EDITOR'S CHOICE

Elucidation of the structure of retroviral proteases: a reminiscence

Mariusz Jaskolski[1], Maria Miller[2], J. K. Mohana Rao[2], Alla Gustchina[2] and Alexander Wlodawer[2]

1 Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University and Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland
2 Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD, USA

Determinations of only a very few protein structures had consequences comparable to the impact exerted by the structure of the protease encoded by HIV-1, published just over 25 years ago. The structure of this relatively small protein and its cousins from other retroviruses provided a clear target for a spectacularly successful structure-assisted drug design effort that offered new hope for controlling the then-escalating AIDS epidemic. This reminiscence is limited primarily to work conducted at the National Cancer Institute, and is not meant to be a comprehensive history of the field, but is rather an attempt to provide a very personal account of how the structures of this most thoroughly studied crystallographic target were determined.

An account of the discovery of the structure of retroviral protease (PR) must begin almost 30 years ago. It is a striking coincidence that this discovery occurred almost exactly 30 years after the first protein structure was announced in 1958 by Kendrew [1]. As is often the case, the start of the project was quite fortuitous – through an introduction in 1987 of a member of our team (A. Wlodawer) to J. Leis, who, at that time, worked at Case Western Reserve University, Cleveland, Ohio. Leis had been working for a long time on the biochemical characterization of various retroviral proteins, and had successfully purified milligram quantities of PR from Rous sarcoma virus (RSV) (now usually called avian sarcoma virus). Our team (Fig. 1) immediately decided to investigate its three-dimensional structure as a stand-in for the struc-

Abbreviations

HTLV, human T-cell lymphotropic virus; PDB, Protein Data Bank; PR, retroviral protease; RSV, Rous sarcoma virus.

Fig. 1. The RSV PR team at the Frederick Cancer Research Facility, Frederick, MD, in 1988. Left to right: Alexander Wlodawer, Maria Miller, Mariusz Jaskolski, and J. K. Mohana Rao.

ture of the medically much more important enzyme encoded by HIV-1. This was the first and the most important project for the newly established Crystallography Laboratory at the Frederick (Maryland) campus of the National Cancer Institute.

One should remember that the 1980s was a time of a looming global epidemic caused by an unknown virus that was triggering an invariably fatal disease, i.e. AIDS. This virus was considered to be the most dangerous emerging threat to human health. HIV-1 was conclusively shown to be the causative agent of AIDS only in 1983 [2], and its RNA genomic sequence was first published in 1985 [3]. It was shown that the HIV genome encodes only three enzymes, one of which is a PR. The PR plays a crucial role in the life cycle of HIV-1 (and of other related retroviruses, such as RSV), as it is needed to cleave the precursor viral polyproteins into mature, individual proteins (Fig. 2). Inactivation of retroviral PR was shown to prevent the viral particles from maturing into their infectious form [4], making PR a potential target for antiviral drugs. However, genuine proteins from HIV were very difficult to come by at that time, although, as we found out later, researchers at the Merck Sharp and Dohme Research Laboratories succeeded in the large-scale production and purification of recombinant HIV-1 PR [5].

RSV PR was known to be composed of 124 residues and to behave like an aspartic PR; for instance, it could be inhibited by pepstatin, which also inhibits pepsin, but other details were at all not certain. Retrospectively, one wonders how little faith we had in the information that was already at hand. These retroviral enzymes indeed looked like an aspartic PR, but there

was only one copy (rather than two) of the catalytic Asp-Thr-Gly (DTG) motif, and the enzyme was three times smaller than a typical pepsin-like aspartic PR. However, an article published ~ 10 years earlier [6] predicted that cell-encoded, two-domain, pepsin-like aspartic PRs might have evolved via gene duplication from much smaller, homodimeric ancestral enzymes. A hypothetical model of HIV-1 PR was built in 1987 [7] in a bold modeling exercise performed despite almost no sequence conservation between cell-derived and retroviral PRs. We knew that model, and looked at it with interest, but did not fully believe in its correctness.

