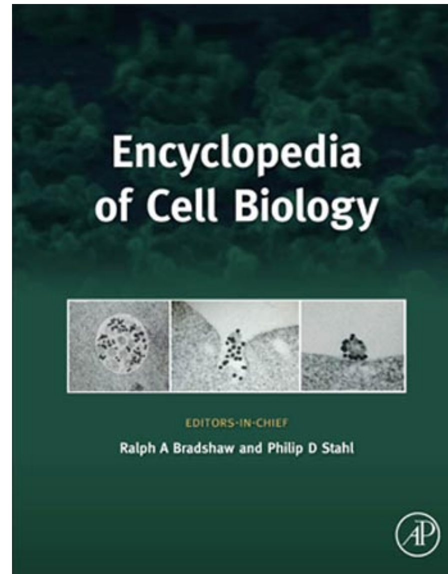
Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

This article was originally published in the *Encyclopedia of Cell Biology*, published by Elsevier, and the copy attached is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

http://www.elsevier.com/locate/permissionusematerial

Wlodawer A., and Jaskolski M., Inhibitors of HIV Protease. In: Ralph A Bradshaw and Philip D Stahl (Editors-in-Chief), Encyclopedia of Cell Biology, Vol 1, Waltham, MA: Academic Press, 2016, pp. 738-745.

© 2016 Elsevier Inc. All rights reserved.

Author's personal copy

PROTEIN SYNTHESIS/DEGRADATION: PROTEIN DEGRADATION – PATHOLOGICAL ASPECTS

Contents

Inhibitors of HIV Protease
Blood Pressure, Proteases and Inhibitors
Cancer – Proteases in the Progression and Metastasis
Lysosomal Diseases
Alpha-1-Antitrypsin Deficiency: A Misfolded Secretory Glycoprotein Damages the Liver by Proteotoxicity and Its Reduced Secretion Predisposes to Emphysematous Lung Disease Because of Protease-Inhibitor Imbalance

Inhibitors of HIV Protease

A Wlodawer, National Cancer Institute, Frederick, MD, USA
M Jaskolski, Adam Mickiewicz University, Poznan, Poland and Polish Academy of Sciences, Poznan, Poland

© 2016 Elsevier Inc. All rights reserved.

Glossary

Allosteric site A site other than a protein's (usually enzyme's) active site, which affects its activity when an effector molecule binds to it.

Aspartic protease An enzyme cleaving peptide bonds in proteins that has an active site consisting of two closely spaced, coplanar aspartate residues that activate a water molecule for the hydrolysis reaction. Enzymes belonging to the AA clan of aspartic peptidases, typified by pepsin found in gastric juice, are present in most organisms higher than bacteria, as well as in retroviruses, where they are called retropepsins and function as symmetric homodimers.

Bioavailability The fraction of an administered dose of a drug that reaches the systemic circulation, usually less than 100% in oral administration.

Diastereomer A stereoisomer of an optically active compound, in which the configuration of one or several chiral centers has been inverted (e.g., from *R* to *S*). Diastereomers should not be confused with (a pair of) enantiomers, in which the configuration of all chiral centers has been inverted. Consequently, enantiomers are mirror symmetric while diastereomers are not, and for this reason display different physical and chemical properties.

HIV (-1,-2) Human immunodeficiency virus (type 1 or 2), the causative agent of AIDS. The most common and virulent variant is designated HIV-1, whereas HIV-2 is more closely related to simian immunodeficiency virus and is less virulent.

HTLV Human T-cell leukemia virus, the first retrovirus to be associated with human disease (lymph node leukemia).

Integrase A virally encoded enzyme responsible for the incorporation of the double-stranded DNA copy of the retroviral genome into host cell genome. After integration, the infection is permanent and the infected cell cannot be cured.

Lipodystrophy A pathological condition characterized by abnormal or degenerative conditions or distribution of the body's fat.

M-PMV Mason–Pfizer monkey virus, a retrovirus similar to HIV, causing acquired immunodeficiency syndrome (AIDS)-like syndrome in macaque (rhesus) monkeys.

Nucleophile A chemical entity (such as a water molecule or hydroxyl group) rich in electrons, that is capable (usually after additional activation) of attacking another chemical group (such as the C atom of a peptide bond (O=)C–N) with depleted electron density, during a chemical transformation, such as hydrolysis.

Pepstatin A universal inhibitor of aspartic proteases, originally isolated from cultures of various species of *Actinomyces*.

Retrovirus A virus with a single-stranded mRNA genome, which is reverse-transcribed into DNA in a process that inverts the normal flow (DNA→RNA) of genetic information.

