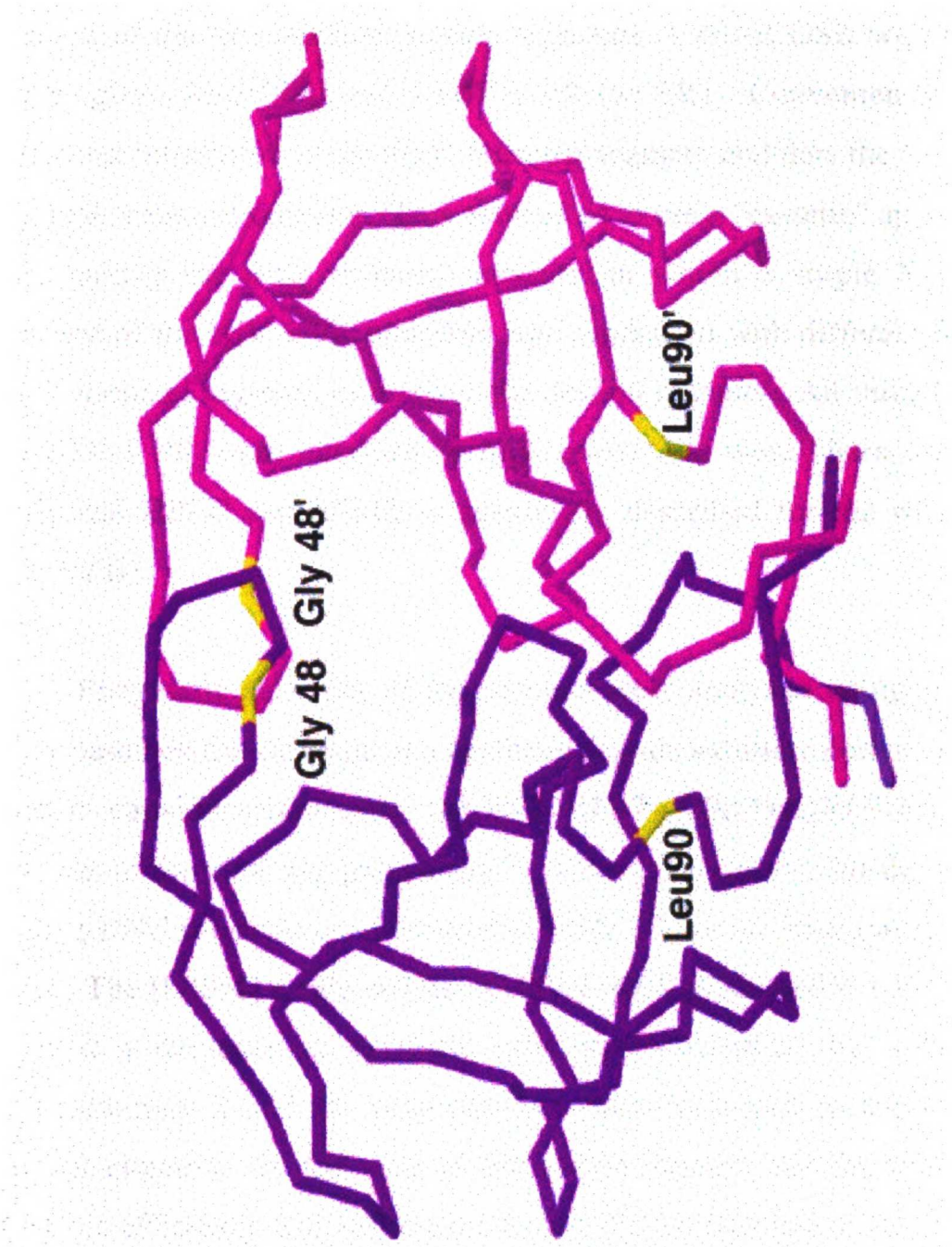
total chemical synthesis and evaluated against two compounds as representative examples of canonical classes of inhibitors. In addition, these data were used as a baseline to evaluate the comparative properties of a pair of HIV-1 PR inhibitors developed by comparative studies against both HIV-1 and feline immunodeficiency virus (FIV) proteases.

3.2 Results

The variants of the HIV-1 PR molecule synthesized were based on those identified in resistant strains of the virus with decreased susceptibility to Ro 31-8959 during clinical trials[28-30]. The resistance mutants derived from this inhibitor were chosen for investigation since Ro 31-8959 was one of the first protease inhibitors designed and has just recently been approved for clinical use. The sequences of the most common Ro 31-8959-resistant HIV-1 PR variants have been found to contain two amino acid substitutions, a Gly to Val exchange at position 48 and a Leu to Met exchange at position 90[28]. As indicated in Figure 3.1, Gly 48 is positioned in the region of the PR known as the "flaps", which are flexible β-hairpin structures that cover the substrate/inhibitor binding site in the HIV-1 PR[31]. L90 sits in the interior of the molecule and has no obvious contact with inhibitor or substrate. We chose to study the [L90M] HIV-1 PR and the double mutant [G48V/L90M] HIV-1 PR variants because these had been associated with drug resistance in clinical trials[32].

In order to characterize the kinetic properties and susceptibility to inhibiton of these variants, the enzymes were prepared by total chemical synthesis. A native sequence control ("wild type") molecule was also

Figure 3.1 Cα backbone structure of HIV-1 PR indicating the positions of drug-resistant mutations. The mutations are highlighted in yellow at positions 48 and 90. The figure was prepared using the program RasMol v.2.6 (Roger Sayle, Biomolecular Structure, Glaxo Research and Development, Middlesex, UK.).


Gly 48
Gly 48'
Leu90
Leu90'

synthesized for comparison. Synthesis of the three molecular forms of HIV-1 PR was carried out by native chemcial ligation. (Refer to Figure 2.5 in Chapter II for the synthetic strategy). This involved the chemoselective reaction of the two purified peptide segments (1-40)αCOSR with Cys(42-99) to give [Aba^{67,95},Cys^{41}] HIV-1 PR (wt PR). Conveniently, all the mutations introduced are in the C-terminal segment and thus the N-terminal segment remained constant. This allowed us to use a "cassette" approach for constructing the drug resistant variants, in which a single N-terminal segment of the HIV-1 PR molecule was condensed with different forms of the C-terminal segment to produce the desired mutants. All three forms of the HIV-1 PR were obtained in high purity and were characterized by analytical HPLC and ESMS as previously described for the wt PR (see Chapter II).

Enzymatic activities of the variants were analyzed using an HPLC assay based on the cleavage of a synthetic peptide substrate derived from the second capsid/nucleocapsid (p1/NC) [RNTATIM*MQRGNFR-amide] junction of the viral gag polyprotein[33]. Kinetic parameters for the cleavage of the p1/NC substrate are presented in Table 3.1 for all three forms of HIV-1 PR. The [L90M] single and the [G48V/L90M] double HIV-1 PR variants showed a decrease in catalytic efficiency (kcat/Km) by a factor of approximately 2 and 15, respectively, when compared to the wild-type control enzyme. In the case of the L90M variant, the decrease in the catalytic efficiency was influenced mainly by the increase in Km. For the [G48V/L90M] double mutant, both kcat and Km were equally affected. Therefore, the inhibitor-resistant mutations produce enzymes that are significantly less active than the wild type enzyme.

Table 3.1 Kinetic Constants for Drug Resistant Mutants

Proteases	K_m (μM)	k_{cat} (sec^{-1})	k_{cat}/K_m ($\mu M^{-1}sec^{-1}$)
wt*	141	4.4	0.031
L90M	232	4	0.017
G48V/L90M	447	1	0.002

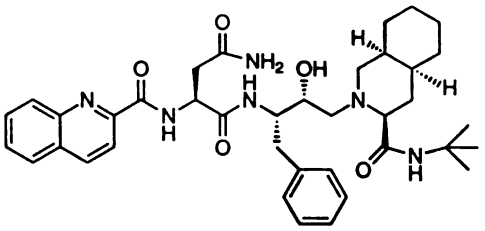
Kinetic parameters were determined for the cleavage of the synthetic peptide substrate RNTATIM*MQRGNFR-amide, where * indicates the site of cleavage. The enzyme concentrations used in assays was approximately 11nM for wt PR, 12 nM for L90M PR, and 30nM for G48V/L90M PR. The substrate concentration varied between 50-900μM in the assays. wt PR and variants were assayed in a buffer containing 100mM sodium acetate, 2mM DTT, 1mM EDTA, 0.5 M NaCl, 10% glycerol and 0.5 mg/ml BSA at pH=5.5, 37°C. *wt = [$Aba^{67,95}$, Cys^{41}] HIV-1 PR.

Inhibition constants (Ki) were compared for the wild-type HIV-1 PR and two drug-resistant variants using the inhibitors shown in Figure 3.2. The inhibition constants for the enzymes with the four different inhibitors are shown in Table 3.2. Ro 31-8959 and Slee 1 displayed similar inhibition effects, with an increase in Ki values of approximately two-fold for the L90M variant and approximately 14-fold for the [G48V/L90M] double mutant. These effects were almost identical to the changes observed for the catalytic efficiency, kcat/Km (Table 3.1). DMP 323 showed the most interesting results, in that identical levels of inhibition for the wt and double mutant were observed. A close to two-fold increase in Ki, however, was observed for the L90M variant. Even so, DMP 323 remained a highly potent (nanomolar) inhibitor of all three enzyme forms. The compounds Slee 1 and Slee 2 were of equal inhibitory activity against the wild-type enzyme. However, Slee 2, the inhibitor evaluated against both HIV and FIV, had significantly better (approximately 3-fold) Ki values than the Slee 1 inhibitor for both resistant variants of the HIV-1 PR. Against the double mutant, Slee 2 showed a 5-fold increase in Ki when compared to wild-type. An approximately 15-fold increase in Ki (compared to wild-type) was observed with Slee 1 for the same enzyme.

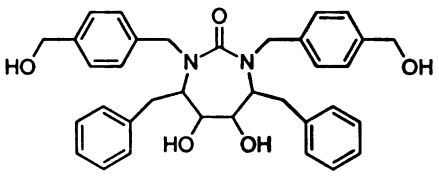
3.3 Discussion

Inhibitors of HIV-1 PR are of considerable promise as therapeutic agents in the treatment of AIDS. However, the clinical effectiveness of these compounds may be compromised by the emergence of drug-resistant variants, as shown with inhibitors of HIV-1 reverse transcriptase. In this study, we chose to evaluate drug-resistant variants with decreased sensitivity

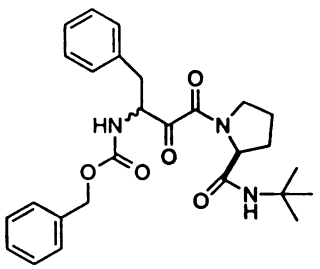
Figure 3.2 Inhibitor Structures



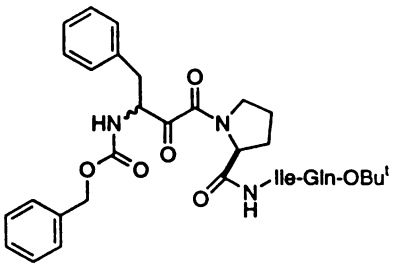
Ro 31-8959



DMP-323



Slee 1



Slee 2

Table 3.2 Inhibitory Constants [K_i (nM)] for Drug Resistant Mutants

Proteases	Ro31-8959†	DMP-323†	Slee 1§	Slee 2§
wt	0.68	0.60	234	215
L90M	1.10	1.14	414	163
G48V/L90M	14	0.60	3580	1140

Ki values were determined in a chromatographic assay using the same conditions as for the analysis of kinetic parameters. The enzyme concentations used in assays was approximately 11 nM for wt PR, 12nM for L90M, and 30nM for G48V/L90M. All enzymes were assayed in a buffer containing 100mM sodium acetate, 2mM DTT, 1mM EDTA, 0.5 M NaCl, 10% glycerol and 0.5 mgs/ml BSA at pH=5.5, 37°C. †Inhibition constants were calculated from Henderson plots. §Inhibition constants were determined using Dixon plots.

to the inhibitor Ro 31-8959. These mutants were chosen because Ro 31-8959 was one of the first protease inhibitors developed and has just been approved for clinical use as an anti-AIDS therapeutic. [L90M] and [G48V/L90M] HIV-1 PR variants were specifically chosen for evaluation because these inhibitor-resistant variants had been associated with drug resistance in clinical trials[28,32].

The [L90M] HIV-1 PR variant has been observed more commonly in vivo[32]. Also, in the case of some of the drug-resistant variants of HIV-1 PR observed, the amino acid changes found correspond to the structurally aligned residues found in FIV-PR[34]. FIV-PR contains a methionine residue (M107) in the structurally analogous position for L90 in HIV-1 PR. Since FIV causes an immunodeficiency disease in cats which is remarkably similar to AIDS in humans, this may provide a basis for which to test inhibitors of HIV-1 PR in the FIV model[35] and potentially contribute to an understanding of the mechanism of drug resistance. [The G48V/L90M] double mutant has been reported to be the most resistant variant of Ro 31-8959 observed[28,32]. Although this variant has been found to be rare in vivo, it seemed appropriate to test this mutant to see what type of inhibitory effect the Slee inhibitors would have, since one of the rationales behind their development was to see if they would show improved relative activity against inhibitor resistant forms of HIV-1 PR.

The 3 different classes of inhibitors used in this study were chosen because they were developed using different rationales (Figure 3.2). Ro 31-8959, one of the first high potency PR inhibitors developed, is a substrate-based inhibitor which contains a hydroxyethylamine transition state mimetic

and was designed by analogy with renin inhibitors. This inhibitor was designed based on Class 1 substrates which contain the cleavage site Phe /Tyr*Pro. It is believed that inhibitors designed based on this class of substrates would have greater selectivity for viral aspartyl proteases over their cell-encoded counterparts. This compound has just been approved by the FDA (as of November 1995) for use as an anti-AIDS therapeutic.

The inhibitors DMP-323 and Slee inhibitors were designed by structure-based methods. DMP-323 is a structure-based, C2-symmetric, cyclic urea inhibitor. One of the key features of the DMP 323 structure is the cyclic urea carbonyl oxygen that mimics the hydrogen-bonding features of a key structural water molecule, designated as water 301. Water 301 is tetrahedrally coordinated with two hydrogen bond donors from the enzyme (NH of Ile 50) and two acceptors from the inhibitors (usually backbone carbonyl groups). This water molecule is observed only in the crystal structures of inhibitor-bound HIV-PR and not in the pepsin-like enzymes.

The Slee inhibitors, which contain an α-keto amide core structure, were designed to serve as mechanism-based inhibitors. The α-keto amide core structure in these inhibitors is not hydrated until bound to the active site of the HIV-1 PR, which then gives rise to the active species. As previously mentioned, the rationale for the development of these inhibitors was to design compounds which would inhibit both the HIV and FIV proteases since they are both mechanistically identical. In addition, it was hoped that these compounds would be effective against inhibitor-resistant HIV-1 PR and thus make it possible to use FIV-infected cats as a model system on which to test inhibitors.

The "wild-type" PR used in this study ([Aba^{67,95},Cys^{41}] HIV-1 PR) possessed full enzymatic activity, with kinetic values comparable to those of native-backbone HIV-1 PR produced either by recombinant or synthetic means[36]. When compared to wild-type PR, the Km values obtained for the single and double mutant PR's increased approximately 2-and 3-fold, respectively. In addition, the catalytic efficiency (kcat/Km) for the enzymes decreased by 2-fold for [L90M] HIV-1 PR and 15-fold for [G48V/L90M] HIV-1 PR. In the case of the single mutant, most of the reduction in the catalytic efficiency is due to the increase in Km (kcat was identical to wt PR). This result was surprising, since the L90M mutation is remote from the enzyme's substrate binding site (Figure 3.1). This "non-active site mutant" of the HIV-1 PR clearly has an indirect effect on substrate binding[24]. The [G48V/L90M] double mutant exhibited changes in both kcat and Km, which contributed to the reduced catalytic efficiency(kcat/Km) of this variant. The changes observed can likely be attributed to the increased steric interaction introduced in the substrate binding pocket by the presence of the valine side chain, which may not allow the substrate to bind as efficiently.