Although we entered the race for the determination of the structures of retroviral PRs relatively late, a series of serendipitous events contributed to our success. Our laboratory was officially opened in November 1987, but setting up the wet laboratory and data collection facility took a few more months. The bench space for crystallization trials was kindly provided by the Laboratory of Chemical and Physical Carcinogenesis. There, the abundance of chemicals that could be used as additives to the 'classic' set of crystallization screening solutions helped us to find the right conditions for growing single crystals of RSV PR within 1 month. Importantly, the crystals diffracted to high resolution and were stable in the X-ray beam. Derivatization of the crystals with a uranyl compound, an excellent anomalous scatterer of Cu$K\alpha$ radiation, yielded a single-site derivative (which marked, as it later turned out, the active site) that enabled the proper choice of the space group enantiomorph and helped in setting the additional derivatives in common origin and handedness. The electron density map,

Fig. 2. The life cycle and partial genome organization of retroviruses. (A) A simplified schematic diagram of the life cycle of HIV-1. RNA molecules are in gray, DNA is in dark gray, polyproteins are shown as thick gray 'sausages', and surface glycoproteins are shown as small circles on stalks. The targets of therapeutic intervention are shown in gray rectangles. (B) Schematic organization of the *gag–pol* parts of the genomes of HIV-1 and RSV that are relevant to the activity of retroviral PR. These parts of the genomes are translated into two polyproteins, gag and pol, with the protease (shaded) located in pol of HIV-1 and gag of RSV. A frameshift or read-through is responsible for the division into the two ORFs. The sites of cleavage by PR are marked by vertical lines. Viral enzymes: IN, integrase; PR, protease; RT, reverse transcriptase. Other viral proteins: CA, capsid; MA, matrix; NC, nucleocapsid; RNaseH, ribonuclease H, part of the RT dimer in RSV, but cleaved off from one RT monomer in HIV-1. Whereas some details may differ for other retroviruses, the general principles of the activity and specificity of their PRs are the same.
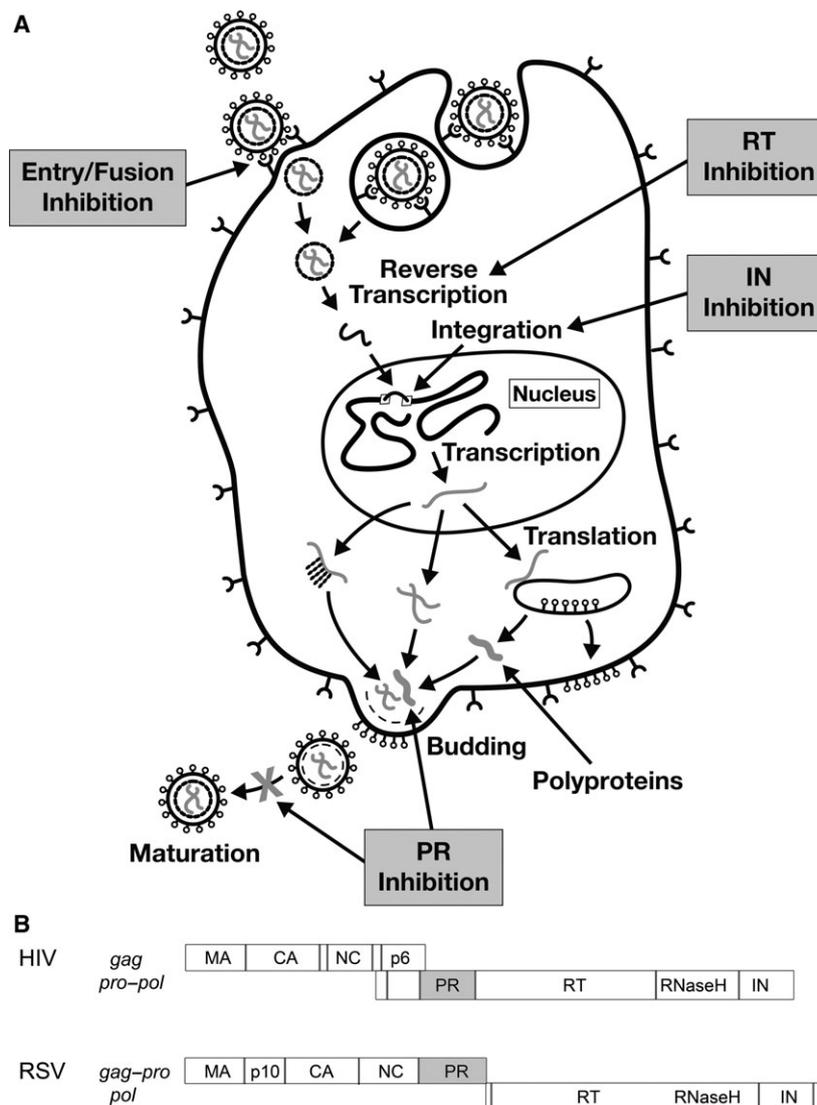
based on multiple isomorphous replacement phases from the four best derivatives, calculated by M. Miller, showed clearly the molecular boundary of a protein dimer and several characteristic sequence features, including the active site. The loop containing two tryptophans served as the starting point for the sequence fitting. Because the program packages that now make macromolecular crystallography quite routine were not available then, the programming skills of J. K. Mohana Rao were needed to identify the two protomers that make up the RSV PR dimer. The phases were further improved by noncrystallographic symmetry averaging, with a unique 'double-averaging technique' that is useful when the masks of the subunits are not clearly delineated. With the exception of eight residues in the flap region (see below), which were disordered in the crystal, the entire polypeptide chains