Reverse transcriptase (RT) A virally encoded enzyme that produces a double-stranded DNA copy of a retroviral RNA genome. RT is an RNA-dependent DNA polymerase, with ribonuclease activity.

RSV Rous sarcoma virus, a retrovirus causing sarcoma in chickens. The first retrovirus to be described, over a century ago, by Peyton Rous. The currently preferred name is avian sarcoma virus (ASV).

 doi:10.1016/B978-0-12-394447-4.10083-5

Author's personal copy

Introduction

The emergence of the acquired immunodeficiency syndrome (AIDS) epidemic in the early 1980s and the subsequent identification of the human immunodeficiency virus (HIV) as its causative agent brought into focus the need to accelerate research on retroviruses, in order to assist the efforts to create anti-HIV drugs. No such drugs were available during the decade after the first reports of the new disease started appearing and initially the disease itself was invariably lethal since the retrovirus, which infects T4 leukocytes, devastates the immune system and leads to its complete failure. With lack of any treatment options, the outlook was obviously very grim.

Retroviruses have been known for over a century, since the identification of an infectious agent causing cancer in chickens, later named after its discoverer Rous sarcoma virus (RSV) (Rous, 1911). However, the exact nature of the agent, of its life cycle and the mode of infectivity were not established until much later. It is now known that the genetic material of retroviruses consists of single-stranded RNA of positive polarity (mRNA) that becomes transcribed into DNA in a reverse (or retro) transcription reaction that is catalyzed by a retrovirus-specific and virally encoded reverse transcriptase (RT). Another retroviral enzyme, integrase (IN), incorporates the resulting double-stranded viral DNA into the host genome, thus making the genetic material of the retrovirus (provirus) a permanent part of the infected cell.

One of the unusual characteristics of many viruses, including all retroviruses, is that their proteins are not translated as individual final units, but rather as one or more large polyproteins that need to be processed (cleaved) into the mature viral enzymes and structural proteins. In retroviruses, the enzyme responsible for such an activity is retroviral protease (PR). Analyzing the retrotransposon and retroviral genomic sequences, including human T-cell leukemia virus (HTLV) (closely related to HIV) and RSV, Toh *et al.* (1985) found a single copy of a signature sequence D(S/T)G (aspartate, serine or threonine, and glycine) in the translated proteins that could be matched with the active-site motif of aspartic proteases from the pepsin family. Pepsin and similar cell-derived aspartic proteases, however, contain this motif in two copies in their pseudo-twofold-symmetric active site, each contributed by a separate domain of a single polypeptide chain. To reconcile these puzzling observations, Toh *et al.* postulated the presence of a pepsin-like protease in retroviruses, with the caveat that the retroviral proteases would need to be symmetric homodimers composed of two identical subunits. This hypothesis was consistent with an earlier speculation by Tang *et al.* (1978) that cell-derived aspartic proteases have arisen from smaller proteins by gene duplication and divergent evolution. The postulate about the homodimeric nature of retroviral proteases (retropepsins) was experimentally confirmed when the crystal structure of RSV protease was determined (Miller *et al.*, 1989a).

Without a functional protease (which could be inactivated by mutation or by inhibitors) the retrovirus is still capable of replicating, but only immature viral particles can be formed, which are noninfectious, i.e., cannot infect other T4 cells or other patients. Although inhibition of HIV PR does not cure the infection *per se*, it can be hoped that with prolonged treatment, inhibition of HIV PR will allow sufficient time for the elimination of the pool of the infected T4 cells.

Structure and Enzymatic Mechanism of HIV Protease

The structure of HIV-1 (type 1) PR itself, in unliganded state, was independently determined in 1989 by three groups (Navia *et al.*, 1989; Wlodawer *et al.*, 1989; Lapatto *et al.*, 1989). As in the case of RSV PR, HIV-1 PR is also a homodimer of two protein chains, 99 residues each. The protein fold of one subunit topologically resembles a truncated version of a single domain of pepsin, with the inter-subunit interface formed by four intertwined β strands from all termini (residues 1–4 and 95–99 from each subunit). The N-terminal strand is followed by two more β strands comprising residues 9–15 and 18–25, the latter including the catalytic Asp25. The next strand consists of residues 30–35 and is followed by a broad loop 36–42. Two more β strands, 43–49 and 52–58 form a flexible 'flap' loop that acts as a gating element for the active site. The second half of the HIV-1 PR subunit is topologically related to the first half and consists of β strands 52–66, 69–78, and 83–85. A prominent helix comprising residues 86–94 (absent in the N-terminal fold) has a clear counterpart in pepsins.