The inhibitor Ro 31-8959 showed decreased levels of inhibition with the single and double mutant as compared to wt PR, with increases in Ki values of approximately 2- and 20-fold respectively. These results were expected based on available literature for these mutants[28,32]. What was not expected were the results obtained with the inhibitor DMP 323, which possessed approximately the same inhibitory potency with all variants studied (Table 3.2). With the double mutant, G48V seems to be a compensatory change for DMP-323 inhibition, but is devastating for inhibition by Ro 31-8959. One explanation for this effect is based on the

structures of the two inhibitors and their interactions with the HIV-1 PR. Based on known structures of the HIV-1 PR with various substrate-based inhibitors, it was observed that G48 is poised to interact with the P3 and P3' positions of a substrate/inhibitor[28,37]. Ro 31-8959 contains a quinoline-2-carbonyl group in its P3 position, which interacts with the S3 binding pocket of the enzyme, whereas the inhibitor DMP-323 does not extend into this binding pocket (this inhibitor extends from the S2-S2' sites). Therefore, the Gly to Val exchange at position 48 may directly interfere with Ro 31-8959 binding by steric interference of the valine side chain. In general, however, it was found that both of these high-potency inhibitors still retain a decent level of inhibiton in the low nanomolar range--Table 3.2. This phenomenon has also been observed in other studies of drug-resistant HIV-1 PR variants[15,38-40].

The two Slee inhibitors gave interesting results. Both Slee1 and Slee 2 inhibited the wild-type HIV-1 PR to approximately the same degree, with Ki's of 234 and 215nM, respectively. However Slee 2, which was designed for better inhibition of FIV-PR, had a clearly better inhibitory effect with both resistant variants than Slee 1. Although these compounds were not stellar in their inhibitory potency, they clearly represent a good beginning in the design of compounds which will inhibit the HIV-1 PR even in the face of resistance.

Some kinetic and inhibition data has been published for a few of the inhibitor-resistant HIV-1 PR mutants[15,38-40]. Although the inhibitor Ro 31-8959 was used in all but one of these studies, none included data for any of the variants generated by prolonged exposure to this compound. The data

inhibitor-resistant HIV-1 PR variants generated by one of the most advanced protease inhibitors in clinical studies.

Total chemical synthesis was used to construct and drug-resistant variants of the HIV-1 PR in a modular fashion. The utility of this method has been demonstrated in the present study, where resistant mutants were constructed and analyzed against 3 different classes of inhibitors. The results indicate that it may be possible to combat resistance by either further developing mechanism-based inhibitors[21,22,34], such as the Slee class, or by the use of potent protease inhibitors given in a sequential manner. In addition, this method may aid in determining the molecular basis of drug resistance by providing a straightforward method to the production of resistant variants of the HIV-1 PR for biochemical and structural characterization.

3.4 Experimental section

3.4.1 Materials and methods

The inhibitor DMP-323 was obtained from DuPont-Merck Pharmaceuticals (Wilmington DE.). Ro 31-8959 was obtained from Roche Products LTD, (Welwyn Garden City, UK). Boc-amino acids and 2-(1-H-benzotriazol-1-yl)-1,1,3,3,-tetramethyluronium hexafluorophosphate (HBTU) were obtained from Nova Biochem (San Diego, CA.). Boc-Phe-OCH_2Pam-resin and diisopropylethylamine were purchased from Applied Biosysyems, Inc. (Foster City, CA.). Aminomethylcopoly(styrene-divinylbenzene) resin was either purchased from Applied Biosystems, Inc.

or prepared in house (prepared by L. E. Canne, TSRI). N,N-Dimethylformamide (Synthesis-grade) was obtained from either Mallinckrodt Chemical Co. (Paris, KY) or Fisher. AR-grade methylene chloride was obtained from Fisher. TFA was purchased from Halocarbon (River Edge, N. J.) and HF obtained from Matheson Gas. HPLC-grade acetonitrile was obtained from either EM Science or Fisher. Vydac HPLC columns were purchased form Western Analytical Co. (Temecula, CA.). All other reagents were AR grade or better, and purchased from either Aldrich Chemical or Fisher.

Machine-assisted solid-phase peptide syntheses were performed as previously described (Chapter II).

3.4.2 Solid-phase peptide synthesis

All peptides were synthesized in a stepwise fashion according to the *in situ* neutralization/HBTU activation protocols for machine-assisted Boc solid-phase chemistry[41]. Side chain protection was as follows: Arg(*p*-toluenesulfonyl), Asn(xanthyl), Asp(*O*-cyclohexyl), Glu(*O*-cyclohexyl), His(Dnp), Lys(2-Cl-Z), Ser(benzyl), Trp(formyl), Thr(benzyl), and Tyr(2-Br-Z). Gln and Met (and, when synthesizing thioacid peptides Trp) were used without side chain protection. Coupling yields were determined by the quantitative ninhydrin assay on peptide-resin sapmles taken under machine control during the chain assembly[41,42]. The HIV-1 PR (SF2 isolate) C-terminal fragments, corresponding to residues (41-99) and incorporating the appropriate changes, were prepared on a Boc-Phe-4-(carboxamidomethyl)benzyl ester-copoly(styrene-1%-divinylbenzene) resin

(Boc-Phe-OCH_2-Pam-resin). The N-terminal fragment corresponding to HIV-1 PR (1-40) was constructed on a 4-[α-(Boc-Gly-S)benzyl]-phenoxyacetamidomethyl resin, which yields a peptide with a C-terminal Gly-αCOSH post cleavage[43-45]. Upon completion of the chain assembly, peptides were deprotected as necessary and cleaved from the resin by treatment with liquid HF, containing 10% *p*-cresol, for 1 hour at 0 °C. After evapoartion of HF under reduced pressure, crude peptide products were precipitated in ice-cold anhydrous diethyl ether, dissolved in 50% aqueous acetic acid, and either purified immediately or after prior lyophilization.

3.4.3 Purification and characterization of peptide segments

Crude lyophilized peptides were dissolved in either acidic aqueous buffers or 50% aqueous acetic acid and purified by preparative or semipreparative HPLC. Purified peptides were stored as lyophilized powders at -20°C. The purified peptide segments were characterized by electrospray MS, and all possessed observed masses within experimental error of the calculated masses.

3.4.4 Preparation of wt HIV-1 PR and inhibitor-resistant variants

[$Aba^{67,95}$, Cys^{41}] HIV-1 PR (wt HIV-1 PR) and resistant variants were constructed using native chemical ligation as described in Chapter II. All enzymes possessed the correct masses as determined by ESMS and were within experimental error of the calculated masses. For wt HIV-1 PR, 10,706 ± 2 Da [calculated 10,706 Da (average isotope composition)]; for [L90M] HIV-1 PR, 10,724 ± 3 Da [calculated 10,724 Da (average isotope

composition)]; for [G48V/L90M] HIV-1 PR, 10,766 ± 3 Da [calculated 10,766 Da (average isotope composition)].

3.4.5 Enzyme folding

All enzymes produced were folded against decreasing concentrations of guanidine hydrochloride (6M-0M) into dialysis buffer (25mM sodium acetate, pH=5.0, 2mM DTT, 1mM EDTA/20% glycerol and contained appropriate amounts (0.80 equivalents based on HIV-1 PR) of the competitive inhibitor, MVT-101) over a 2-day period. Following dialyisis, the sample was centrifuged to remove any precipitate and stored at 4°C. Concentrations of folded protein were determined by measuring the peak area obtained in an analytical HPLC run and comparing to the area obtained from a sample of denatured enzyme of known concentration. (The concentration of the denatured enzyme was determined by absorbance at 280 nm, using a molar extinction coefficient of 25,500 M^{-1} cm^{-1}[146]).

3.4.6 Enzyme assays

Kinetic parameters were determined for the cleavage of the peptide substrate RNTATIM*MQRGNFR-amide in the following assay buffer for the native-ligated PR's: pH 5.5 buffer containing 100mM sodium acetate, 2mM DTT, 1mM EDTA, 0.5M NaCl, 10% glycerol and bovine serum albumin (BSA) at 0.5 mg/ml. DMSO was used to solubilize the synthetic peptide substrate. Because DMSO has been found to weakly inhibit the enzyme[36], its final assay concentration was kept standardized at 2%. Cleavage products were quantitated by HPLC with peak detection by UV

absorbance (214 nm). Initial rates were determined by single time points of reactions in which consumption of substrate was <20% of the starting concentration (15-30 minute reaction time). Kinetic parameters were obtained by fitting data to the Michaelis-Menten equation using a nonlinear regression computer program[47].

Inhibition constants of the drug resistant variants and wt enzyme for the various inhibitors were assessed in the chromatographic assay as above at the following fixed substrate concentrations: wt and L90M, 300mM; G48V/L90M, 550mM. For inhibitors Slee 1 and Slee 2, IC_{50} values were determined by least squares analysis using Dixon plots[48], and these values were converted into inhibitor dissociation constants (K_i) by the use of the expression $K_i=IC_{50}/(1+[S]/K_m)$. For the tight-binding inhibitors DMP-323 and Ro 31-8959, Ki values were determined by the method of Henderson[49].

References

(1). 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.; White, G.; Foster, P.; Markham, P. D. *Science* **1984**, *224*, 500-503.

(2). 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.; Montagnier, L. *Science* **1983**, *220*, 868-871.

(3). Gallo, R. C.; Sarin, P. S.; Gelmann, E. P.; Robert-Guroff, M.; Richardson, E.; Kalyanaraman, V. S.; Mann, D.; Sidhu, G. D.; Stahl, R. E.; Zolla-Pazner, S.; Leibowitch, J.; Popovic, M. *Science* **1983**, *220*, 865-867.

(4). 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.; Sigal, I. S. *Biochem. Biophys. Res. Commun.* **1988**, *156*, 297-303.

(5). Kramer, R. A.; Schaber, M. D.; Skalka, A. M.; Ganguly, K.; Wong-Staal, F.; Reddy, E. P. *Science* **1986**, *231*, 1580-1584.

(6). Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K. J.; Starcich, B.; Josephs, S. F.; Doran, E. R.; Rafalski, J. A.; Whitehorn, E. A.; Baumeister, K.; Ivanoff, L.; Petteway, S. R.; Pearson, M. L.; Lautenberger, L. A.; Papas, T. S.; Ghrayeb, J.; Chang, N. T.; Gallo, R. C.; Wong-Stahl, F. *Nature* **1985**, *313*, 277-284.

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

(8). Mous, J.; Heimer, E. P.; Le Grice, S. F. *J. Virol.* **1988**, *62*, 1433-6.

(9). Martin, J. A. *Antiviral Res.* **1992**, *17*, 265-78.

(10). Tomasselli, A. G.; Howe, W. J.; Sawyer, T. K.; Wlodawer, A.; Heinrikson, R. L. *Chimica Oggi* **1991**, *9*, 6-27.

(11). Wlodawer, A.; Erickson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543-585.

(12). Debouck, C. *AIDS Res. Hum. Retroviruses* **1992**, *8*, 153-164.

(13). Lam, P. Y.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; Chang, C. H.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R.; Erickson-Viitanen, S. *Science* **1994**, *263*, 380-384.

(14). Roberts, N. A.; Martin, J. A.; Kinchington, D.; Broadhurst, A. V.; Craig, J. C.; Duncan, I. B.; Galpin, S. A.; Handa, B. K.; Kay, J.; Krohn, A.; Lambert, R. W.; Merrett, J. H.; Mills, J. S.; Parkes, K. E. B.; Redshaw, S.; Ritchie, A. J.; Taylor, D. L.; Thomas, G. J.; Machin, P. J. *Science* **1990**, *248*, 358-361.

(15). 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.; et, a. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 4096-4100.

(16). 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.; Wideburg, N. E.; Saldivar, A.; Ruiz, L.; Kati, W.; Sham, H. L.; Robins, T.; Stewart, K. D.; Hsu, A.; Plattner, J. J.; Leonard, J. M.; Norbeck, D. W. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2484-2488.

(17). el-Farrash, M. A.; Kuroda, M. J.; Kitazaki, T.; Masuda, T.; Kato, K.; Hatanaka, M.; Harada, S. *J. Virol.* **1994**, *68*, 233-9.

(18). Ho, D. D.; Toyoshima, T.; Mo, H.; Kempf, D. J.; Norbeck, D.; Chen, C. M.; Wideburg, N. E.; Burt, S. K.; Erickson, J. W.; Singh, M. K. *J. Virol.* **1994**, *68*, 2016-20.

(19). Kaplan, A. H.; Michael, S. F.; Wehbie, R. S.; Knigge, M. F.; Paul, D. A.; Everitt, L.; Kempf, D. J.; Norbeck, D. W.; Erickson, J. W.; Swanstrom, R. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5597-601.

(20). Otto, M. J.; Garber, S.; Winslow, D. L.; Reid, C. D.; Aldrich, P.; Jadhav, P. K.; Patterson, C. E.; Hodge, C. N.; Cheng, Y. S. *Proc. Natl. Acad. Sci USA* **1993**, *90*, 7543-7547.

(21). 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.; Titus, D.; Yang, T.; Teppler, H.; Squires, K. E.; Deutsch, P. J.; Emini, E. A. *Nature* **1995**, *374*, 569-571.

(22). Richman, D. D. *Nature* **1995**, *374*, 494.

(23). Baldwin, E. T.; Bhat, T. N.; Liu, B.; Pattabiraman, N.; Erickson, J. W. *Nature Structural Biology* **1995**, *2*, 244-249.

(24). Erickson, J. W. *Nature Structural Biology* **1995**, *2*, 523-529.

(25). Baldwin, E. T.; Bhat, T. N.; Gulnik, S.; Liu, B.; Topol, I. A.; Kiso, Y.; Mimoto, T.; Mitsuya, H.; Erickson, J. W. *Structure* **1995**, *3*, 581-590.
(26) Collins, J. R.; Burt, S. K.; Erickson, J. W. *Nature Structural Biology* **1995**, *2*, 334-338.

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

(28). Jacobsen, H.; Yasargil, K.; Winslow, D. L.; Craig, J. C.; Kröhn, A.; Duncan, I. B.; Mous, J. *Virology* **1995**, *206*, 527-534.

(29). Eberle, J.; Bechowsky, B.; Rose, D.; Hauser, U.; von der Helm, K.; Gürtler, L.; Nitschko, H. *AIDS Research and Human Retroviruses* **1995**, *11*, 671-676.

(30). Tisdale, M.; Myers, R.; Parry, N. R.; Oliver, N.; Mashera, B.; Blair, E. *Abstracts of the Third International Workshop on HIV Drug Resistance* **1994**, p. 14.

(31). Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. *Science* **1989**, *246*, 1149-1152.

(32). Mellors, J. W.; Larder, B. A.; Schinazi, R. F. *International Antiviral News* **1995**, *3*, 8-13.

(33). Sanchez-Pescador, R.; Power, M. D.; Barr, P. J.; Steimer, K. S.; Stempien, M. M.; Brown-Shimer, S. L.; Gee, W. W.; Renard, A.; Randolph, A.; Levy, J. A.; Dina, D.; Luciw, P. A. *Science* **1985**, *227*, 484-492.

(34). Slee, D. H.; Laslo, K. L.; Elder, J. H.; Ollmann, I. R.; Gustchina, A.; Kervinen, J.; Zdanov, A.; Wlodawer, A.; Wong, C. H. *JACS* **1995**, in press.

(35). North, T. W.; LaCasse, R. A. *Nature Medicine* **1995**, *1*, 410-411.

(36). Bergman, D. A.; Alewood, D.; Alewood, P. F.; Andrews, J. L.; Brinkworth, R. I.; Englebretsen, D. R.; Kent, S. B. H. *Letters in Peptide Science* **1995**, in press.

(37). Moody, M. D.; Pettit, S. C.; Shao, W.; Everitt, L.; Loeb, D. D.; Hutchison, I., C. A.; Swanstrom, R. *Virology* **1995**, *207*, 475-485.

(38). Gulnik, S. V.; Suvorov, L. I.; Liu, B.; Yu, B.; Anderson, B.; Mitsuya, H.; Erickson, J. W. *Biochemistry* **1995**, *34*, 9282-9287.

(39). Lin, Y.; Lin, X.; Hong, L.; Foundling, S.; Heinrikson, R. L.; Thaisrivongs, S.; Leelamanit, W.; Raterman, D.; Shah, M.; Dunn, B. M.; Tang, J. *Biochemistry* **1995**, *34*, 1143-1152.

(40). Sardana, V. V.; Schlabach, A. J.; Graham, P.; Bush, B. L.; Condra, J. H.; Culberson, J. C.; Gotlib, L.; Graham, D. J.; Kohl, N. E.; LaFemina, R. L.; Schneider, C. L.; Wolanski, B. S.; Wolfgang, J. A.; Emini, E. A. *Biochemistry* **1994**, *33*, 2004-2010.

(41). Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180-193.

(42). Sarin, V. K.; Kent, S. B. H.; Tam, J. P.; Merrifield, R. B. *Analytical Biochemistry* **1981**, *117*, 147-157.

(43). Yamashiro, D.; Li, C. H. *Int. J. Pept. Protein Res.* **1988**, *31*, 322-334.

(44). Canne, L. E.; Walker, S. M.; Kent, S. B. H. *Tetrahedron Letters* **1995**, *36*, 1217-1220.