corresponding to the two protomers were defined by contiguous electron density in the final map. The atomic model of RSV PR was complete in October 1988.

The structure of RSV PR immediately showed that the early predictions were correct: we saw a dimeric aspartic PR resembling the monomeric two-domain pepsin (Fig. 3A). The active site had the same architecture, including the fireman's grip of the two DSG elements [8], and the catalytic water molecule could be seen between the aspartates. The similarity to pepsin allowed us to name the secondary structure elements as in a single domain of pepsin, and to note that there was yet another level of two-fold similarity faintly preserved in the structure of a single subunit and described in detail by Mohana Rao and Wlodawer [9]. However, there were also significant differences from
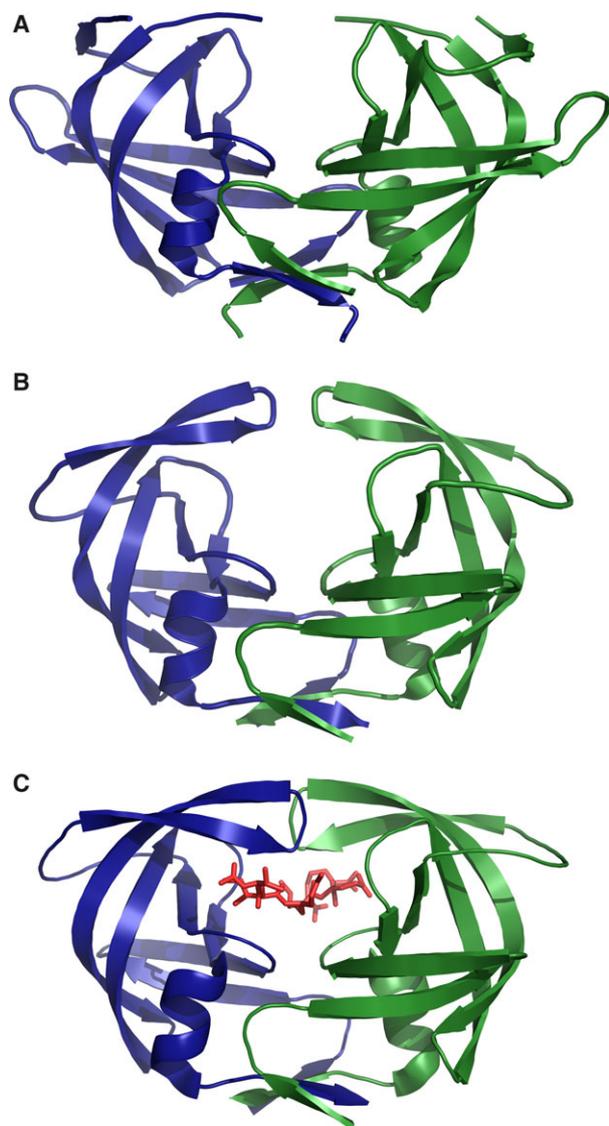
Fig. 3. Crystal structures of retroviral proteases. The two subunits of the homodimers are shown in blue and green. (A) The structure of RSV PR (PDB ID 2RSP) determined for the enzyme isolated directly from the virus. (B) The structure of synthetic HIV-1 PR (PDB ID 3HVP). The flap arms are lifted symmetrically over the empty active site. (C) The structure of a complex of synthetic HIV-1 PR with a substrate-based peptidic inhibitor MVT-101 (shown as red sticks), in which the scissile peptide bond was reduced (PDB ID 4HVP).