In the catalytic mechanism of aspartic proteases, a nucleophilic water molecule is activated by hydrogen bonding between the two (hemiprotonated) aspartate carboxylic groups. Such a catalytic water molecule was also found in the active sites of RSV PR (Miller *et al.*, 1989a) and HIV-1 PR (Wlodawer *et al.*, 1989) and its involvement in the enzymatic mechanism was gleaned from structural data (Jaskólski *et al.*, 1991).

Inhibitors of Retroviral Proteases

Strategies aimed at deactivation of HIV PR are focusing on relatively small chemical molecules that could inhibit its catalytic function. In a broad classification, the inhibitors can be covalent (irreversible) or non-covalent (reversible). Covalent inhibitors, such as molecules with a reactive oxirane (epoxy) group (Ro *et al.*, 1999) that attaches itself to the active-site aspartate, are not good candidates, even though they would block the enzyme permanently, because they can also modify many important host proteins. Non-covalent inhibitors typically compete with substrates for the active site, but bind more strongly and are resistant to cleavage. In the simplest approach, a competitive inhibitor is designed using as template the amino acid sequence of a good peptidic substrate (with some optimization), and replacing the scissile bond by a non-cleavable surrogate. Non-peptidic inhibitors use a similar strategy but the sequences of chemical moieties designed for docking in the respective enzyme binding sites do not have amino acid or peptidic character.

Whereas the structures of ligand-free retroviral proteases were crucial for delineating their structural and evolutionary relationship to pepsins, they were by themselves not sufficient to guide the design of potent inhibitors of HIV-1 PR.

Author's personal copy

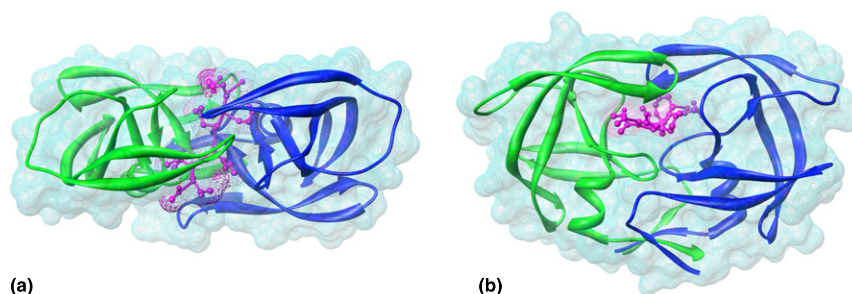


Figure 1 Two canonical views of HIV-1 PR complexed with MVT-101, the first inhibitor cocrystallized with this enzyme. The main-chain traces of the two subunits of the enzyme are colored green and blue, and the ball-and-stick model of the inhibitor is in magenta. The molecular surface of the protein is shown in light blue, and the surface of the inhibitor in pink. In (a) the homodimeric enzyme is viewed along its twofold axis. In (b) this axis is vertical. Figure prepared by Dr. Jiri Vondrasek.

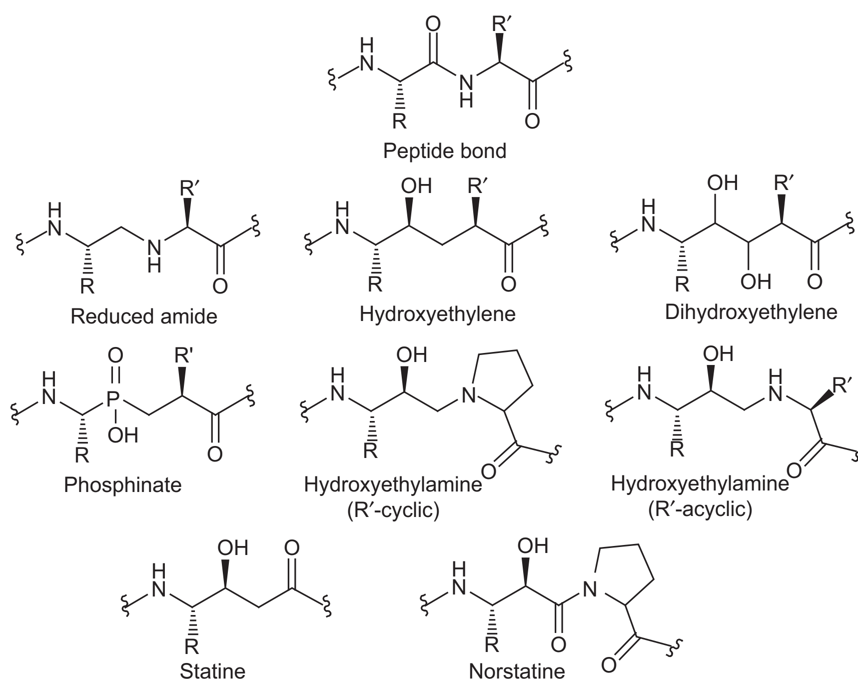


Figure 2 Chemical formulas of the peptide bond and its selected non-cleavable replacements.