(45). Blake, J. *Int. J. Pept. Protein Res.* **1981**, *17*, 273-274.

(46). Gill, S. C.; von Hippel, P. H. *Anal. Biochem.* **1989**, *182*, 319-326.

(47). Duggleby, R. G. *Comput. Biol. Med.* **1984**, *14*, 447-455.

(48). Dixon, M. *Biochem. J.* **1953**, *55*, 170-171.

(49). Henderson, P. J. F. *Biochem. J.* **1972**, *127*, 321-333.

Chapter IV

Re-engineering the Binding Pocket of the HIV-1 Protease Through Total Chemical Synthesis

4.1 Introduction

The HIV-1 protease (HIV-1 PR) is essential for viral replication and for the production of infectious virions[1,2]. The essential nature of the protease in the viral life cycle has made it an important target for the design of antiviral drugs for AIDS. The HIV-1 PR is a member of the aspartyl proteinase family, but unlike the cell-encoded members of that class of enzymes--such as renin and pepsin--the active form of the HIV-1 PR is a homodimer[3-5]. In the enzyme molecule, the two 99-amino acid monomers of the enzyme are related by an approximate twofold axis of symmetry, with the active site of the enzyme lying at the dimer interface. Each monomer contributes one of the two conserved active site aspartyl residues to the catalytic machinery and the substrate binding site is formed by residues from both momomers[6].

Because of the importance of the HIV-1 PR for drug design, it has become one of the best characterized enzymes, particularly with regard to structural information[7,8]. Structural analyses of inhibitor complexes of HIV-1 PR show that substrate binds in an extended conformation, which is maintained by a network of hydrogen bonds between the main chain amide and carbonyl groups of the substrate and the protease[8-11]. The HIV-1 PR recognizes a linear array of 6 or 7 amino acids in the polypeptide substrates and is thus able to cleave large proteins at a few well-defined sites[12,13]. Most importantly, interaction with an array of amino acids contained in a particular substrate allows the HIV-1 PR to recognize precise cleavage sites in the diverse set of sequences observed in the eight different gag and

gag/pol polyprotein cleavage sites which are processed by the enzyme during viral maturation[12].

Thus, it appears that overall choice of a substrate is dictated by the sum of all of the interactions of the enzyme subsites with the substrate[14]. In general, however, the rules governing substrate recognition are not well understood. Therefore, we set out to test our understanding of the structural features of the HIV-1 PR protein molecule which are involved in the recognition and cleavage of specific substrates. This chapter describes the use of total chemical synthesis to re-engineer the S1/S1' binding pockets of the enzyme molecule and the resulting effects on substrate specificity.

4.2 Results and discussion

4.2.1 Design of HIV-1 PR analogues

Based on crystal structure analysis of the HIV-1 PR[9], a proline residue (Pro $^{81/181}$) was shown to line the S1, S1' substrate binding pockets, as shown in Figure 4.1. Substituents introduced at the 4-*cis* position on the cyclic side chain of proline should point directly into the S1/S1' substrate binding pockets. We hoped to make use of this predictable molecular geometry to introduce novel functionality into the S1/S1' binding pockets and, as a result, to alter the residues accepted in P1 and P1' positions of the substrate. Thus, proline analogues containing guanidino and guanidinomethyl moieties at the 4-*cis* position were prepared and

Figure 4.1 Connolly surface of the substrate binding pocket of HIV-1 PR modelled based on the crystal structure of the HIV-1 PR complexed to the inhibitor MVT-101[9] using the computer program Insight II (Biosym Technologies, Inc., San Diego, CA.). The inhibitor MVT-101 (N-acetyl-Thr-Ile Nle-Ψ[CH_2-NH]-Nle-Gln-Arg-amide) is highlighted in pink. Proline 81 and 181, which line the S1 and S1' binding site, are shown in white.

Pro 81
P1'
P1
Pro 181

incorporated at position 81 in the 99 residue HIV-1 PR monomer by total chemical synthesis.

4.2.2 Synthesis of proline analogues

The 4-*cis*-(guanidino) and (guanidinomethyl)-containing proline derivatives, respectively designated as Pro1 and Pro2, are shown in Figure 4.2 together with the synthetic strategy used in their preparation. These compounds were generated from commercially available, enantiomerically pure *trans*-4-hydroxyproline in 5 steps in good overall yield and high purity by the method of Webb and Eigenbrot[15]. The two proline analogues can also be considered conformationally-constrained arginine analogues, and were α-N-Boc and N^G-tosyl protected for direct introduction into peptides via Boc solid-phase peptide synthesis.

4.2.3 Total chemical synthesis of HIV-1 PR analogues

In order to construct the HIV-1 PR analogues, we used the chemical ligation strategy previously described by Schnölzer and Kent[16] (Figure 4.3). Two unprotected peptide segments were prepared by optimized SPPS[17]. HIV-1 PR (1-51)αCOSH, containing a C-terminal α-thioacid, and BrAc(53-99) HIV-1 PR, containing an N-terminal α-bromoacetyl moiety, were chemoselectively ligated to generate a backbone thioester bond between Gly^{51} and Gly^{52} instead of the native amide bond. This ($[NHCH_2COSCH_2CO]^{51-52}$) HIV-1 PR analogue had previously been found to possess the normal enzymatic properties of the native HIV-1 PR molecule[16]. The two proline derivatives Pro1 and Pro2 were separately

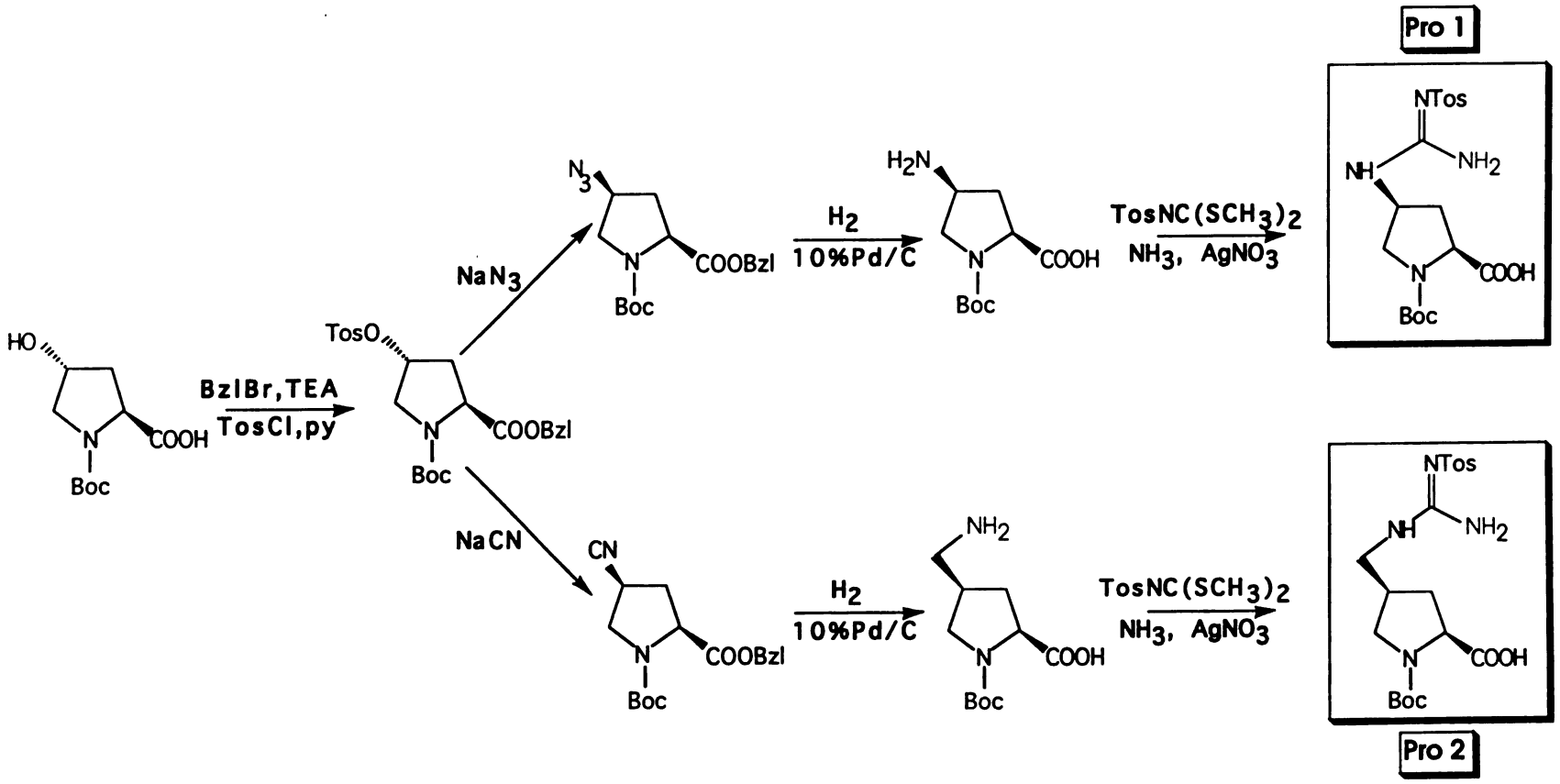
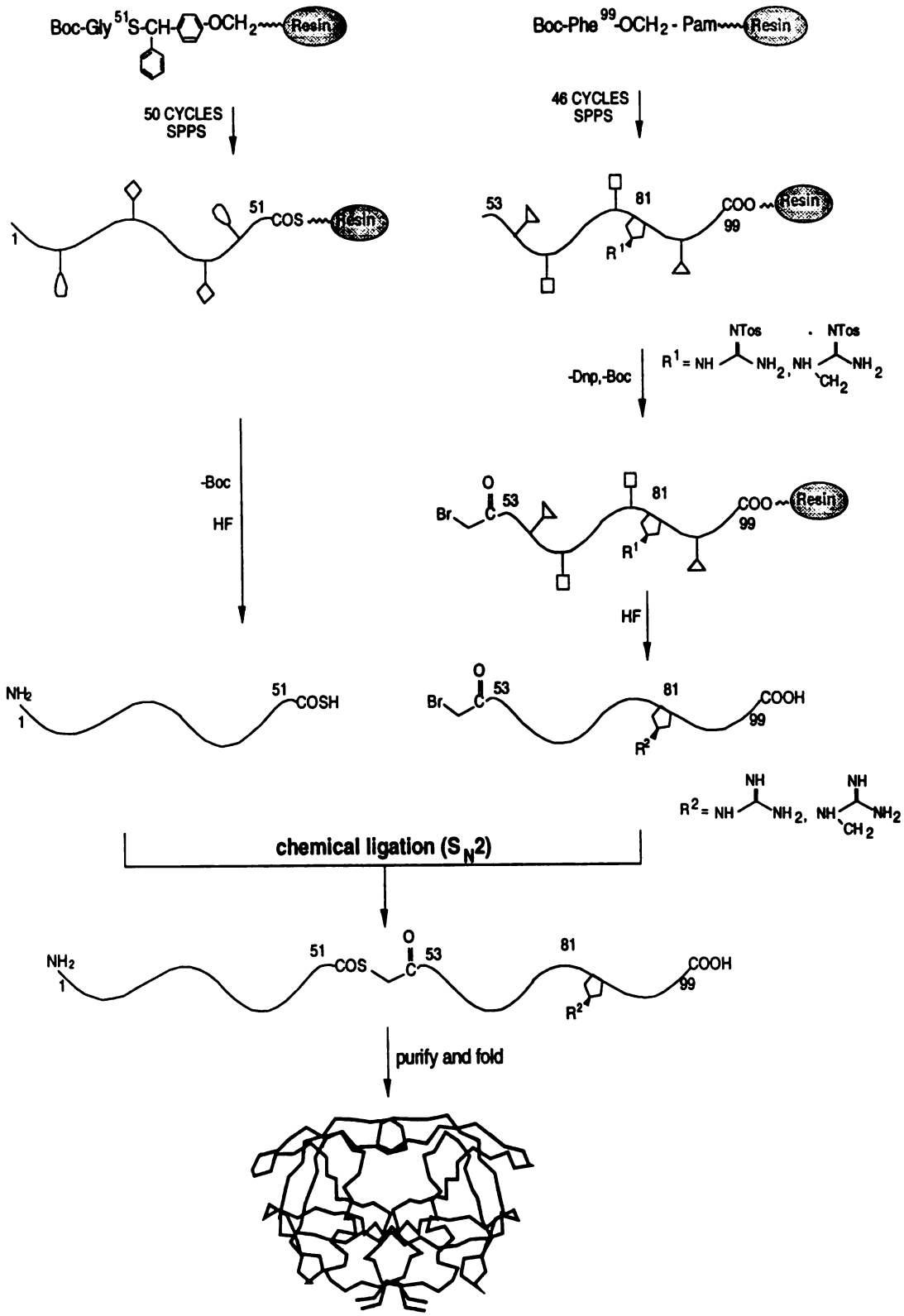


Figure 4.2 Synthesis of geometrically-constrained proline analogues.

Figure 4.3 Schematic illustration of the chemical ligation strategy used for the preparation of HIV-1 PR analogues.


Boc-Gly 51S-CH-OCH2 Resin
50 CYCLES SPPS
Boc-Phe 99-OCH2 - Pam Resin
46 CYCLES SPPS
-Dnp,-Boc
-Boc
HF
HF
chemical ligation (SN2)
purify and fold

incorporated at position 81 into the BrAc(53-99) HIV-1 PR segment during the course of manual stepwise solid-phase assembly of the protected polypeptide chain[17]. A native sequence control ("wild type") molecule ($[NHCH_2COSCH_2CO]^{51\text{-}52}$-$Aba^{67,95}$) HIV-1 PR was also synthesized in order to act as a reference point for the kinetic and specificity properties of the enzyme analogues.

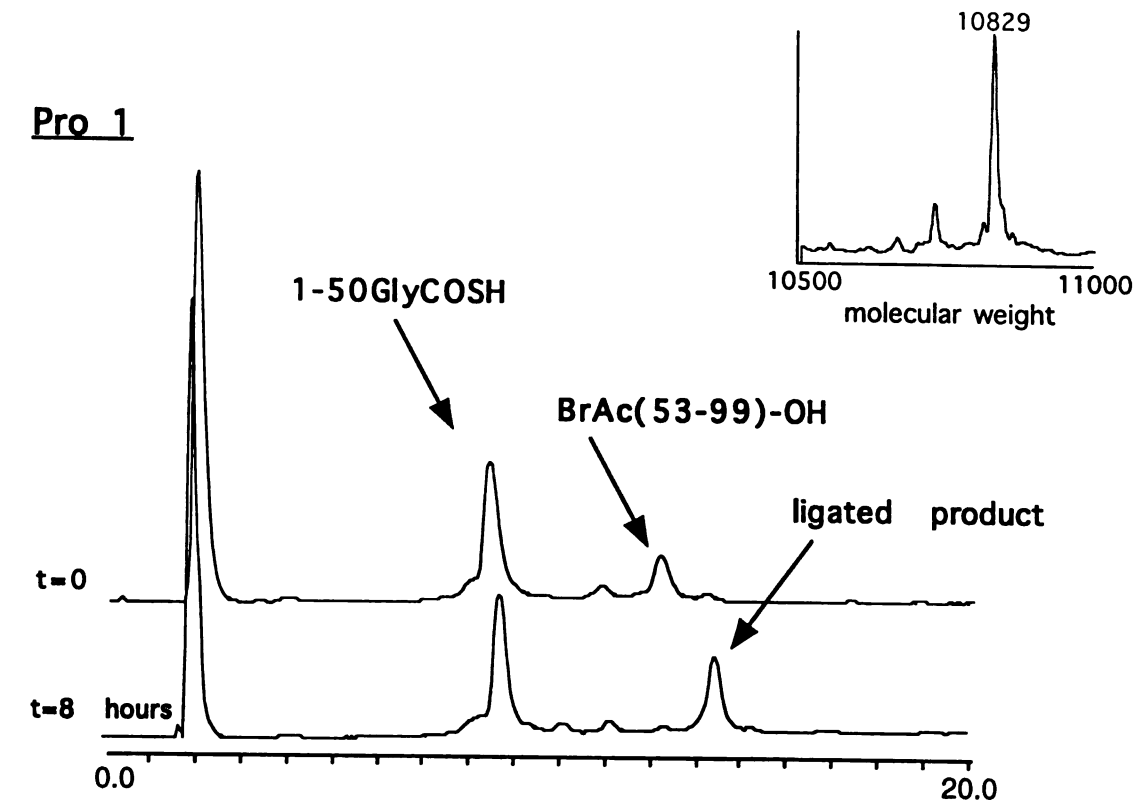
Ligation of HIV-1 PR segments (1-51)αCOSH and BrAc(53-99) was performed under aqueous conditions in 8M urea, 0.1 M sodium phosphate, pH= 5.5. The progress of each ligation reaction was monitored by analytical reverse-phase HPLC and electrospray MS (Figure 4.4). As indicated in Figure 4.4, the ligation reactions went in high yield, cleanly, and showed little or no side product formation. Because of the essentially quantitative yield of the ligation reaction, the crude ligation mixtures were then subjected to folding and subsequently characterized for specificity and enzymatic activity without further purification.