pepsin [10]. Because of the symmetry, both flaps of the homodimer were the same length and were prominent, although the tips in their elevated position over the empty active site were disordered in our RSV PR structure. The subunit interface was formed by a tight four-stranded antiparallel β-sheet woven from all the termini of molecules A and B, in the order N(A)–C (B)–C(A)–N(B).

As soon as the first RSV PR model was complete, our colleague, I. Weber, used it to skillfully build a homology model of the HIV-1 enzyme. The model looked very plausible, and had all the features of the template, with differences limited to the loop regions. The structure of RSV PR was published in *Nature* in early February 1989 [11]. A week later, in the same journal, the crystal structure of HIV-1 PR was unveiled by M. Navia, P. Fitzgerald and coworkers from Merck Sharp and Dohme [12], and, in that same week, Weber's model was published in *Science* [13]. After the first burst of joy, there was suddenly consternation, because the crystal structures of the RSV and HIV-1 PRs, although similar in their basic features, also showed some perplexing differences, especially in the C-terminal regions of the molecules. Where the RSV PR model had a clear α-helix, the HIV-1 PR structure had a straight β-strand, and the topology of the dimer interface was completely different. Instead of the interlaced termini with three intersubunit β-sheet connections found in RSV PR, the HIV-1 PR crystal structure had a hairpin with only one area of intersubunit contact, and a disordered N terminus. The latter difference was not trivial, because it had profound consequences for dimer stability and for the PR's ability to excise itself from the viral gag–pol fusion polyprotein (Fig. 2B) synthesized in the infected cell. Moreover, the question concerning the correct features of retroviral PR was not purely academic, because an accurate HIV-1 PR model was badly needed for the structure-guided design of inhibitors that might be developed into AIDS drugs. These points were clearly elucidated by Blundell and Pearl [14].

As mentioned above, the retroviral PR is translated as part of a polyprotein containing all of the structural (gag) and enzymatic (pol) proteins (Fig. 2B). For the virion particles to mature, all of the proteins, including the PR, must be liberated from the precursor polyproteins. This maturation process is carried out by the viral PR itself, which poses the puzzling topological question of how the PR can fold properly while still being embedded in the polyprotein, form an active dimer, and ultimately cut itself out, and accomplish all this in the restricted environment of the viral particle. The disorder of the N terminus suggested by the Merck model would allow PR excision not only in *trans* but even in *cis*, and, if correct, would provide an attractive paradigm for virion maturation, which is required for infectivity.

The situation in February 1989 became rather uncomfortable: which HIV-1 PR model should be used for designing AIDS drugs? Which one was correct? The dilemma could be resolved only by experiment,

but the question was where to obtain the protein. Help came from S. Kent, then at the California Institute of Technology, who was pioneering the methodology for synthesizing proteins with a purely chemical process [15]. He and J. Schneider quickly sent us 0.2 mg of chemically synthesized HIV-1 PR. This was enough to grow a few crystals with the modified protocol of McKeever [5]. Our molecular replacement calculations had to rely on Weber's model of HIV-1 PR, as the coordinates of the Merck structure were not made available. However, more material was needed to produce heavy-atom derivatives, because it was critical to obtain phase information experimentally, to avoid model bias and to produce an independent model of the protein. More protein was also needed for co-crystallization trials with inhibitors. The Kent group set a precedent by producing for us, within a period of just 2 weeks, milligram quantities of HPLC-purified enzyme for crystallographic studies [16]. It is worth noting that the enzyme was highly active and its autolysis, which occurred in the absence of an inhibitor, precluded formation of well-diffracting crystals. With only a few larger crystals at hand, the derivatization of synthetic HIV-1 PR, in which both cysteine residues were replaced by α-butyric acid, was not an easy task. M. Miller succeeded with an application of the same platinum compound that we had previously found bound to methionine residues in the crystals of RSV PR.