Nevertheless, a model of an inhibitor complex was created on the basis of the structure of the apoenzyme of RSV PR superimposed on rhizopuspepsin, a fungal pepsin-like enzyme, in complex with a peptidic inhibitor with a reduced peptide replacing the cleavage site, His-Pro-Phe-His-Phe-ψ[CH-NH]Phe-Val-Tyr (where ψ denotes the reduced bond between the two Phe residues). Analysis of this model allowed successful delineation of seven subsites for the interactions between the side chains of the inhibitor and HIV-1 PR (Weber *et al.*, 1989).

Progress toward full understanding of the details of the interactions between the inhibitors of HIV-1 PR and the enzyme was boosted by the elucidation of the crystal structures of numerous complexes, now most likely going into thousands (although it is not possible to ascertain the exact count). The first such structure involved a complex with MVT-101 (Miller *et al.*, 1989b), a peptidic inhibitor created on the basis of a good natural substrate of HIV-1 PR with the sequence Ac-Thr-Ile-Met-Met-Gln-Arg.amide (K_m 1.4 mM) (**Figure 1**). In an approach established previously for the synthesis of peptidic inhibitors of human renin (which is a pepsin-like aspartic protease involved in the blood pressure control cascade), both methionines in this inhibitor were replaced by norleucine isosteres, and the scissile peptide between them was replaced by a reduced analog. The inhibition constant K_i for MVT-101 was 0.78 μM, indicating that this compound acted as a comparatively weak inhibitor and could not be considered as a good drug candidate, although it was powerful enough to form a stable complex with the protease during crystallization experiments.

Binding of MVT-101 led to quite substantial rearrangements of the structure of the enzyme, with the tips of the flaps, locked over the active-site-bound inhibitor, moving as much as 7 Å away from their positions in the free protease (where they were most likely fixed by crystal packing). In addition, in the complex with MVT-101 and all other peptidic inhibitors, a tightly bound water molecule was found at the interface

Author's personal copy

between the locked flap arms and the central part of the bound inhibitor, in variance with analogous pepsin-like complexes, where essentially only one flap is long enough to gate access to the active site. Although the HIV PR dimer is symmetric in the absence of a ligand with unique directionality, its binding induces some asymmetry which, if not transmitted to the surface during crystallization, leads to apparent twofold orientational disorder of the inhibitor that was noted in a number of structures. This binding, in turn, allowed delineation of the pockets that would accommodate the side chains of MVT-101. In the convention of Schechter and Berger (1967), in which the N-terminal (N-Pn…P1-) and C-terminal (-P1′…Pn′-C) substrate/inhibitor residues (linked by the P1-P1′ scissile bond/analog) are docked in the corresponding Sn…S1…S1′…Sn′ subsites of the enzyme, subsite S3 of HIV-1 PR included Arg8′, Asp29, and Gly48; subsite S2 was lined by Ala28, Ile47, Ile50′, and Ile84; and subsite S1 included Leu23′, Asp25′, Ile50, Pro81′, Val82′, and Ile84′. (Primed residues refer to the second subunit of the homodimer.) The subsites on the other side of the non-scissile P1-P1′ bond were generally similar with few exceptions. Subsite S1′ included Leu23′, Gly27, Asp25, Ile50′, Pro81, Val82, and Ile84; subsite S2′ consisted of Val32′, Ile47′, Gly48′, and Ile50; and subsite S3′ was surrounded by Arg8, Gly27′, Asp29′, Gly48′, and Val82.

A large number of structures of complexes with other peptide-based inhibitors that included non-scissile peptide mimetics of different chemistry (**Figure 2**) followed, providing a wealth of information about the plasticity of the enzyme subsites. The structure of a complex with acetyl-pepstatin, in which the scissile bond was replaced by the unnatural amino acid statine, elucidated the binding of this standard inhibitor of cell-derived aspartic proteases to HIV-1 PR (Fitzgerald *et al.*, 1990). Another early structure included a complex with JG-365, a hydroxyethylamine-containing peptide analog (Swain *et al.*, 1990). The structural data on a large number of inhibitors and inhibitor complexes of HIV PR were previously summarized in detailed reviews (Wlodawer and Erickson, 1993; Fitzgerald, 1993; Wlodawer and Vondrasek, 1998; Qiu and Liu, 2011).