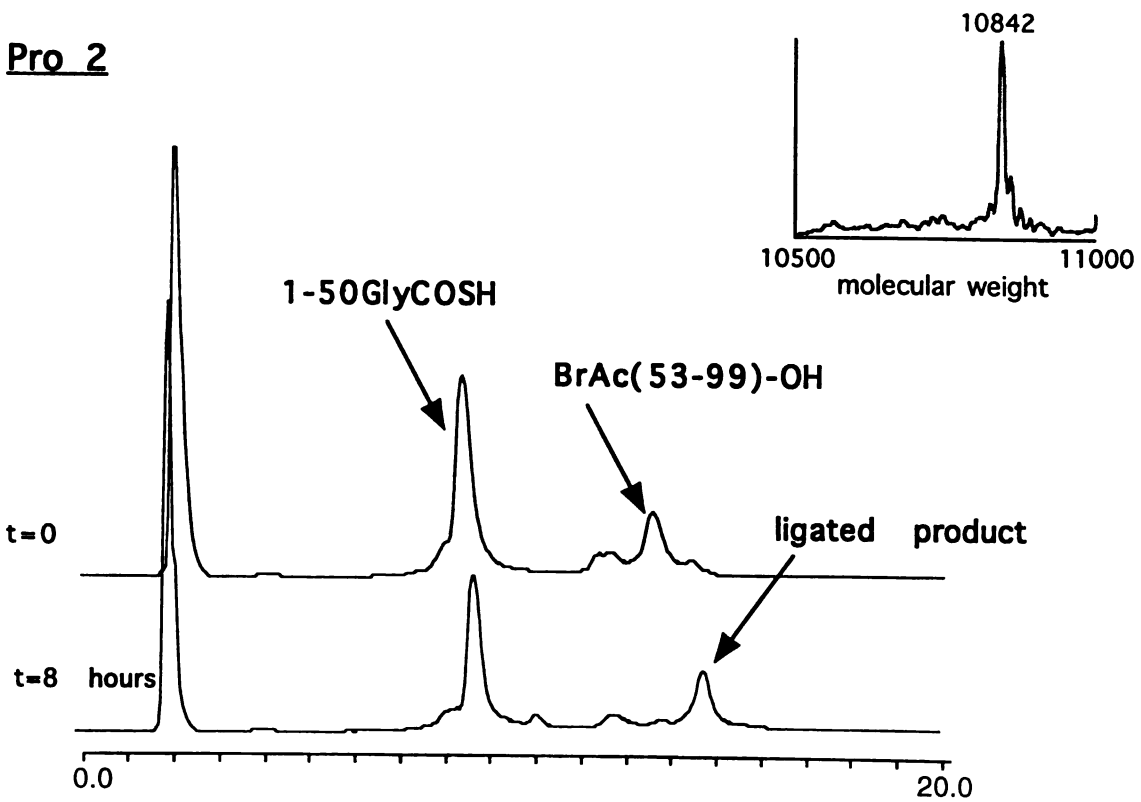
4.2.4 Enzymatic activity of HIV-1 PR analogues

Specificity

All three molecular forms of HIV-1 PR cleaved a series of synthetic peptide analogues of gag and pol polypeptides at the natural processing sites only, indicating unchanged specificity. Molecular modelling was performed on the $(Pro1)^{81/181}$ and $(Pro2)^{81/181}$ HIV-1 PR analogues based on the structure of the HIV-1 PR complexed with MVT-101 (4hvp)[9] (Figures 4.5 and 4.6). As indicated in Figures 4.5 and 4.6, the guanidino and

Figure 4.4 Synthesis of (Pro1)^{81} and (Pro2)^{81} HIV-1 PR monomers by thioester-forming chemical ligation of HIV-1 PR segments (1-50)GlyCOSH and BrAc(53-99)-OH monitored by analytical HPLC and ESMS.


Pro 1
10829
10500
11000
molecular weight
1-50GlyCOSH
BrAc(53-99)-OH
ligated product
t=0
t=8 hours
0.0
20.0


Pro 2
10842
10500
11000
molecular weight
1-50GlyCOSH
BrAc(53-99)-OH
ligated product
t=0
t=8 hours
0.0
20.0

Figures 4.5 and 4.6 Molecular modelling of the substrate binding pocket of the HIV-1 PR incorporating the Pro1 and Pro2 amino acid analogues at positions 81 and 181 of the HIV-1 PR dimer. The binding pocket was modelled based on the crystal structure of the HIV-1 PR complexed to the inhibitor MVT-101[9] using the computer program Insight II (Biosym Technologies, Inc., San Diego, CA.). The inhibitor MVT-101 (N-acetyl-Thr-Ile Nle-Ψ[CH_2-NH]-Nle-Gln-Arg-amide) is highlighted in pink. Pro1 and Pro 2 are shown in white. The guanidinium and guanidinium methyl substituents in the 4-*cis* position of Pro1 and Pro2, respectively, point directly into the S1 and S1' positions of the substrate binding pocket.

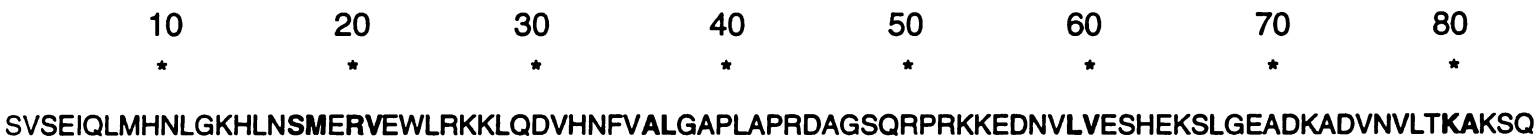
guanidium Pro 81
P1'
P1
guanidium Pro 181

guanidiniummethyl Pro 81
P1'
P1
guanidiniummethyl Pro 181

guanidinomethyl substituents point directly into the binding pocket. In the absence of substrate, the binding pocket of the HIV-1 PR is freely accessible to water. The guanidino functionality is strongly basic (pKa ≅ 13) and will pick up a proton from bulk solvent even at relatively high pH values. Based on the presence of positively charged guanidino groups in the binding pocket, it seemed likely that the HIV-1 PR analogues would recognize substrates which contained acidic side chains (Asp or Glu) at the P1 and P1' positions. We speculated that these acidic side chains would form salt bridges with the guanidino and guanidinomethyl moieties of $(Pro1)^{81/181}$ and $(Pro2)^{81/181}$ introduced into the S1 and S1' binding pockets.

A series of peptides based on the sequence of the p1/NC peptide substrate which contained Asp and/or Glu in the P1 and P1' positions of the substrate were synthesized and tested for cleavage by the HIV-1 PR analogues. None of these substrates were cleaved by either HIV-1 PR or the $(Pro1 \text{ or } 2)^{81,181}$ HIV-1 PR analogues. In addition, we searched for any new cleavage sites by incubating the polypeptide substrates human parathyroid hormone (hPTH) (1-84) and one of the ligation segments HIV-1 PR (1-51)αCOSH. HIV-1 PR (1-51)αCOSH segment was known to contain two autolysis sites, L^{5}-W^{6} and L^{33}-E^{34}[18-21]. The two HIV-1 PR analogues cleaved hPTH and the HIV-1 PR (1-51)αCOSH segment with the same specificity as the "wild-type" HIV-1 PR (Figure 4.7). The unchanged substrate specificities of the $(Pro1)^{81,181}$ and $(Pro2)^{81,181}$ HIV-1 PR analogues were unexpected based on the results of our molecular modelling analysis and on the work of Leis and coworkers[22,23], who have suggested that the individual subsites of the substrate binding pocket in the HIV-1 PR act independently in the selection of substrates.

(a).



(b).

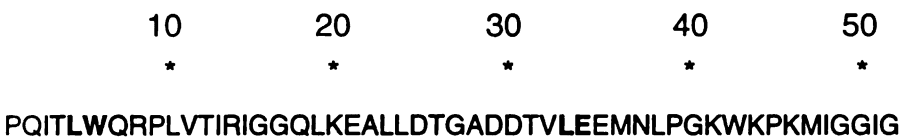


Figure 4.7 Amino acid sequences of human parathyroid hormone and HIV-1 PR fragment (1-50)αCOSH. Cleavage sites are highlighted in bold.

Kinetics

The enzymatic activity of "wild type" HIV-1 PR and the two analogues were determined using an HPLC assay based on the cleavage of a synthetic peptide substrate [RNTATIM*MQRGNFR-amide] derived from the second capsid/nucleocapsid (p1/NC) junction of the viral gag polyprotein[24]. Kinetic parameters for the cleavage of the p1/NC synthetic substrate are summarized Table 4.1. Based on the kinetic data, $(Pro1)^{81,181}$ and $(Pro2)^{81,181}$ HIV-1 PR analogues were shown to be approximately 100- and 10- fold less active than "wt" PR on p1/NC substrate based on comparison of the catalytic efficiency (kcat/Km) of the enzymes. The reduction in the catalytic efficiency of the enzyme analogues is due to both the elevation of the Km (15-fold for $(Pro1)^{81,181}$ HIV-1 PR and 3.5 fold for $(Pro2)^{81,181}$ HIV-1 PR) in addition to the reduction in the kcat values (by approximately 7- and 3-fold, respectively).

4.2.5 Significance

In this study, we have made use of total chemical synthesis to prepare S1/S1'-modified analogues of the HIV-1 PR in order to test the molecular origins of substrate recognition. Two HIV-1 PR analogues were constructed which were designed to introduce polar, charged substituents in the S1 and S1' substrate binding pockets with the notion that the substrate specificites of these enzymes would be altered in a predictable fashion. Interestingly, this was not the case, as the HIV-1 PR analogues still maintained the same substrate specificity as wild type, as shown by the cleavage of a range of substrates at identical sites. Even though the HIV-1 PR analogues had the

Table 4.1. Kinetic Constants for HIV-1 PR Mutants

Proteases	Km (μM)	kcat (sec^{-1})	kcat/Km ($\mu M^{-1}min^{-1}$)
wild-type†	190	3.00	0.95
$(Pro1)^{81,181}$ (guanidino)	2870	0.45	0.01
$(Pro2)^{81,181}$ (guanidinomethyl)	670	1.22	0.11

Kinetic parameters were determined for the cleavage of the synthetic peptide substrate RNTATIM*MQRGNFR-amide, where * indicates the site of cleavage. The substrate concentration varied between 20-2000μM in the assays. wt PR and variants were assayed in a buffer containing 100mM sodium acetate, 0.5 M NaCl, 10% glycerol and 0.5 mg/ml BSA at pH=5.3, 37°C. †wt = $([NHCH_2COSCH_2CO]^{51\text{-}52}\text{-}Aba^{67,95})$HIV-1 PR.

same specificity, the substrates were recognized less well as indicated by the observed Km values (Table 4.1).

The (Pro2)^{81,181} HIV-1 PR containing the guanidinium methyl substituent in the S1 and S1' binding pocket showed only a 10-fold reduction in catalytic efficiency (kcat/Km), as compared to the 100-fold reduction observed with the (Pro1)^{81,181} HIV-1 PR which contains the guanidinium moiety (Table 4.1). These results may be explained by the extra degree of freedom introduced by the incorporation of the methylene group in the modified proline analogue Pro2, which would allow the guanidinium group to move away and reorient itself in order to accommodate substrate more productively. The guanidinium substituent in the 4-*cis* position of Pro1 is highly constrained, since it is linked directly to the proline ring, and would not have the ability to reorient itself to interact more productively with the substrate once incorporated in the HIV-1 PR. The reorientation of a side chain in the S1/S1' binding pocket was also observed in the case of a drug-resistant mutant, V82I[25]. The structure of the V82I mutant with DMP-323 shows a loss of interaction between the cγ1 methyl group of the valine side chain and the inhibitor, even though isoleucine has a larger side chain. As indicated by the structure, the energetically favorable side-chain rotomer for isoleucine resulted in a repositioning of the Cδ1 methyl group away from the inhibitor and resulted in reduced interaction between the inhibitor and the enzyme.

It seems that the notion of altering the substrate specificity by introducing substituents in the substrate binding pocket was not unfounded. A study appeared recently that exploited a similar design as the present

study[26]. Moody and coworkers observed that the addition of a side chain at postion 48 of the HIV-1 PR can alter the substrate specificity in a sequence-specific manner and that compensatory changes can be made in the substrate to restore activity. In this study, the G48E mutation in the flap of the HIV-1 PR produced an enzyme with reduced activity (approximately 7-fold compared to wt) on one of the processing sites in the viral pol polyprotein precursor. This mutation results in the placement of a side chain in the S3 and S3' subsites of the enzyme. Cleavage at this site was restored by making a change in the P3' position of the substrate, by replacing an aspartic acid residue with either glycine or asparagine. Additionally, cleavage at another processing site of the viral polyprotein was enhanced, and the enhancement was dependent on the presence of an arginine residue in the P3' position of the substrate. Thus, the altered specificity of the G48E mutant examined in this study was shown to result from charge interaction between the glutamic acid side chain and the residue present in the P3' position of the substrate.

Additionally, it is possible to think of a substrate (or an inhibitor) of the HIV-1 PR as a missing component of the enzyme molecule[25]. In this case, when a substrate or inhibitor is bound to the enzyme, one might consider any changes made in the HIV-1 PR binding pocket analogous to making a substitution in the interior of a protein. Certainly, when a substrate or inhibitor is bound to the HIV-1 PR, its binding pocket does begin to possess many of the aspects of a solvent-inaccessible protein core[25]. Then, similar to the observations of Matthews and coworkers in their studies with T4 lysozyme[27,28], it is possible that shifts in the backbone of the modified HIV-1 PR analogues and slight alterations of side chain configuration in the substrate could result in the repacking of elements in the binding pocket (or

protein core) and allow the analogue enzyme molecules to bind and cleave the natural substrates. The predicted result would then be, as observed in this study, unaltered substrate specificity. However there is a cost, as reflected in the reduced levels of binding (increased Km values) and catalytic activity (diminished kcat/Km). Again, a similar phenomenon has been observed in the context of the structure of a drug-resistant mutant as described by Baldwin et al.[29]. The structure of the V82A mutant of the HIV-1 PR showed that the mutation gave rise to backbone shifts that resulted in repacking of the enzyme and inhibitor atoms in the S1/S1' and S3/S3' subsites, which was unexpected. These results support the notion that substrate or inhibitor binding to the HIV-1 PR resemble the packing interactions that occur inside the core of a protein and, additionally, that the HIV-1 PR is highly flexible and can accommodate a number of subsite mutations and still be functional. Additionally, it would appear that the recognition of substrates is highly dependent on the cumulative interactions between the enzyme and substrate[14], as opposed to independent recognition of substrate within the individual subsites proposed by Leis and coworkers[22,23].

Extensions of the present work could involve investigating the effects of changes introduced at either S1 or S1' on the specificity of HIV-1 PR. One might be able to alter specificity more effectively by incorporating novel moieties at one site in the enzyme binding pocket (i.e., S1 only). The HIV-1 PR molecule is a C2-symmetric homodimer in its unliganded form. However, the HIV-1 PR loses its symmetry upon inhibitor/substrate binding[9,30,31] and thus the binding pockets are nonequivalent[22]. Therefore, it might make more sense to introduce an asymmetric change in the PR and

to is dimeric nature, a single mutation at the monomer level results in changes at two sites in the functional HIV-1 PR dimer molecule. The only way to make a single, asymmetric change in the HIV-1 PR is to make a tethered version of the enzyme. This has been accomplished by recombinant means[32-35] and more recently by total chemical synthesis[36]. One would predict that if a large, hydrophobic substituent, a benzyl group for example, were substituted in the 4-cis position of Pro81, cleavage may be observed at Gly*Pro (a site not normally cleaved) instead of Tyr/Phe*Pro. The incorporation of a benzyl group would be analogous to exchanging or incorporating an element of the substrate into the HIV-1 PR. The tethered dimer construct of the HIV-1 PR would also make it possible to probe individual subsites of the HIV-1 PR to see what effect changes made in the individual subsites would have on specificity.

It appears that more light could be shed on the results obtained in this study if crystal structures of these analogues were available. As has been shown with their drug-resistant relatives[25,29], the crystal structures of the (Pro1)^{81,181} and (Pro2)^{81,181} HIV-1 PR analogues generated in this study would aid in better understanding the structural basis of our kinetic observations.