The electron density map based on two derivatives was of sufficient quality to demonstrate that all of the features deduced from the RSV PR structure were indeed present (Fig. 3B) and consistent with the model built by Weber. The dimer had fully visible, elevated flap arms, a tightly interlaced intersubunit β-sheet, the C-terminal α-helix, and even a well-defined water molecule between the catalytic aspartates. The definitive structure of the HIV-1 PR apoenzyme was published in *Science* in August 1989 [17]. The structure of the chemically synthesized HIV-1 PR was quickly confirmed by the structure of the recombinant enzyme, obtained solely by molecular replacement with a model built with RSV PR coordinates in Blundell's laboratory [18].

The structures of HIV-1 and RSV PRs from our laboratory were released for public access through the Protein Data Bank (PDB) promptly upon the completion of the refinement in 1989. In fact, the coordinates, including those of the Weber model, were made available to all genuine scientists who might be interested in using them as targets for inhibitor/drug design.

The next goal was to determine the structure of HIV-1 PR in complex with inhibitors. The first such inhibitor, MVT-101, was provided to us by G. Marshall (Washington University). Cocrystals with the synthetic enzyme grew overnight, and we were able to complete and publish the structure of the complex (Fig. 3C) 4 months after the publication of the structure of the apoenzyme [19]. It is worth stressing that the coordinates of the synthetic HIV-1 PR–MVT-101 complex were deposited in the PDB in April 1990, and, for the two most critical years, were the only ones freely available to all researchers worldwide who were working on the design of specific retroviral PR inhibitors.

The inhibitors initially utilized were the obvious choice: oligopeptides with the substrate sequence, but with the scissile peptide bond replaced by a nonhydrolyzable surrogate, such as a reduced peptide, or various hydroxylated ethyl groups [20]. Also, the existing inhibitors of cell-derived aspartic PRs, such as pepstatin or renin inhibitors designed as potential hypertension drugs, could be immediately tried. However, if selective inhibitors of retroviral PRs were to be found, they should not interfere with the host enzymes. Instead, they should exploit the unique features of retroviral PR, such as the perfect symmetry of the binding site, the existence of two flaps, or the presence of a structural water molecule in complex with peptidic inhibitors. This water molecule, with perfect tetrahedral coordination at the inhibitor–flap interface, was first observed in the crystal structure of the HIV-1 PR–MVT-101 complex [19]. Later, this interface water molecule was included in a novel class of inhibitors, based on the cyclic urea scaffold.

Some of the inhibitors were developed into potent drugs for treating HIV infection. The first HIV-1 PR inhibitor to become a drug, saquinavir (Ro-8959), was developed by Roche [21], and was approved for clinical use in December 1995 – only 6 years after the structure of the first inhibitor complex had been published, and < 7 years from the moment when an experimental model of the protein saw the light of day. This accomplishment marked a real triumph of structural biology, giving it the power to quickly lead to efficient therapies against a disease that, only a few years earlier, had been considered to be a global threat. By now, 10 PR inhibitors had gained Food and Drug Admistration approval for the treatment of HIV infection [22,23]. These molecules are known to be competitive inhibitors, meaning that they compete for the active site with PR substrates. MVT-101, the first competitive inhibitor to be structurally investigated in a complex with HIV-1 PR, was characterized by a micromolar dissociation constant. Picomolar and better inhibitors were subsequently developed through

fine-tuning to the enzyme binding sites. Competitive inhibition, however, was not the only option to be tried. One could also imagine irreversible modification of the active site or binding of the inhibitor molecule in a place other than the active site in order to, for instance, hinder flap closure or disrupt dimer formation. However, these options did not result in usable pharmacological agents.