First generation of HIV-1 PR Inhibitors Approved as AIDS Drugs

A complete summary of the status of the inhibitors of HIV-1 PR that have been approved for clinical use by the food and drug administration (FDA) can be found in Wikipedia. The first such inhibitor, approved on 6 December 1995, was saquinavir (**Figure 3**), developed by Hoffmann-La Roche. This compound, originally designated as Ro 31–8959, was the subject of intensive biochemical, biological, and structural studies (Craig *et al.*, 1991). It has a molecular weight (MW) of 671. In common with JG-365, saquinavir utilizes hydroxyethylamine as a non-cleavable peptide isostere. However, early crystallographic studies (Krohn *et al.*, 1991) suggested that the more potent version of this compound should be the *S* diastereomer of the –C*(OH)– chiral center, rather than the *R* form found in JG-365. Another major difference between these two compounds is the replacement of Pro, located in the P1′ position of JG-365, by DIQ ((*S*,*S*,*S*)-decahydroisoquinoline-3-carbonyl) in saquinavir. The structure of a saquinavir complex reported in 1991 (Krohn *et al.*, 1991) was deposited in the Protein Data Bank only in 1996 and released in 1997 (PDB ID 1HXB), but a number of other structures using the native and mutant forms of HIV-1 PR became available since then, including the atomic-resolution (0.97 Å) structure (**Figure 4(a)**) of a complex with the V82A mutant (PDB ID 2NMZ; (Tie *et al.*, 2007)). The main problem with the original formulation of the drug (sold under the name of Invirase) was poor bioavailability, as low as 3–4%. However, the later reformulation (brand name Fortovase) increased the bioavailability very significantly.

Three more inhibitors of HIV-1 PR gained FDA approval in 1996 and 1997 (Wlodawer and Vondrasek, 1998). Abbott Laboratories (now AbbVie) developed ritonavir (brand name Norvir), a relatively large (MW 721) peptidomimetic inhibitor (**Figure 3**) that was designed as a result of testing the concept of making the inhibitors fully, or nearly, C_2-symmetric. Even though ultimately such symmetric inhibitors turned out to be more difficult to manipulate for improved solubility and bioavailability, the concept resulted in the inclusion of a $–CH_2–C(OH)–$ group as a non-cleavable linker between two phenylalanines that occupy the S1 and S1′ subsites of the enzyme. Ritonavir has much better bioavailability than saquinavir, but it quickly became apparent that it had an unanticipated ability to act as a very potent inhibitor of cytochrome P450 (particularly of its isoform Cyp3A4), slowing down the removal of other pharmaceutical agents from circulation (Lea and Faulds, 1996). For that reason, ritonavir is currently used only as a booster for other protease inhibitors (see below).

Another peptidomimetic inhibitor of HIV-1 PR is indinavir (**Figure 3**; brand name Crixivan), developed by Merck (Vacca *et al.*, 1994). Slightly smaller (MW 614) than either saquinavir or ritonavir, it utilizes a hydroxyethylamine insert as a non-cleavable group. The compound is highly selective for HIV-1 PR (K_i 0.52 nM) and HIV-2 PR (K_i 3.3 nM) while it shows no inhibitory activity against any mammalian (including human) aspartic proteases.

An inhibitor that finally shed any peptidic character is viracept (**Figure 3**; brand name Nelfinavir). Although this drug utilizes the same DIQ group at its 'C terminus' as saquinavir, it does not retain any actual peptide bonds. Viracept was developed by Agouron Pharmaceuticals (Kaldor *et al.*, 1997), a company, now part of Pfizer, especially created with the task of introducing structure-based drug design (Appelt *et al.*, 1991). Viracept is smaller (MW 568) than the previous approved inhibitors of HIV-1 PR and was the first protease inhibitor approved for the treatment of pediatric AIDS.

The next inhibitor to reach the market, amprenavir (**Figure 3**; brand name Agenerase), was originally discovered at Vertex Pharmaceuticals, another venture capital company explicitly created to apply rational structure-based drug design in practice. The company was later acquired by GlaxoSmithKline, which became the distributor of this drug. Amprenavir is smaller still (MW 506) than the previous drugs but it was withdrawn from the market in 2004 after introduction of its prodrug form, fosamprenavir (brand name Lexiva (US) or Telzir (Europe)), in which a main-chain hydroxyl group was modified by phosphorylation. The prodrug is converted to amprenavir by host enzymes, thus slowing the release (and excretion) of the active form.