4.3 Experimental Section

4.3.1 Materials and methods

Boc-trans-4-hydroxy-L-proline was purchased from Bachem (Bubendorf, Switzerland). All other Boc-amino acids and 2-(1-H-

benzotriazol-1-yl)-1,1,3,3,-tetramethyluronium hexafluorophosphate (HBTU) were obtained from Nova Biochem (San Diego, CA.). Boc-Phe-OCH_2Pam-resin, aminomethylcopoly(styrene-divinylbenzene) resin and diisopropylethylamine were purchased from Applied Biosysyems, Inc. (Foster City, CA.). N,N-Dimethylformamide (Synthesis-grade) was obtained from Auspep (Parkville, Australia). AR-grade methylene chloride was obtained from Fisher. TFA was purchased from Halocarbon (River Edge, N. J.) and HF obtained from Matheson Gas. HPLC-grade acetonitrile was obtained from either EM Science or Fisher. Vydac HPLC columns were purchased form Western Analytical Co. (Temecula, CA.). Flash chromatography was perfomed using Merck silica gel 60 (230-400 mesh) obtained from EM Science. TLC was performed using Whatman silica gel 60A plates and were developed by heating followed by spraying with ninhydrin. All other reagents were AR grade or better, and purchased from either Aldrich Chemical or Fisher.

Machine-assisted solid-phase peptide syntheses were performed on a custom-modified Applied Biosystems 430A peptide/protein synthesizer[17]. This instrument was modified by installation of bypass valves to accelerate flows and to prevent clogging of valve blocks. Preparative HPLC was performed on a Waters Delta-prep 4000 HPLC system. Semipreparative HPLC was performed on a Rainin HPXL dual pump system with detection on a Dynamax UV detector. Analytical HPLC was performed on either a Rainin HPXL dual pump system with detection on a Dynamax UV detector or on an integrated Hewlett-Packard 1050 system. Preparative HPLC was run on a Vydac C18 column (15-20μ, 5.0 x 25 cm) at a flow rate of 30mL/min, semipreparative HPLC was run on a Vydac C4 column (10μ, 1.0

x 25 cm) at a flow rate of 3mL/min, and analytical HPLC was run on a Vydac C4 column (5μ, 0.46 x 15 cm) at 1mL/min. Chromatographic separations were achieved using linear gradients of 0.1% TFA in water (solvent A) versus 90% acetonitrile/ 10% water, 0.09% TFA (solvent B). Mass analyses of all peptide segments and proteins were performed on a Sciex API-III triple quadrupole electrospray mass spectrometer as previously described[37]. Masses were derived from the experimental m/z values for all observed protonation states of a molecular species using the program MacSpec (Sciex). Calculated masses were based on average isotope composition and were derived using the program MacProMass (Sunil Vemuri and Terry Lee, Beckman Research Institute, Duarte, CA.). Non-protein mass spectrometry was performed at The Scripps Research Institute Mass Spectrometry Facility. ^{1}H NMR spectra were recorded on a Bruker 250MHz spectrometer.

4.3.2 Synthesis of Pro1 and Pro2

The fully protected 4-*cis*-(X)-proline derivatives (Pro1 and Pro2) were synthesized from *trans*-4-hydroxy-L-proline according to the method of Webb and Eigenbrot[15] with the following modifications: *trans*-4-hydroxy-L-proline was transformed into the benzyl ester by treatment with triethylamine and benzyl bromide in refluxing ethyl acetate. In addition, p-toluenesulfonyl chloride was used instead of methanesulfonyl chloride. The high resolution MS, NMR data and optical purities obtained for all compounds were in agreement with reported literature values.

4.3.3 Solid-phase peptide synthesis

All peptides were synthesized in a stepwise fashion according to the *in situ* neutralization/HBTU activation protocols for machine-assisted Boc solid-phase chemistry[17]. Side chain protection was as follows: Arg(*p*-toluenesulfonyl), Asn(xanthyl), Asp(*O*-cyclohexyl), Glu(*O*-cyclohexyl), His(Dnp), Lys(2-Cl-Z), Ser(benzyl), Trp(formyl), Thr(benzyl), and Tyr(2-Br-Z). Gln and Met (and when synthesizing thioacid peptides, Trp) were used without side chain protection. Coupling yields were determined by the quantitative ninhydrin assay on peptide-resin samples taken under machine control during the chain assembly[17,38].

The HIV-1 PR (SF2 strain) carboxyl-terminal fragment (53-99) was prepared manually on a Boc-Phe-4-(carboxamidomethyl)benzyl ester-copoly(styrene-1%-divinylbenzene) resin (Boc-Phe-OCH_2-Pam-resin). This segment contained L-α-amino-n-butyric acid as an isosteric replacement for cysteine at positions 67 and 95[39]. The fully protected Pro1 and Pro2 amino acid analogues were coupled at position 81 by normal coupling procedures. After coupling of the proline analogues, any remaining free amine was acetylated by reaction with 10% acetic anhydride/10% pyridine in DMF for 15 minutes. Capping by acetylation was performed because the coupling yields for Pro1 and Pro2 at position 81 were between 85-90% due to the limited amounts of each compound available. The bromoacetyl group was introduced at the amino terminus of the assembled peptide chain by manually coupling of the preformed symmetric anhydride of bromoacetic acid after removal of the His(Dnp) and Boc protecting groups from the resin-bound peptide as previously described[16,40]. The N-terminal fragment

corresponding to HIV-1 PR (1-51) was constructed on a 4-[α-(Boc-Gly-S)benzyl]-phenoxyacetamidomethyl resin, which on treatment with anhydrous HF yields a peptide with a C-terminal Gly-αCOSH$^{41-43}$. Side chain-unprotected Trp was used in the synthesis of this peptide because the thiolysis conditions used to remove the formyl protecting group caused a side reaction with the -αCOSH group. Upon completion of the chain assembly, peptides were deprotected as necessary and cleaved from the resin by treatment with liquid HF, containing 10% *p*-cresol, for 1 hour at 0 °C. After evaporation of HF under reduced pressure, crude peptide products were precipitated in ice-cold anhydrous diethyl ether, dissolved in 50% aqueous acetic acid, and either purified immediately or after prior lyophilization.

4.3.4 Purification and characterization of peptide segments

Crude lyophilized peptides were dissolved in either acidic aqueous buffers or 50% aqueous acetic acid and purified by preparative HPLC. Purified peptides were stored as lyophilized powders at -20°C. The purified peptide segments were characterized by electrospray MS, and all possessed observed masses within experimental error of the calculated masses.

4.3.5 Synthesis of "wt" control, (Pro1)^{81,181} and (Pro2)^{81,181} HIV-1 PR analogues by chemical ligation

Ligation of HIV-1 PR segments (1-51)αCOSH and BrAc(53-99) was performed under the aqueous conditions using 8M urea, 0.1 M sodium phosphate, pH= 5.5. For each ligation reaction, the following quantities of

the HIV-1 PR segments (1-51)αCOSH and BrAc(53-99) were used, respectively: wt--2.54 mgs (0.45 μmol) and 1.98 mgs (0.38 μmol) in 300 μl ligation buffer; Pro1--1.53 mgs (0.27 μmol) and 1.17 mgs (0.22 μmol) in 200 μl ligation buffer; Pro2--2.86 mgs (0.51 μmol) and 1.8 mgs (0.34 μmol) in 240 μl ligation buffer. The progress of each reaction was monitored by reversed-phase analytical HPLC and electrospray MS. The ligation products charaterized by electrospray MS possessed masses of 10,770 ± 2 Da for "wt" control HIV-1 PR [calculated 10,771 Da (average isotope composition)], 10,829 ± 3 Da for Pro1 HIV-1 PR [calculated 10,829 Da (average isotope composition)] and 10,842 ± 2 Da for Pro2 HIV-1 PR [calculated 10,843 Da (average isotope composition)].

4.3.6 Enzyme folding

The crude reaction mixtures of wt, $(Pro1)^{81,181}$ and $(Pro2)^{81,181}$ HIV-1 PR analogues were folded by dialysis over a two day period against decreasing concentrations of guanidine hydrochloride (6M-0M) into dialysis buffer (20mM sodium acetate, 20% glycerol at pH=5.0, containing appropriate amounts of the competitive inhibitor, pepstatin A). Following dialysis, the sample was centrifuged to remove any precipitate and stored at 4°C. Concentrations of folded protein were determined by measuring the peak area obtained in an analytical HPLC run and comparing to the area obtained from a sample of denatured enzyme of known concentration. (The concentration of the denatured enzyme was determined by absorbance at 280 nm, using a molar extinction coefficient of 25,500 M^{-1} cm^{-1}[144]).

4.3.7 Enzyme assays

Kinetic parameters were determined for the cleavage of the peptide substrate RNTATIM*MQRGNFR in the following assay buffer: pH=5.3 buffer containing 100mM sodium acetate, 0.5M NaCl, 10% glycerol and bovine serum albumin (BSA) at 0.5 mg/ml. DMSO was used to solubilize the synthetic substate. Because DMSO has been found to weakly inhibit the enzyme[45], its final assay concentration was kept standardized at 2%. Cleavage products were quantitated by HPLC with peak detection by UV absorbance (214 nm). Initial rates were determined by single time points of reactions in which consumption of substrate was <20% of the starting concentration (60-90 minute reaction times). Kinetic parameters were obtained by fitting data to the Michaelis-Menten equation using a nonlinear regression computer program[46].

References

(1). Debouck, C. *AIDS Res. Hum. Retroviruses* **1992**, *8*, 153-164.

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

(3). Meek, T. D.; Dayton, B. D.; Metcalf, B. W.; Dreyer, G. B.; Strickler, J. E.; Gorniak, J. G.; Rosenberg, M.; Moore, M. L.; Magaard, V. W.; Debouck, C. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 1841-1845.

(4). Pearl, L. H.; Taylor, W. R. *Nature* **1987**, *329*, 351-354.

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

(6). Miller, M.; Jaskólski, M.; Rao, J. K.; Leis, J.; Wlodawer, A. *Nature* **1989**, *337*, 576-579.

(7). Tomasselli, A. G.; Howe, W. J.; Sawyer, T. K.; Wlodawer, A.; Heinrikson, R. L. *Chimica Oggi* **1991**, *9*, 6-27.

(8). Wlodawer, A.; Erickson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543-585.

(9). Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. *Science* **1989**, *246*, 1149-1152.

(10). Fitzgerald, P. M.; McKeever, B. M.; VanMiddlesworth, J. F.; Springer, J. P.; Heimbach, J. C.; Leu, C. T.; Herber, W. K.; Dixon, R. A.; Darke, P. L. *J. Biol. Chem.* **1990**, *265*, 14209-14219.

(11). Gustchina, A.; Sansom, C.; Prevost, M.; Richelle, J.; Wodak, S. Y.; Wlodawer, A.; Weber, I. T. *Protein Eng.* **1994**, *7*, 309-17.

(12). Oroszlan, S. In *Viral Proteinases as Targets For Chemotherapy*; H. Kräusslich; S. Oroszlan and E. Wimmer, Ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989; pp 87-100.

(13). Tomasselli, A.; Heinrikson, R. L. *Methods in Enzymology* **1994**, *241*, 279-301.

(14). Kent, S. B. H.; Schneider, J.; Clawson, L.; Selk, L.; Delahunty, C.; Chen, Q. In *Viral Proteinases as Targets For Chemotherapy*; H. Kräusslich; S. Oroszlan and E. Wimmer, Ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989; pp 223-230.

(15). Webb, T. R.; Eigenbrot, C. *J. Org. Chem.* **1991**, *56*, 3009-3016.

(16). Schnölzer, M.; Kent, S. B. H. *Science* **1992**, *256*, 221-225.

(17). Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180-193.

(18). Mildner, A. M.; Rothrock, D. J.; Leone, J. W.; Bannow, C. A.; Lull, J. M.; Reardon, I. M.; Sarcich, J. L.; Howe, W. J.; Tomich, C. C.; Smith, C. W.; Heinrikson, R. L.; Tomasselli, A. G. *Biochemistry* **1994**, *33*, 9405-9413.

(19). Rosé, J. R.; Salto, R.; Craik, C. S. *J. Biol. Chem.* **1993**, *268*, 11939-11945.

(20). Strickler, J. E.; Gorniak, J.; Dayton, B.; Meek, T.; Moore, M.; Magaard, V.; Malinowski, J.; Debouck, C. *Proteins* **1989**, *6*, 139-154.

(21). Hui, J. O.; Tomasselli, A. G.; Reardon, I. M.; Lull, J. M.; Brunner, D. P.; Tomich, C. S.; Heinrikson, R. L. *J. Protein Chem.* **1993**, *12*, 323-7.

(22). Cameron, C. E.; Grinde, B.; Jacques, P.; Jentoft, J.; Leis, J.; Wlodawer, A.; Weber, I. T. *J. Biol. Chem.* **1993**, *268*, 11711-20.

(23). Cameron, C. E.; Ridky, T. W.; Shulenin, S.; Leis, J.; Weber, I. T.; Copeland, T.; Wlodawer, A.; Burstein, H.; Bizub-Bender, D.; Skalka, A. M. *J. Biol. Chem.* **1994**, *269*, 11170-7.

(24). Sanchez-Pescador, R.; Power, M. D.; Barr, P. J.; Steimer, K. S.; Stempien, M. M.; Brown-Shimer, S. L.; Gee, W. W.; Renard, A.; Randolph, A.; Levy, J. A.; Dina, D.; Luciw, P. A. *Science* **1985**, *227*, 484-492.

(25). Erickson, J. W. *Nature Structural Biology* **1995**, *2*, 523-529.

(26). Moody, M. D.; Pettit, S. C.; Shao, W.; Everitt, L.; Loeb, D. D.; Hutchison, I., C. A.; Swanstrom, R. *Virology* **1995**, *207*, 475-485.

(27). Baldwin, E. P.; Hajiseyedjavadi, O.; Baase, W. A.; Matthews, B. W. *Science* **1993**, *262*, 1715-1718.

(28). Baldwin, E. P.; Matthews, B. W. *Current Opinion in Biotechnology* **1994**, *5*, 396-402.

(29). Baldwin, E. T.; Bhat, T. N.; Liu, B.; Pattabiraman, N.; Erickson, J. W. *Nature Structural Biology* **1995**, *2*, 244-249.

(30). Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Wideburg, N. *Science* **1990**, *249*, 527-533.

(31). Bone, R.; Vacca, J. P.; Anderson, P. S.; Holloway, K. M. *J. Am. Chem. Soc.* **1991**, *113*, 9382-9384.

(32). Cheng, Y. S.; Yin, F. H.; Foundling, S.; Blomstrom, D.; Kettner, C. A. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9660-9664.

(33). Kräusslich, H. G. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 3213-3217.

(34). Griffiths, J. T.; Tomchak, L. A.; Mills, J. S.; Graves, M. C.; Cook, N. D.; Dunn, B. M.; Kay, J. *J. Biol. Chem.* **1994**, *269*, 4787-4793.

(35). Dilanni, C. L.; Davis, L. J.; Holloway, M. K.; Herber, W. K.; Darke, P. L.; Kohl, N. E.; Dixon, R. A. *J. Biol. Chem.* **1990**, *265*, 17348-17354.

(36). Baca, M.; Muir, T. W.; Schnölzer, M.; Kent, S. B. H. *J. Am. Chem. Soc.* **1994**, *117*, 1881-1887.

(37). Schnölzer, M.; Jones, A.; Alewood, P. F.; Kent, S. B. H. *Anal. Biochem.* **1992**, *204*, 335-343.

(38). Sarin, V. K.; Kent, S. B. H.; Tam, J. P.; Merrifield, R. B. *Analytical Biochemistry* **1981**, *117*, 147-157.

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

(40). Robey, F. A.; Fields, R. L. *Anal. Biochem.* **1989**, *177*, 373-377.

(41). Yamashiro, D.; Li, C. H. *Int. J. Pept. Protein Res.* **1988**, *31*, 322-334.

(42). Canne, L. E.; Walker, S. M.; Kent, S. B. H. *Tetrahedron Letters* **1995**, *36*, 1217-1220.

(43). Blake, J. *Int. J. Pept. Protein Res.* **1981**, *17*, 273-274.

(44). Gill, S. C.; von Hippel, P. H. *Anal. Biochem.* **1989**, *182*, 319-326.

(45). Bergman, D. A.; Alewood, D.; Alewood, P. F.; Andrews, J. L.; Brinkworth, R. I.; Englebretsen, D. R.; Kent, S. B. H. *Letters in Peptide Science* **1995**, in press.

(46). Duggleby, R. G. *Comput. Biol. Med.* **1984**, *14*, 447-455.