When faced with a potent drug, the virus counteracts, and resistance to PR inhibitors arises through either selection of existing variants or mutations. An important aspect of the continuing structural research is to understand the mutations of this arms race, and design even more sophisticated drugs or combinations of such drugs. Another area of activity is focused on the structures of PRs from different retroviruses. In addition to RSV and HIV-1 PRs, the enzymes corresponding to HIV-2, simian immunodeficiency virus, feline immunodeficiency virus and equine infectious anemia virus PRs were studied, most of them in the National Cancer Institute laboratory [24,25]. These proteins, of course, share the same fold, domain organization, and quaternary structure. However, knowledge of the differences in their structural details, especially in the context of inhibitor complexes, also contributes to our understanding of drug resistance through sequence alterations. Detailed analysis of all known retropepsins as a subclass of aspartic proteases was primarily pioneered by A. Gustchina [26,27]. A more recent addition to the collection of retroviral PR structures is the enzyme from human T-cell lymphotropic virus (HTLV)-1 [28], a retrovirus that causes human leukemia. With this addition, the efforts to cure AIDS and cancer suddenly found common structural ground [29]. When the structure of HTLV-1 PR was solved, it became obvious why the AIDS drugs tried on patients with HTLV-related leukemia had no effect. Although HTLV-1 was discovered before HIV-1, the PR from HTLV-1 had resisted structural characterization for a long time, partly because of various crystallographic obstacles. For instance, the rmsd between the Cα traces of HTLV-1 PR and the molecules from other retroviruses is as high as 1.93 Å (RSV PR) and is 1.72 Å on average, complicating molecular replacement calculations and showing that, on closer observation, there are indeed significant variations in the canonical retroviral PR fold. Incidentally, similar rmsd values are obtained in comparisons with pepsin, albeit for a smaller number of superposed atoms. However, when only the atoms of the active sites are compared, the match is nearly perfect, with an rmsd of ~ 0.5 Å in superpositions of retroviral and cell-derived aspartic PRs.

An interesting variation on the theme was provided by the recently determined structure of the PR encoded by the xenotropic murine leukemia virus-related virus, with a typical retroviral PR fold but a completely different dimer interface [30]. Although the virus itself turned out to be created in a laboratory, the structure of its PR shows considerable deviation from the template.

Structural studies of a broad range of retroviral PRs have the added advantage of providing a different perspective on bottlenecks and obstacles. One such difficulty stems from the mixed blessing of the two-fold symmetry of retroviral PRs. With HIV-1 PR, this symmetrical arrangement has led to ambiguity of space group assignment and to a two-fold disorder of the bound inhibitors, and also caused difficulty in obtaining useful heavy-atom derivatives. This drawback was turned into a benefit when C_2 symmetric (or pseudosymmetric) inhibitors were synthesized.

As with many other PRs, a serious problem with HIV-1 PR is autodigestion on prolonged incubation, but this problem was overcome by mutations of the autolysis sites [31]. Another solution to this problem is a mutation, usually Asp→Asn, in the active site. In the simplest variant, this approach leads to the simultaneous replacement of both catalytic aspartates. However, asymmetric mutations also became possible through a clever engineering trick, in which the two subunits are tethered via a linker connecting the carboxy terminal of molecule A with the amino terminal of molecule B. The struggle to determine the crystal structure of a retroviral PR in its monomeric form reached an extraordinary conclusion when the structure of Mason–Pfizer monkey virus PR was solved with the massive help of 'citizen scientists' playing the protein-folding game Foldit on the Internet [32]. The objective of studying the monomeric form of the protein was to provide an innovative target for drug design that would prevent the formation of active dimer in the first place [33].

The structural studies of retroviral PRs over the last 27 years have generated a tremendous stimulus for structural biology and an enormous amount of information. HIV-1 PR has become the most studied protein, in structural terms, in the world. There are hundreds of structure determinations, and the actual count has been lost, despite early efforts at book-keeping in a designated database [34]. The overwhelming majority of the structures were determined by protein crystallography, but NMR has also been used to determine structures.

HIV-1 PR has helped to advance the frontiers of structural biology in many different ways. We have

already mentioned the demonstration that total chemical synthesis can be an option for making proteins for structural studies, and that such synthetic proteins can fold spontaneously outside of the context of a biological cell. Chemical synthesis was also used by Kent to obtain the D-enantiomer of HIV-1 PR crystallized by M. Miller. It was possible to demonstrate that this mirror twin of the natural enzyme behaves identically in a looking-glass world [16]. The growth of huge single crystals of HIV-1 PR has enabled neutron diffraction experiments, which, by visualization of the hydrogen atoms, confirmed the protonation of one of the aspartates in the active site and the deprotonation of the other, thus providing direct experimental evidence for the catalytic mechanism [35]. HIV-1 PR has already been characterized at a breathtaking, ultrahigh resolution of 0.84 Å [36].