Chapter V. Summary and Future Directions

Over the past few years, a number of new methods have emerged that make the total chemical synthesis of proteins and their analogues a viable alternative to recombinant DNA-based protein biosynthesis. The chemical synthesis techniques described in this body of work and elsewhere have made it possible to construct proteins containing novel structures, which allow the exploration of protein structure and function in an unprecedented manner. By integrating organic chemistry with synthetic peptide chemistry, these techniques have opened up a whole new world of protein structure-function analysis not previously explored. The total chemical synthesis of HIV-1 PR analogues described in this work illustrates the versatility inherent in a chemical synthesis approach and gives some sense of the scope and limitations of the method.

In Chapter II, two different methods for the total chemical synthesis of the HIV-1 PR containing native backbone structure, stepwise SPPS and the native chemical ligation of unprotected peptide segments, were compared. Both methods produced high purity HIV-1 PR in good yield. Synthetic HIV-1 PR possessed activities characteristic of the native enzyme active in viral maturation. The K41C mutation present in the native ligated HIV-1 PR did not have a negative effect on the activity. In addition, this form of the HIV-1 PR could be produced consistently, in high purity and yield. This method simplifies the synthesis of HIV-1 PR analogues in allowing for their construction via a modular or "cassette" approach, where modified peptide segments can be combined so as to produce the desired mutant.

Chapter III describes the modular construction of drug-resistant forms of the HIV-1 PR. These drug-resistant forms were tested for inhibiton by

three different types of inhibitors. The work in this chapter illustrates the utility of a chemical synthesis approach to the study of resistant forms of the HIV-1 PR, in addition to providing a good system for the production of multiple variants for biochemical characterization. In the future, rapid chemical synthesis could be used to explore the properties of resistant variants produced from exposure to different types of inhibitors in order to gain a better understanding of the effects the different mutations have on HIV-1 PR activity. In addition, this work could be extended by subsequent crystal structural analysis of the variants constructed in order to gain an understanding of the structural basis of resistance (see Baldwin et al.[1]).

The work described in Chapter IV demonstrated the use of molecular modelling and chemistry in engineering rational changes in the HIV-1 PR structure. Unnatural amino acids were incorporated into the S1 and S1' positions of the substrate binding pocket of the HIV-1 PR in order to probe the molecular basis of substrate recognition. The aim was to design changes in the HIV-1 PR with predictable consequences. Contrary to our expectations, the engineered enzyme retained a remarkably robust ability to cleave at the native processing sites. This study, however, demonstrated that it is not all that simple to design "rational" changes in a molecule that will produce predictable effects. Ultimately, the investigation of single-site mutations in a chemically synthesized tethered dimer form of the HIV-1 PR should be explored to see what effect an asymmetric change in the binding pocket will have on the specificity of the HIV-1 PR.

Table 5.1 Summary of HIV-1 protease analogues produced by total chemical synthesis.
1. autolysis-resistant HIV-1 PR*
2. (Cys^{41}) autolysis-resistant HIV-1 PR*
3. (Cys^{41}/M^{90}) autolysis-resistant HIV-1 PR*
4. ($Cys^{41}/V^{48}/M^{90}$) autolysis-resistant HIV-1 PR*
5. $(COS)^{51\text{-}52}$ HIV-1 PR ("backbone engineered")
6. $(COS)^{51\text{-}52}$ $(Pro1)^{81,181}$ HIV-1 PR
7. $(COS)^{51\text{-}52}$ $(Pro2)^{81,181}$ HIV-1 PR
All enzyme analogues contain l-α-amino-n-butyric acid at positions 67 and 95 as isosteric replacements for native cysteine residues. *Autolysis-resistant mutations were incorporated where indicated at positions 7 (Q7K) and 33 (L33I) of the HIV-1 PR.

A summary of all the HIV-1 PR analogues examined in this body of work is provided in Table 5.1. The 99-amino acid monomer of HIV-1 PR was synthesized by stepwise SPPS and using two different chemoselective ligation methods. These HIV-1 PR analogues can be divided into two groups, those which could have been produced by conventional means (i.e., recombinant DNA technology) [analogues 1-4], and those only accessible by total chemical synthesis [analogues 5-7]. The various forms of HIV-1 PR characterized in this body of work demonstrate that the HIV-1 PR is a very robust enzyme that can tolerate a number of changes in its primary structure. It is quite remarkable that so many changes could be introduced into this molecule, either in the backbone and/or at the amino acid level, and still lead to high levels of enzymatic activity. Due to its essential nature in the viral lifecycle, it may be that the HIV-1 PR has evolved to maintain a certain level of activity even in the presence of a diversity of point mutations. This seems

quite likely, especially in light of the fact that the viral reverse transcriptase enzyme has a high rate of mutation and can generate any number of mutations in a variety of locations the HIV genome, including the protease coding region, during viral replication[2,3].

The observations reported in this body of work may offer important insights into the structural origins of drug-resistance in the HIV-1 PR. These drug-resistant forms of the HIV-1 PR are produced in response to exposure to protease inhibitors. A number of drug-resistant variants of the HIV-1 PR have been described which contain mutations in the substrate binding pocket (Refer to the analysis of the drug-resistant variants described in Chapter III and see Erickson[4], Mellors et al.[5], Gulnik[6] and references contained within.). These variants tend to display both varying degrees of reduced catalytic activity and less efficient binding of substrates, as demonstrated by reduced kcat/Km values. Therefore, the $(Pro1)^{81,181}$ and $(Pro2)^{81,181}$ HIV-1 PR analogues could be considered analogous to the "active-site" escape mutants of the HIV-1 PR produced by exposure to different protease inhibitors in that these variants can still recognize and cleave substrates, but just less efficiently due to the effect of the mutation. Given the enormous changes introduced into the molecular structure of the HIV-1 PR by the incorporation of Pro1 and Pro2 with minimal impact on specificity and catalysis, it is likely that the problem of inhibitor-resistant mutant forms of HIV-1 PR will be an intractable one. Crystal structures of these active site HIV-1 PR analogues would likely aid in our understanding the structural basis for our kinetic observations, as it has for some resistance mutants[4,7].

It is my hope that the work described in this thesis will contribute to a better understanding of the molecular basis of specificity and catalytic activity of HIV-1 PR and, in addition, demonstrate the applicability and versatility inherent in a total chemical synthesis approach to the study of proteins.

References

(1). Baldwin, E. T.; Bhat, T. N.; Liu, B.; Pattabiraman, N.; Erickson, J. W. *Nature Structural Biology* **1995**, *2*, 244-249.

(2). Preston, B. D.; Poiesz, B. J.; Loeb, L. A. *Science* **1988**, *242*, 1168-1171.

(3). Roberts, J. D.; Bebenek, K.; Kunkel, T. A. *Science* **1988**, *242*, 1171-1173.

(4). Erickson, J. W. *Nature Structural Biology* **1995**, *2*, 523-529.

(5). Mellors, J. W.; Larder, B. A.; Schinazi, R. F. *International Antiviral News* **1995**, *3*, 8-13.

(6). Gulnik, S. V.; Suvorov, L. I.; Liu, B.; Yu, B.; Anderson, B.; Mitsuya, H.; Erickson, J. W. *Biochemistry* **1995**, *34*, 9282-9287.

(7). Baldwin, E. T.; Bhat, T. N.; Gulnik, S.; Liu, B.; Topol, I. A.; Kiso, Y.; Mimoto, T.; Mitsuya, H.; Erickson, J. W. *Structure* **1995**, *3*, 581-590.

Bibliography

(1994). Centers for Disease Control and Prevention. HIV/AIDS Surveillance Report **6**(2): 5.

(1995). In Point of Fact. World Health Organization Press Office.

Ashorn, P., T. J. McQuade, et al. (1990). "An inhibitor of the protease blocks maturation of human and simian immunodeficiency viruses and spread of infection." Proc. Natl. Acad. Sci. USA **87**: 7472-6.

Baca, M., P. F. Alewood, et al. (1993). "Structural engineering of the HIV-1 protease molecule with a beta-turn mimic of fixed geometry." Protein Science **2**: 1085-1091.

Baca, M. and S. B. H. Kent (1993). "Catalytic contribution of flap-substrate hydrogen bonds in "HIV-1 protease" explored by chemical synthesis." Proc. Natl. Acad. Sci. USA **90**: 11638-11642.

Baca, M., T. W. Muir, et al. (1994). "Chemical ligation of cysteine-containing peptides: synthesis of a 22 kDa tethered dimer HIV-1 protease analogue." J. Am. Chem. Soc. **117**(7): 1881-1887.

Baldwin, E. P., O. Hajiseyedjavadi, et al. (1993). "The role of backbone flexibility in the accommodation of variants that repack the core of T4 lysozyme." Science **262**: 1715-1718.

Baldwin, E. P. and B. W. Matthews (1994). "Core-packing constraints, hydrophobicity and protein design." Curr. Opin. Biotech. **5**: 396-402.

Baldwin, E. T., T. N. Bhat, et al. (1995). "Structure of HIV-1 protease with KNI-272, a tight-binding transition-state analog containing allophenylnorstatine." Structure **3**(6): 581-590.

Baldwin, E. T., T. N. Bhat, et al. (1995). "Structural basis of drug resistance for the V82A mutant of HIV-1 proteinase." Nature Structural Biology **2**(3): 244-249.

Barre-Sinoussi, F., J. C. Chermann, et al. (1983). "Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)." Science **220**: 868-871.

Bergman, D. A., D. Alewood, et al. (1995). "Kinetic properties of HIV-1 protease produced by total chemical synthesis with cysteine residues replaced by isosteric L-a-amino-n-butyric acid." Letters in Peptide Science : in press.

Billich, S., M. Knoop, et al. (1988). "Synthetic peptides as substrates and inhibitors of human immunodeficiency virus-1 protease." J. Biol. Chem. **263**: 17905-17908.

Blake, J. (1981). "Peptide segment coupling in aqueous medium: silver ion activation of the thiolcarboxyl group." Int. J. Pept. Protein Res. **17**: 273-274.

Bone, R., J. P. Vacca, et al. (1991). "X-ray crystal structure of the HIV protease complex with L-700,417, an inhibitor with pseudo C_2 symmetry." J. Am. Chem. Soc. **113**: 9382-9384.

Cameron, C. E., B. Grinde, et al. (1993). "Comparison of the substrate-binding pockets of the Rous sarcoma virus and human immunodeficiency virus type 1 proteases." J. Biol. Chem. **268**: 11711-20.

Cameron, C. E., T. W. Ridky, et al. (1994). "Mutational analysis of the substrate binding pockets of the Rous sarcoma virus and human immunodeficiency virus-1 proteases." J. Biol. Chem. **269**: 11170-7.

Canne, L. E., A. R. Ferré-D'Amaré, et al. (1995). "Total chemical synthesis of a unique transcription factor-related protein: cMyc-Max." J. Am. Chem. Soc. **117**: 2998-3007.

Canne, L. E., S. M. Walker, et al. (1995). "A general method for the synthesis of thioester resin linkers for use in the solid phase synthesis of peptide-α-thioacids." Tetrahedron Letters **36**(8): 1217-1220.

Cheng, Y. S., F. H. Yin, et al. (1990). "Stability and activity of human immunodeficiency virus protease: comparison of the natural dimer with a homologous, single-chain tethered dimer." Proc. Natl. Acad. Sci. USA **87**: 9660-9664.

Clavel, F., D. Guétard, et al. (1986). "Isolation of a new human retrovirus from West African patients with AIDS." Science **233**: 343-346.

Coffin, J., A. Haase, et al. (1986). "Human immunodeficiency viruses." Science **232**: 697.

Collins, J. R., S. K. Burt, et al. (1995). "Flap opening in HIV-1 protease simulated by 'activated' molecular dynamics." Nature Structural Biology **2**(4): 334-338.

Condra, J. H., W. A. Schleif, et al. (1995). "In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors." Nature **374**: 569-571.

Cooper, J., S. Foundling, et al. (1987). "The structure of a synthetic pepsin inhibitor complexed with endothiapepsin." Eur. J. Biochem. **169**: 215-221.

Cornish, V. W., D. Mendel, et al. (1995). "Probing protein structure and function with an expanded genetic code." Angew. Chem. Int. Ed. Engl. **34**: 621-633.

Cornish, V. W. and P. G. Schultz (1994). "A new tool for studying protein structure and function." Curr. Opin. Struct. Biol. **4**: 601-607.

Craig, J. C., I. B. Duncan, et al. (1991). "Antiviral properties of Ro 31-8959, an inhibitor of human immunodeficiency virus (HIV) proteinase." Antiviral Res. **16**: 295-305.

Daniel, M. D., N. L. Letvin, et al. (1985). "Isolation of T-cell tropic HTLV-III-like retrovirus from macaques." Science **228**: 1201-1204.

Darby, S. C., D. W. Ewart, et al. (1995). "Mortality before and after HIV infection in the complete UK population of haemophiliacs." Nature **377**: 79-82.

Darke, P. L. and J. R. Huff (1994). "HIV protease as an inhibitor target for the treatment of AIDS." Adv. Pharmacol. **25**: 399-454.

Darke, P. L., R. F. Nutt, et al. (1988). "HIV-1 protease specificity of peptide cleavage is sufficient for processing of gag and pol polyproteins." Biochem. Biophys. Res. Commun. **156**: 297-303.

Dawson, P. E., T. W. Muir, et al. (1994). "Synthesis of proteins by native chemical ligation." Science **266**: 776-779.

Debouck, C. (1992). "The HIV-1 protease as a therapeutic target for AIDS." AIDS Res. Hum. Retroviruses **8**: 153-164.

DesJarlais, R. L., G. L. Seibel, et al. (1990). "Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease." Proc. Natl. Acad. Sci. USA **87**: 6644-6648.

Dilanni, C. L., L. J. Davis, et al. (1990). "Characterization of an active single polypeptide form of the human immunodeficiency virus type 1 protease." J. Biol. Chem. **265**: 17348-17354.

Dixon, M. (1953). "The determination of enzyme inhibitor constants." Biochem. J. **55**: 170-171.

Duggleby, R. G. (1984). "Regression analysis of nonlinear Arrhenius plots: an empirical model and a computer program." Comput. Biol. Med. **14**: 447-455.

Eberle, J., B. Bechowsky, et al. (1995). "Resistance of HIV-1 to proteinase inhibitor Ro 31-8959." AIDS Research and Human Retroviruses **11**(6): 671-676.

el-Farrash, M. A., M. J. Kuroda, et al. (1994). "Generation and characterization of a human immunodeficiency virus type 1 (HIV-1) mutant resistant to an HIV-1 protease inhibitor." J. Virol. **68**: 233-9.

Ellman, J. A., D. Mendel, et al. (1992). "Site-specific incorporation of novel backbone structures into proteins." Science **255**: 197-200.

Erickson, J., D. J. Neidhart, et al. (1990). "Design, activity, and 2.8 Å crystal structure of a C2 symmetric inhibitor complexed to HIV-1 protease." Science **249**: 527-533.

Erickson, J. W. (1995). "The not-so-great escape." Nature Structural Biology **2**(7): 523-529.

Fitzgerald, P. M., B. M. McKeever, et al. (1990). "Crystallographic analysis of a complex between human immunodeficiency virus type 1 protease and acetyl-pepstatin at 2.0-Å resolution." J. Biol. Chem. **265**: 14209-14219.

Gaertner, H. F., K. Rose, et al. (1992). "Construction of protein analogues by site-specific condensation of unprotected peptides." Bioconjugate Chem. **3**: 262-268.

Gallo, R. C. and L. Montagnier (1988). "AIDS in 1988." Sci Am **259**: 41-48.

Gallo, R. C., S. Z. Salahuddin, et al. (1984). "Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS." Science **224**: 500-503.

Gallo, R. C., P. S. Sarin, et al. (1983). "Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS)." Science **220**: 865-867.

Gill, S. C. and P. H. von Hippel (1989). "Calculation of protein extinction coefficients from amino acid sequence data." Anal. Biochem. **182**: 319-326.

Greenlee, W. J. (1990). "Renin inhibitors." Med. Res. Rev. **10**: 173-236.

Griffiths, J. T., L. H. Phylip, et al. (1992). "Different requirements for productive interaction between the active site of HIV-1 proteinase and substrates containing -hydrophobic*hydrophobic- or -aromatic*pro-cleavage sites." Biochemistry **31**: 5193-5200.

Griffiths, J. T., L. A. Tomchak, et al. (1994). "Interactions of substrates and inhibitors with a family of tethered HIV-1 and HIV-2 homo- and heterodimeric proteinases." J. Biol. Chem. **269**: 4787-4793.

Gulnik, S. V., L. I. Suvorov, et al. (1995). "Kinetic characterization and cross resistance patterns of HIV-1 protease mutants selected under drug pressure." Biochemistry **34**(29): 9282-9287.

Gustchina, A., C. Sansom, et al. (1994). "Energy calculations and analysis of HIV-1 protease-inhibitor crystal structures." Protein Engineering **7**: 309-17.

Guyader, M., M. Emerman, et al. (1987). "Genome organization and transactivation of the human immunodeficency virus type 2." Nature **326**: 662-669.

Henderson, L. E., T. D. Copeland, et al. (1988). Analysis of proteins and peptides from sucrose gradient banded HTLV-III. Human Retroviruses, Cancer, and AIDS: Approaches to Prevention and Therapy. New York, Alan R. Liss, Inc. 135-147.

Henderson, P. J. F. (1972). "A linear equation that describes the steady-state kinetics of enzymes and subcellular particles interacting with tightly bound inhibitors." Biochem. J. **127**: 321-333.

Hilvert, D. (1994). "Chemical synthesis of proteins." Chemistry and Biology **1**: 201-203.

Ho, D. D., T. Toyoshima, et al. (1994). "Characterization of human immunodeficiency virus type 1 variants with increased resistance to a C2-symmetric protease inhibitor." J. Virol. **68**: 2016-20.

Huff, J. R. (1991). "HIV protease: a novel chemotherapeutic target for AIDS." J. Med. Chem. **34**: 2305-2314.

Hui, J. O., A. G. Tomasselli, et al. (1993). "Large scale purification and refolding of HIV-1 protease from Escherichia coli inclusion bodies." J. Protein Chem. **12**: 323-7.

Hui, J. O., A. G. Tomasselli, et al. (1990). "Ribonuclease A as a substrate of the protease from human immunodeficiency virus-1." J. Biol. Chem. **265**: 21386-9.

Hyland, L. J., T. T.A., et al. (1991). "Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism." Biochemistry **30**: 8454-8463.

Hyland, L. J., T. T.A., et al. (1991). "Human immunodeficiency virus-1 protease. 1. Initial velocity studies and kinetic characterization of reaction intermediates by ^{18}O isotope exchange." Biochemistry **30**: 8441-8453.

Jacks, T., M. D. Power, et al. (1988). "Characterization of ribosomal frameshifting in HIV-1 gag-pol expression." Nature **331**: 280-283.

Jacks, T., K. Townsley, et al. (1987). "Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins." Proc. Natl. Acad. Sci. USA **84**(12): 4298-4302.

Jacks, T. and H. E. Varmus (1985). "Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting." Science **230**: 1237-1242.

Jacobsen, H., K. Yasargil, et al. (1995). "Characterization of human immunodeficiency virus type 1 mutants with decreased sensitivity to proteinase inhibitor Ro 31-8959." Virology **206**: 527-534.

Kanki, P. J., M. F. McLane, et al. (1985). "Serologic identification and characterization of a macaque T-lymphotropic retrovirus closely related to HTLV-III." Science **228**: 1199-1201.

Kaplan, A. H., S. F. Michael, et al. (1994). "Selection of multiple human immunodeficiency virus type 1 variants that encode viral proteases with decreased sensitivity to an inhibitor of the viral protease." Proc. Natl. Acad. Sci. USA **91**: 5597-601.

Katoh, I., T. Yasunaga, et al. (1987). "Inhibition of retroviral protease activity by an aspartyl proteinase inhibitor." Nature **329**: 654-656.

Kempf, D. J., K. C. Marsh, et al. (1995). "ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans." Proc. Natl. Acad. Sci. USA **92**: 2484-2488.

Kent, S. B. H. (1988). "Chemical synthesis of peptides and proteins." Ann. Rev. Biochem. **57**: 957-989.

Kent, S. B. H., J. Schneider, et al. (1989). Chemical synthesis approach to the molecular biology of retroviral proteases: structure-function studies of the HIV-1 and HIV-2 enzymes and their substrates. Viral Proteinases as Targets For Chemotherapy. Cold Spring Harbor, Cold Spring Harbor Laboratory Press. 223-230.

Kohl, N. E., E. A. Emini, et al. (1988). "Active human immunodeficiency virus protease is required for viral infectivity." Proc. Natl. Acad. Sci. USA **85**: 4686-4690.

Kramer, R. A., M. D. Schaber, et al. (1986). "HTLV-III gag protein is processed in yeast cells by the virus pol-protease." Science **231**: 1580-1584.

Kräusslich, H. G. (1991). "Human immunodeficiency virus proteinase dimer as component of the viral polyprotein prevents particle assembly and viral infectivity." Proc. Natl. Acad. Sci. USA **88**: 3213-3217.

Kuntz, I. D., J. M. Blaney, et al. (1982). "A geometric approach to macromolecule-ligand interactions." J. Mol. Biol. **161**: 269-288.

Lam, P. Y., P. K. Jadhav, et al. (1994). "Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors." Science **263**: 380-384.

Larder, B. A. (1993). Inhibitors of HIV reverse transcriptase as antiviral agents and drug resistance. Reverse Transcriptases. New York, Cold Spring Harbor Laboratory Press. 205-222.

Larder, B. A., G. Darby, et al. (1989). "HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy." Science **243**: 1731-1734.

Letvin, N. L., M. D. Daniel, et al. (1985). "Induction of AIDS-like disease in macaque monkeys with T-cell tropic retrovirus STLV-III." Science **230**: 71-73.

Lillehoj, E. P., F. H. Salazar, et al. (1988). "Purification and structural characterization of the putative gag-pol protease of human immunodeficiency virus." J. Virol. **62**: 3053-3058.

Lin, Y., X. Lin, et al. (1995). "Effect of point mutations on the kinetics and the inhibition of human immunodeficiency virus type 1 protease: relationship to drug resistance." Biochemistry **34**(4): 1143-1152.

Loll, P. J., D. Picot, et al. (1995). "The structural basis of aspirin activity inferred from the crystal structure of inactivated prostaglandin H_2 synthase." Nature Structural Biology **2**(8): 637-643.

Martin, J. A. (1992). "Recent advances in the design of HIV proteinase inhibitors." Antiviral Res. **17**: 265-78.

McQuade, T. J., A. G. Tomasselli, et al. (1990). "A synthetic HIV-1 protease inhibitor with antiviral activity arrests HIV-like particle maturation." Science **247**: 454-456.

Meek, T. D. (1992). "Inhibitors of HIV-1 protease." J. Enzym. Inhib. **6**: 65-98.

Meek, T. D., B. D. Dayton, et al. (1989). "Human immunodeficiency virus 1 protease expressed in Escherichia coli behaves as a dimeric aspartic protease." Proc. Natl. Acad. Sci. USA **86**: 1841-1845.

Meek, T. D. and G. B. Dreyer (1990). "HIV-1 protease as a potential target for anti-AIDS therapy." Ann. N. Y. Acad. Sci. **616**: 41-53.

Meek, T. D., E. J. Rodriguez, et al. (1994). "Use of steady state kinetic methods to elucidate the kinetic and chemical mechanisms of retroviral proteases." Methods in Enzymology **241**: 127-156.

Mellors, J. W., B. A. Larder, et al. (1995). "Mutations in HIV-1 reverse transcriptase and protease associated with drug resistance." International Antiviral News **3**(1): 8-13.

Mildner, A. M., D. J. Rothrock, et al. (1994). "The HIV-protease as enzyme and substrate: mutagenesis of autolysis sites and generation of a stable mutant with retained kinetic properties." Biochemistry **33**: 9405-9413.

Miller, M. (1993). D-HIV-1 protease structure.

Miller, M., M. Jaskólski, et al. (1989). "Crystal structure of a retroviral protease proves relationship to aspartic protease family." Nature **337**: 576-579.

Miller, M., J. Schneider, et al. (1989). "Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 A resolution." Science **246**: 1149-1152.

Milton, R. C., S. C. Milton, et al. (1992). "Total chemical synthesis of a D-enzyme: the enantiomers of HIV-1 protease show reciprocal chiral substrate specificity." Science **256**: 1445-8.

Milton, R. C., S. C. F. Milton, et al. (1993). Synthesis of proteins by chemical ligation of unprotected peptide segments: mirror-image enzyme molecules, D- & L-HIV protease analogs. Techniques in Protein Chemistry IV. San Diego, Academic Press. 257-267.

Mitsuya, H. (1992). "Development of inhibitors of reverse transcriptase and protease as therapeutics against HIV infection." J. Enzym. Inhib. **6**: 1-8.

Mitsuya, H. and S. Broder (1987). "Strategies for antiviral therapy in AIDS." Nature **325**: 773-778.

Mitsuya, H., R. Yarchoan, et al. (1990). "Molecular targets for AIDS therapy." Science **249**: 1533-1544.

Moody, M. D., S. C. Pettit, et al. (1995). "A side chain at position 48 of the human immunodeficiency virus type-1 protease flap provides an additional specificity determinant." Virology **207**: 475-485.

Mous, J., E. P. Heimer, et al. (1988). "Processing protease and reverse transcriptase from human immunodeficiency virus type I polyprotein in Escherichia coli." J. Virol. **62**: 1433-6.

Muesing, M. A., D. H. Smith, et al. (1985). "Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus." Nature **313**: 450-458.

Muir, T. M. (1995). "A chemical approach to the construction of multimeric protein assemblies." Structure **3**: 649-652.

Muir, T. W. and S. B. H. Kent (1993). "The chemical synthesis of proteins." Curr. Opin. Biotech. **4**: 420-427.

Muir, T. W., M. J. Williams, et al. (1994). "Design and chemical synthesis of a neoprotein structural model for the cytoplasmic domain of a multisubunit cell-surface receptor: integrin $\alpha_{IIb}\beta_3$ (platelet GPIIb-IIIa)." Biochemistry **33**: 7701-7708.

Munroe, D. G. and C. Y. Lau (1995). "Turning down the heat: new routes to inhibition of inflammatory signaling by prostaglandin H_2 synthases." Chemistry and Biology **2**: 343-350.

Navia, M. A., P. M. Fitzgerald, et al. (1989). "Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1." Nature **337**: 615-620.

North, T. W. and R. A. LaCasse (1995). "Testing anti-HIV drugs in the FIV model." Nature Medicine **1**(5): 410-411.

Nutt, R. F., S. F. Brady, et al. (1988). "Chemical synthesis and enzymatic activity of a 99-residue peptide with a sequence proposed for the human immunodeficiency virus protease." Proc. Natl. Acad. Sci USA **85**: 7129-7133.

Oroszlan, S. (1989). Biosynthesis and proteolytic processing of retroviral proteins: an overview. Viral Proteinases as Targets For Chemotherapy. Cold Spring Harbor, Cold Spring Harbor Laboratory Press. 87-100.

Oswald, M. and K. von der Helm (1991). "Fibronectin is a non-viral substrate for the HIV proteinase." FEBS Lett. **292**: 298-300.

Otto, M. J., S. Garber, et al. (1993). "In vitro isolation and identification of human immunodeficiency virus (HIV) variants with reduced sensitivity to C-2 symmetrical inhibitors of HIV type 1 protease." Proc. Natl. Acad. Sci USA **90**: 7543-7547.

Pearl, L. H. and W. R. Taylor (1987). "A structural model for the retroviral proteases." Nature **329**: 351-354.

Pettit, S. C., J. Simsic, et al. (1991). "Analysis of retroviral protease cleavage sites reveals two types of cleavage sites and the structural requirements of the P1 amino acid." J. Biol. Chem. **266**: 14539-14547.

Preston, B. D., B. J. Poiesz, et al. (1988). "Fidelity of HIV-1 reverse transcriptase." Science **242**: 1168-1171.

Ratner, L., W. Haseltine, et al. (1985). "Complete nucleotide sequence of the AIDS virus, HTLV-III." Nature **313**: 277-284.

Richards, A. D., L. H. Phylip, et al. (1990). "Sensitive, soluble chromogenic substrates for HIV-1 proteinase." J. Biol. Chem. **265**: 7733-7736.

Richards, A. D., R. Roberts, et al. (1989). "Effective blocking of HIV-1 proteinase activity by characteristic inhibitors of aspartic proteinases." FEBS Lett. **247**: 113-7.

Richman, D. D. (1995). "Protease uninhibited." Nature **374**: 494.

Riviere, Y., V. Blank, et al. (1991). "Processing of the precursor of NF-kappa B by the HIV-1 protease during acute infection." Nature **350**: 625-626.

Roberts, J. D., K. Bebenek, et al. (1988). "The accuracy of reverse transcriptase from HIV-1." Science **242**: 1171-1173.

Roberts, M. M. and S. Oroszlan (1989). "The preparation and biochemical characterization of intact capsids of equine infectious anemia virus." Biochem. and Biophys. Res. Commun. **160**: 486-494.

Roberts, N. A., J. C. Craig, et al. (1992). "HIV proteinase inhibitors." Biochem. Soc. Trans. **20**: 513-6.

Roberts, N. A., J. A. Martin, et al. (1990). "Rational design of peptide-based HIV proteinase inhibitors." Science **248**: 358-361.

Robey, F. A. and R. L. Fields (1989). "Automated synthesis of N-bromoacetyl-modified peptides for the preparation of synthetic peptide polymers, peptide-protein conjugates, and cyclic peptides." Anal. Biochem. **177**: 373-377.

Rodriguez, E. J., T. S. Angeles, et al. (1993). "Use of nitrogen-15 kinetic isotope effects to elucidate details of the chemical mechanism of human immunodeficiency virus 1 protease." Biochemistry **32**: 12380-12385.

Rosé, J. R. and C. S. Craik (1994). "Structure-assisted design of nonpeptide human immunodeficiency virus-1 protease inhibitors." Am. J. Respir. Crit Care Med. **150**: S176-S182.

Rosé, J. R., R. Salto, et al. (1993). "Regulation of autoproteolysis of the HIV-1 and HIV-2 proteases with engineered amino acid substitutions." J. Biol. Chem. **268**: 11939-11945.

Rose, K. (1994). "Facile synthesis of homogeneous artificial proteins." J. Am. Chem. Soc. **116**(1): 30-33.

Sanchez-Pescador, R., M. D. Power, et al. (1985). "Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2)." Science **227**: 484-492.

Sardana, V. V., A. J. Schlabach, et al. (1994). "Human immunodeficiency virus type 1 protease inhibitors: evaluation of resistance engendered by amino acid substitutions in the enzyme's substrate binding site." Biochemistry **33**: 2004-2010.

Sarin, V. K., S. B. H. Kent, et al. (1981). "Quantitative monitoring of solid-phase peptide synthesis by the ninhydrin reaction." Analytical Biochemistry **117**: 147-157.

Schecter, I. and A. Berger (1967). "On the size of the active site in proteases. I. Papain." Biochem. Biophys. Res. Commun. **27**: 157-162.

Schneider, J. and S. B. H. Kent (1988). "Enzymatic activity of a synthetic 99 residue protein corresponding to the putative HIV-1 protease." Cell **54**: 363-368.

Schnölzer, M., P. Alewood, et al. (1992). "In-situ neutralization in Boc-chemistry solid phase peptide synthesis: rapid, high yield assembly of difficult sequences." Int. J. Pept. Protein Res. **40**: 180-193.

Schnölzer, M., A. Jones, et al. (1992). "Ion-spray tandem mass spectrometry in peptide synthesis: structural characterization of minor by-products in the synthesis of ACP(65-74)." Anal. Biochem. **204**: 335-343.

Schnölzer, M. and S. B. H. Kent (1992). "Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease." Science **256**: 221-225.

Seelmeier, S., H. Schmidt, et al. (1988). "Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A." Proc. Natl. Acad. Sci. USA **85**: 6612-6.

Shoeman, R. L., C. Kesselmier, et al. (1991). "Non-viral cellular substrates for human immunodeficiency virus type 1 protease." FEBS Lett. **278**: 199-203.

Slee, D. H., K. L. Laslo, et al. (1995). "Selectivity in the inhibition of HIV and FIV protease: inhibitory and mechanistic studies of pyrrolidine-containing α-keto amide and hydroxyethylamine core structures." JACS : in press.

Steele, F. (1995). "AIDS drugs lurch towards market." Nature Medicine **1**(4): 285-286.

Strickler, J. E., J. Gorniak, et al. (1989). "Characterization and autoprocessing of precursor and mature forms of human immunodeficiency virus type 1 (HIV 1) protease purified from Escherichia coli." Proteins **6**: 139-154.

Tisdale, M., R. Myers, et al. (1994). Abstracts of the Third International Workshop on HIV Drug Resistance : p. 14.

Toh, H., M. Ono, et al. (1985). "Retroviral protease-like sequence in the yeast transposon Ty1." Nature **315**: 691.

Tomasselli, A. and R. L. Heinrikson (1994). "Specificity of retroviral proteases: an analysis of viral and nonviral protein substrates." Methods in Enzymology **241**(15): 279-301.