It is generally recognized that determining the structure of HIV-1 PR has been the springboard for the development of rational drug design strategies, taking them from a flimsy dream to reality. This accomplishment, which led to the expeditious finding of effective therapies for an incurable and fatal disease, is among the major scientific achievements of the last century. There were several fortunate circumstances that contributed to this success. Personally, we very much cherish one aspect that is sometimes overlooked, namely, the openness and willingness of the members of our group to share scientific information. Such openness, among other things, helps to quickly correct errors, which are inevitable, and leads to success that benefits all.

Author contributions

All authors contributed equally to writing this manuscript.

References

1 Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H & Phillips DC (1958) A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662–666.

2 Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C *et al.* (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868–871.

3 Wain-Hobson S, Sonigo P, Danos O, Cole S & Alizon M (1985) Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**, 9–17.

4 Kohl NE, Emini EA, Schleif WA, Davis LJ, Heimbach JC, Dixon RA, Scolnick EM & Sigal IS (1988) Active

human immunodeficiency virus protease is required for viral infectivity. *Proc Natl Acad Sci USA* **85**, 4686–4690.

5 McKeever BM, Navia MA, Fitzgerald PM, Springer JP, Leu CT, Heimbach JC, Herbert WK, Sigal IS & Darke PL (1989) Crystallization of the aspartylprotease from the human immunodeficiency virus, HIV-1. *J Biol Chem* **264**, 1919–1921.

6 Tang J, James MNG, Hsu IN, Jenkins JA & Blundell TL (1978) Structural evidence for gene duplication in the evolution of the acid proteases. *Nature* **271**, 618–621.

7 Pearl LH & Taylor WR (1987) A structural model for the retroviral proteases. *Nature* **329**, 351–354.

8 Pearl L & Blundell T (1984) The active site of aspartic proteinases. *FEBS Lett* **174**, 96–101.

9 Rao JKM & Wlodawer A (1990) Is the pseudo-dyad in retroviral proteinase monomers structural or evolutionary? *FEBS Lett* **260**, 201–205.

10 Rao JKM, Erickson JW & Wlodawer A (1991) Structural and evolutionary relationships between retroviral and eucaryotic aspartic proteinases. *Biochemistry* **30**, 4663–4671.

11 Miller M, Jaskólski M, Rao JKM, Leis J & Wlodawer A (1989) Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature* **337**, 576–579.

12 Navia MA, Fitzgerald PM, McKeever BM, Leu CT, Heimbach JC, Herber WK, Sigal IS, Darke PL & Springer JP (1989) Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337**, 615–620.

13 Weber IT, Miller M, Jaskólski M, Leis J, Skalka AM & Wlodawer A (1989) Molecular modeling of the HIV-1 protease and its substrate binding site. *Science* **243**, 928–931.

14 Blundell T & Pearl L (1989) Retroviral proteinases. A second front against AIDS. *Nature* **337**, 596–597.

15 Schneider J & Kent SB (1988) Enzymatic activity of a synthetic 99 residue protein corresponding to the putative HIV-1 protease. *Cell* **54**, 363–368.

16 Miller M (2010) The early years of retroviral protease crystal structures. *Biopolymers* **94**, 521–529.

17 Wlodawer A, Miller M, Jaskólski M, Sathyanarayana BK, Baldwin E, Weber IT, Selk LM, Clawson L, Schneider J & Kent SBH (1989) Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science* **245**, 616–621.

18 Lapatto R, Blundell T, Hemmings A, Overington J, Wilderspin AF, Wood S, Merson JR, Whittle PJ, Danley DE, Geoghegan KF *et al.* (1989) X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes. *Nature* **342**, 299–302.

19 Miller M, Schneider J, Sathyanarayana BK, Toth MV, Marshall GR, Clawson L, Selk L, Kent SBH &

Wlodawer A (1989) Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* **246**, 1149–1152.

20 Wlodawer A & Erickson JW (1993) Structure-based inhibitors of HIV-1 protease. *Annu Rev Biochem* **62**, 543–585.

21 Craig JC, Duncan IB, Hockley D, Grief C, Roberts NA & Mills JS (1991) Antiviral properties of Ro 31-8959, an inhibitor of human immunodeficiency virus (HIV) proteinase. *Antiviral Res* **16**, 295–305.

22 Wlodawer A & Vondrasek J (1998) Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu Rev Biophys Biomol Struct* **27**, 249–284.

23 Wlodawer A & Jaskolski M (2015) Inhibitors of HIV protease. In Encyclopedia of Cell Biology (Bradshaw RA & Stahl P, eds), Elsevier, Amsterdam, The Netherlands. in press.

24 Wlodawer A, Gustchina A, Reshetnikova L, Lubkowski J, Zdanov A, Hui KY, Angleton EL, Farmerie WG, Goodenow MM, Bhatt D *et al.* (1995) Structure of an inhibitor complex of the proteinase from feline immunodeficiency virus. *Nat Struct Biol* **2**, 480–488.

25 Gustchina A, Kervinen J, Powell DJ, Zdanov A, Kay J & Wlodawer A (1996) Structure of equine infectious anemia virus proteinase complexed with an inhibitor. *Protein Sci* **5**, 1453–1465.

26 Gustchina A, Sansom C, Prevost M, Richelle J, Wodak SY, Wlodawer A & Weber IT (1994) Energy calculations and analysis of HIV-1 protease-inhibitor crystal structure. *Protein Eng* **7**, 309–317.

27 Wlodawer A & Gustchina A (2000) Structural and biochemical studies of retroviral proteases. *Biochim Biophys Acta* **1477**, 16–34.

28 Li M, Laco GS, Jaskólski M, Rozycki J, Alexandratos J, Wlodawer A & Gustchina A (2005) Crystal structure of human T-cell leukemia virus protease, a novel target for anti-cancer drug design. *Proc Natl Acad Sci USA* **102**, 18322–18337.

29 Gustchina A, Jaskolski M & Wlodawer A (2006) Lessons learned fighting HIV can be applied to anti-cancer drug design. *Cell Cycle* **5**, 463–464.

30 Li M, Gustchina A, Matuz K, Tozser J, Namwong S, Goldfarb NE, Dunn BM & Wlodawer A (2011) Structural and biochemical characterization of the inhibitor complexes of xenotropic murine leukemia virus-related virus protease. *FEBS J* **278**, 4413–4424.

31 Johnson EC, Malito E, Shen Y, Pentelute B, Rich D, Florian J, Tang WJ & Kent SB (2007) Insights from atomic-resolution X-ray structures of chemically synthesized HIV-1 protease in complex with inhibitors. *J Mol Biol* **373**, 573–586.

32 Khatib F, DiMaio F, Cooper S, Kazmierczyk M, Gilski M, Krzywda S, Zabranska H, Pichova I, Thompson J, Popovic Z *et al.* (2011) Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nat Struct Mol Biol* **18**, 1175–1177.

33 Gilski M, Kazmierczyk M, Krzywda S, Zabranska H, Cooper S, Popovic Z, Khatib F, DiMaio F, Thompson J, Baker D *et al.* (2011) High-resolution structure of a retroviral protease folded as a monomer. *Acta Crystallogr D* **67**, 907–914.

34 Vondrasek J, van Buskirk CP & Wlodawer A (1997) Database of three-dimensional structures of HIV proteinases. *Nat Struct Biol* **4**, 8.

35 Adachi M, Ohhara T, Kurihara K, Tamada T, Honjo E, Okazaki N, Arai S, Shoyama Y, Kimura K, Matsumura H *et al.* (2009) Structure of HIV-1 protease in complex with potent inhibitor KNI-272 determined by high-resolution X-ray and neutron crystallography. *Proc Natl Acad Sci USA* **106**, 4641–4646.

36 Kovalevsky AY, Liu F, Leshchenko S, Ghosh AK, Louis JM, Harrison RW & Weber IT (2006) Ultra-high resolution crystal structure of HIV-1 protease mutant reveals two binding sites for clinical inhibitor TMC114. *J Mol Biol* **363**, 161–173.