Tomasselli, A. G., W. J. Howe, et al. (1991). "Calcium-free calmodulin is a substrate of proteases from human immunodeficiency viruses 1 and 2." Proteins **10**: 1-9.

Tomasselli, A. G., W. J. Howe, et al. (1991). "The complexities of AIDS: an assessment of the HIV protease as a therapeutic target." Chimica Oggi **9**: 6-27.

Tomasselli, A. G., J. O. Hui, et al. (1991). "Actin, troponin C, alzheimer amyloid precursor protein and pro-interleukin 1 beta as substrates of the protease from human immunodeficiency virus." J. Biol. Chem. **266**: 14548-53.

Tomasselli, A. G., J. O. Hui, et al. (1990). "Proteases from human immunodeficiency virus and avian myeloblastosis virus show distinct specificities in hydrolysis of multidomain protein substrates." J. Virol. **64**: 3157-61.

Tomasselli, A. G., J. O. Hui, et al. (1990). "Interdomain hydrolysis of a truncated Pseudomonas exotoxin by the human immunodeficiency virus-1 protease." J. Biol. Chem. **265**: 408-13.

Tomasselli, A. G., J. O. Hui, et al. (1991). "The evaluation of non-viral substrates of the HIV protease as leads in the design of inhibitors for AIDS therapy." Adv. Exp. Med. Biol. **306**: 469-82.

Trono, D. (1995). "HIV accessory proteins: leading roles for the supporting cast." Cell **82**: 189-192.

Vacca, J. P., B. D. Dorsey, et al. (1994). "L-735,524: an orally bioavailable human immunodeficiency virus type 1 protease inhibitor." Proc. Natl. Acad. Sci. USA **91**: 4096-4100.

Vella, S. (1994). "Update on a Proteinase Inhibitor." AIDS **8** (suppl 3): S25-S29.

von der Helm, K., L. Gurtler, et al. (1989). "Inhibition of HIV replication in cell culture by the specific aspartic protease inhibitor pepstatin A." FEBS Lett. **247**: 349-52.

Wain-Hobson, S., P. Sonigo, et al. (1985). "Nucleotide sequence of the AIDS virus, LAV." Cell **40**: 9-17.

Wallace, C. J. A. (1995). "Peptide ligation and semisynthesis." Current Opinion in Biotechnology **6**: 403-410.

Webb, T. R. and C. Eigenbrot (1991). "Conformationally restricted arginine analogues." J. Org. Chem. **56**: 3009-3016.

Williams, M. J., T. W. Muir, et al. (1994). "Ligated 10F3, the integrin-binding module of fibronectin." J. Am. Chem. Soc. **116**: 10797-10798.

Wlodawer, A. (1994). "Rational drug design: the proteinase inhibitors." Pharmacotherapy **14**(6 pt 2): 9S-20S.

Wlodawer, A. and J. W. Erickson (1993). "Structure-based inhibitors of HIV-1 protease." Ann. Rev. Biochem. **62**: 543-585.

Wlodawer, A., M. Miller, et al. (1989). "Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease." Science **245**: 616-621.

Yamashiro, D. and C. H. Li (1988). "New segment synthesis of alpha-inhibin-92 by the acyl disulfide method." Int. J. Pept. Protein Res. **31**: 322-334.

Yarchoan, R., H. Mitsuya, et al. (1988). "AIDS therapies." Science **259**(4): 110-119.

Appendix A

A General Method for the Synthesis of Thioester Resin Linkers for Use in the Solid Phase Synthesis of Peptide-α-Thioacids

Published 1995, Tetrahedron Letters **36**, 1217-1220.

Pergamon

Tetrahedron Letters, Vol. 36, No. 8, pp. 1217-1220, 1995
Elsevier Science Ltd
Printed in Great Britain
0040-4039/95 $9.50+0.00

0040-4039(95)00037-2

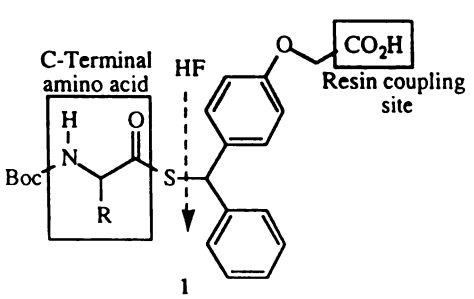
A General Method for the Synthesis of Thioester Resin Linkers for Use in the Solid Phase Synthesis of Peptide-α-Thioacids

Lynne E. Canne,* Sharon M. Walker, and Stephen B.H. Kent
The Scripps Research Institute, 10666 North Torrey Pines Rd,
La Jolla, CA 92037

Abstract: A generalized procedure for the synthesis of thioester linkers for use in stepwise solid phase peptide synthesis is reported. The linkers are compatible with Boc chemistry and, upon cleavage in HF, generate peptide C-terminal thioacids.

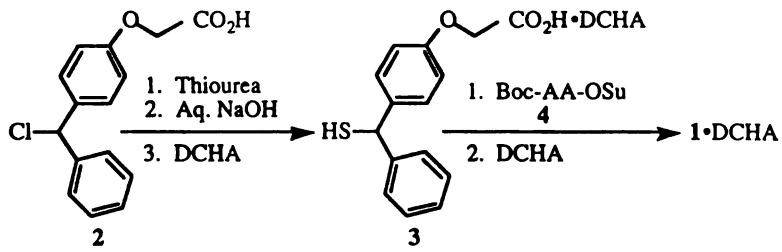
Peptides of up to 60 amino acids can be efficiently produced using stepwise solid phase peptide synthesis (SPPS).[1] Beyond this range, the accumulation of low level by-products makes purification of substantial amounts of the desired peptide difficult.[2] Recently, techniques for the chemical ligation of unprotected peptide segments have been developed that overcome the length limitations of SPPS.[3-7] This approach uses chemoselective reaction between two unique, mutually reactive functionalities to join unprotected peptide segments and gives the target product directly in final form. Several examples of functional proteins made by chemical ligation employ a C-terminal thioacid as a reactive functionality on one of the peptide segments to be reacted.[3-5,8]

The generation of peptide C-terminal thioacids in Boc chemistry SPPS requires the use of a special thioester linker between the peptide and the resin (**1**).[9,10] The appropriate amino acyl form of this linker is coupled to the resin support via the indicated carboxyl group. The desired peptide is then assembled in a stepwise fashion from the C-terminal amino acid which was incorporated into the linker prior to coupling to



1

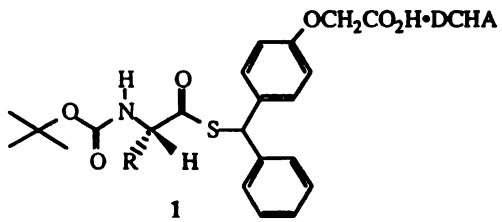
Scheme 1



the resin. Acidolytic cleavage in HF generates a peptide with a C-terminal -$^{\alpha}$COSH.[3-5,8-10]

We report a generalized, more convenient version of the Blake[9] and Yamashiro[10] procedure for the synthesis of thioacid-producing linkers. The Blake and Yamashiro method requires the conversion of Boc-amino acid succinimide esters to the corresponding Boc-amino thioacids with hydrogen sulfide. We have developed an alternative method, based on that of Hojo and Aimoto,[11] that uses the Boc-amino acid succinimide esters directly, avoiding the inconvenience and hazards of hydrogen sulfide gas. In this method (**Scheme 1**), thiol **3** is generated from the reaction of chloride **2**[10] with thiourea, followed by

Table 1. Boc-L-Amino thioester linkers.



1

Starting material (4)[14]	Product (1) R	% yield of 1 based on 3	Product (1) formula[a]
a. Boc-L-Gly-OSu	-H[15]	57	$C_{34}H_{48}N_2O_6S$
b. Boc-L-Ala-OSu	-CH_3[16]	74	$C_{35}H_{50}N_2O_6S$
c. Boc-L-Aba-OSu[b]	-CH_2CH_3[17]	70	$C_{36}H_{52}N_2O_6S$
d. Boc-L-Leu-OSu	-$CH_2CH(CH_3)_2$[18]	65	$C_{38}H_{56}N_2O_6S$

[a]Compounds were analyzed for C, H, N, S. Values were within ±0.4% of the theoretical value for the indicated molecular formula. [b]Aba = α-amino-n-butyric acid.

hydrolysis of the resulting thiouronium salt in aqueous base.[12] Thiol **3** is a general intermediate which can then be reacted with any of a wide range of Boc-amino acid succinimide esters (**4**) to produce the desired thioester linker (**1**) which is conveniently isolated as the dicyclohexylamine (DCHA) salt.

This method has been employed in the synthesis of several different Boc-amino thioester linkers, examples of which are shown in **Table 1**. This convenient new method is a simple and safe alternative to the previously published procedures,[9,10] and in combination with new chemical ligation strategies[3,5,13,19] will find a wide application in peptide chemistry.

Experimental Procedures

4-(α-Mercaptobenzyl)phenoxyacetic acid, DCHA (3). A mixture of **2**[10] (7.5 g, 27 mmol), thiourea (2.3 g, 30 mmol), and ethanol (100 mL) were heated to reflux.[12] After 4 hours, conversion to the thiouronium salt was essentially complete as shown by TLC (90:5:5 $CHCl_3$:MeOH:AcOH). 10N NaOH (30 ml) was added and the reflux continued for 2-3 hours. After cooling to room temperature, the reaction mixture was concentrated *in vacuo* to approximately half the original volume, acidified with concentrated HCl (to pH 2.0), and extracted with EtOAc (4 x 30 mL). The combined EtOAc extracts were washed with saturated NaCl (1 x 30 mL) and dried over $MgSO_4$. The volatile materials were removed *in vacuo*. The resulting oil was dissolved in EtOAc (100 mL) and any insoluble material filtered. DCHA (6.0 mL, 30 mmol) was added to the filtrate with stirring. Within a few minutes, a white solid began to precipitate. Et_2O (150 mL) was added and the suspension cooled at -20 °C for several hours. The resulting white solid was filtered, washed with Et_2O, and dried under vacuum to give **3** (10.3 g, 23 mmol, 84%): ^1H NMR ($CDCl_3$): δ 7.30 (m, 7H), 6.82 (d, 2H, J=8.7 Hz), 5.39 (br s, 1H), 4.40 (s, 2H), 2.81 (m, 2H), 2.23 (br s, 1H, ex D_2O), 1.88-1.02 (comp m, 20H); FAB MS (cesium ion): calc for [$C_{27}H_{37}NO_3S$, H^+] 456.2572, found 456.2572. Anal. Calcd for $C_{27}H_{37}NO_3S$: C, 71.17; H, 8.18; N, 3.07; S, 7.04. Found: C, 71.11; H, 8.41; N, 3.08; S, 7.09.

General Synthesis of Boc-amino thioester linker (1), DCHA salt. A mixture of **3** (3.67 mmol), Boc-AA-OSu (**4**) (3.68 mmol), DIEA (5.74 mmol), DMF (35 mL) and CH_2Cl_2 (4 mL) was stirred at room temperature. After several hours, the initial white suspension completely dissolved to give a clear, colorless solution. After 24 hours, the reaction mixture was poured into 1N HCl (150 mL) and extracted with EtOAc (4 x 35 mL). The combined EtOAc extracts were washed with 1N HCl (2 x 30 mL), H_2O (1 x 30 mL), saturated NaCl (1 x 30 mL) and dried over $MgSO_4$. Volatiles were removed *in vacuo*. The resulting oil was purified by flash chromatography (925:50:25 $CHCl_3$:MeOH:AcOH) to give an oil contaminated with AcOH. To remove residual AcOH, the oil was dissolved in $CHCl_3$ (40 mL) and washed with 0.1 N HCl (7 x 10 mL), saturated NaCl (1 x 10 mL) and dried over $MgSO_4$. Volatiles were removed *in vacuo* to give **1** as an oil. This oil was dissolved in Et_2O (10 mL) to which was added DCHA (1 equiv). Hexane (100 mL) was added with stirring to separate the DCHA salt of **1** as a thick oil from any unreacted DCHA. Solvents were decanted from the oil and the oil dissolved in CH_2Cl_2 (30-40 mL). The resulting solution was concentated *in vacuo* to give the DCHA salt **1** as a white foamy solid.

Acknowledgment

This work was supported by funds from the NIH (S.B.H.K.). SMW was supported by NIH grant PO1 GM48870. We gratefully acknowledge the assistance of: Dr. Gary Suizdak and Dr. Raj K. Chandra, The Scripps Research Institute, for FAB MS measurements and elemental analyses, respectively.

References and Notes

(1) Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180-193.
(2) Kent, S. B. H. *Ann. Rev. Biochem.* **1988**, *57*, 957-989.
(3) Schnölzer, M.; Kent, S. B. H. *Science* **1992**, *256*, 221-225.
(4) Dawson, P. E.; Kent, S. B. H. *J. Am. Chem. Soc.* **1993**, *115*, 7263-7266.
(5) Baca, M.; Kent, S. B. H. *Proc. Natl. Acad. Sci.* **1993**, *90*, 11638-11642.
(6) Rose, K. *J. Am. Chem. Soc.* **1994**, *116*, 30-34.
(7) Muir, T. W.; Williams, M. J.; Ginsberg, M. H.; Kent, S. B. H. *Biochemistry* **1994**, *33*, 7701-7708.
(8) Williams, M. J.; Muir, T. W.; Ginsberg, M. H.; Kent, S. B. H. *J. Am. Chem. Soc.* **in press** (1994).
(9) Blake, J.; C.H., L. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4055-4058.
(10) Yamashiro, D.; Li, C. H. *Int. J. Pept. Protein Res.* **1988**, *31*, 322-334.
(11) Hojo, H.; Aimoto, S. *Bull. Chem. Soc. Jpn.* **1991**, *64*, 111-117.
(12) Koenig, N. H.; Sasin, G. S.; Swern, D. *J. Org. Chem.* **1958**, *23*, 1525-1530.
(13) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776-779.
(14) Boc-AA-OSu esters were purchased from Novabiochem (LaJolla, CA). Any succinimide esters not commercially available were made from the corresponding free acids by the method of Anderson [Anderson, G.W.; Zimmerman, J.E.; Callahan, F.M. *J. Am. Chem. Soc.* **1964**, *86*, 1839-1842.] Boc-Aba-OSu: ^{1}H NMR ($CDCl_3$): δ 5.00 (m, 1H, ex D_2O), 4.62 (m, 1H), 2.83 (s, 4H), 1.89 (m, 2H), 1.45 (s, 9H), 1.07 (t, 3H).
(15) **1a**: ^{1}H NMR ($CDCl_3$): δ 7.24 (m, 7H), 6.80 (d, 2H, J=8.6 Hz), 5.88 (s, 1H), 5.06 (m, 1H, ex D_2O), 4.39 (s, 2H), 4.05 (d, 2H, J=5.7 Hz), 2.81 (m, 2H), 1.87-1.05 (comp m, 29H); FAB MS (cesium ion): calc for [$C_{34}H_{48}N_2O_6S$, H^+] 613.3311, found 613.3308.
(16) **2a**: ^{1}H NMR ($CDCl_3$): δ 7.24 (m, 7H), 6.80 (d, 2H, J=8.6 Hz), 5.81 (s, 1H), 4.95 (m, 1H, J=8.0 Hz, ex D_2O), 4.39 (m, 3H), 2.82 (m, 2H), 1.88-1.06 (comp m, 31H); FAB MS (cesium ion): calc for [$C_{35}H_{50}N_2O_6S$, H^+] 627.3468, found 627.3445.
(17) **3a**: ^{1}H NMR ($CDCl_3$): δ 7.24 (m, 7H), 6.80 (d, 2H, J=8.0 Hz), 5.82 (s, 1H), 4.94 (m, 1H, J=7.5 Hz, ex D_2O), 4.39 (m, 3H), 2.83 (m, 2H), 1.92-1.09 (comp m, 31H), 0.89 (t, 3H, J=7.4 Hz).
(18) **4a**: ^{1}H NMR ($CDCl_3$): δ 7.22 (m, 7H), 6.81 (d, 2H, J=7.5 Hz), 5.80 (s, 1H), 4.78 (d, 1H, J=8.1 Hz, ex D_2O), 4.39 (m, 3H), 2.82 (m, 2H), 1.88-1.06 (comp m, 32H), 0.90 (d, 6H, J=6.0 Hz); FAB MS (cesium ion): calc for [$C_{38}H_{56}N_2O_6S$, H^+] 669.3937, found 669.3960.
(19) Canne, L.E.; Ferré-D'Amaré, A.R.; Burley, S.K.; Kent, S.B.H. *J. Am, Chem. Soc.* **in press** (1995).

(*Received in USA* 10 *November* 1994; *accepted* 19 *December* 1994)