**Chemical Synthesis of an Enzyme Containing an Artificial Catalytic Apparatus**

**Vladimir Torbeev**<sup>A,B</sup> and **Stephen B. H. Kent**<sup>A,C</sup>

<sup>A</sup>Department of Chemistry, The University of Chicago, 929 East 57th Street, Chicago, IL 60637, USA.

<sup>B</sup>Current address: Laboratory of Protein Chemistry, Institut de Science et d’Ingénierie Supramoléculaires, Université de Strasbourg, 8 Allée Gaspard Monge, BP 70028, 67083 Strasbourg ( Cedex ), France.

<sup>C</sup>Corresponding author. Email: skent@uchicago.edu

With the goal of investigating electronic aspects of the catalysis of peptide bond hydrolysis, an analogue of HIV-1 protease was designed in which a non-peptide hydroxy-isoquinolinone artificial catalytic apparatus replaced the conserved Asp<sup>25</sup>–Thr<sup>26</sup>–Gly<sup>27</sup> sequence in each 99-residue polypeptide chain of the homodimeric enzyme molecule. The enzyme analogue was prepared by total chemical synthesis and had detectable catalytic activity on known HIV-1 protease peptide substrates. Compared with uncatalyzed hydrolysis, the analogue enzyme increased the rate of peptide bond hydrolysis by ~10<sup>12</sup>-fold. Extensions of this unique approach to the study of enzyme catalysis in HIV-1 protease are discussed.

**Introduction**

The human immunodeficiency virus type 1 protease (HIV-1 PR) is a retroviral-encoded aspartic acid proteinase<sup>[1]</sup> The 22 kDa HIV-1 PR protein is a homodimer made up of two identical 99-residue polypeptide chains that form a single enzyme molecule that binds a polypeptide substrate (Fig. 1). Two mobile flap structures, one from each monomer, close down over the central six amino acid residues of the bound substrate.<sup>[2]</sup> Two catalytically essential side chain carboxyl groups are contributed by aspartic acid proteinases, including HIV-1 PR, involves general acid–general base catalysis by the side chain carboxyl groups of Asp<sup>25</sup> and Asp<sup>57</sup> (Fig. 2).<sup>[4]</sup> One (protonated) carboxyl donates a proton to the carbonyl oxygen of the scissile peptide bond in the bound substrate, while the other (ionized) carboxylate activates the nucleophilic water by removing a proton, generating a gem-diol tetrahedral intermediate. In the breakdown of the tetrahedral intermediate, the roles of the carboxyl groups are reversed: the now-protonated carboxyl donates a proton to the nitrogen atom leaving group to give one peptide fragment, while the ionized carboxylate removes a hydroxyl proton from the gem diol, to give a terminal carboxyl group on the other fragment of the cleaved substrate peptide.

It has long been appreciated that a detailed understanding of this general acid–general base mechanism could usefully be investigated by controlling the ionization properties of the two catalytic Asp β-carboxyl groups in the HIV-1 PR protein molecule. Key aspects of catalysis could be tested by varying the electronic properties of the carboxyl functionalities proposed to play the roles of general acid–general base. An initial study has been reported in which non-coded β-substituted catalytic Asp residues were introduced into the HIV-1 protease polypeptide chain by use of misacylated tRNAs.<sup>[3]</sup>

Our ultimate goal is to design and build an HIV-1 PR protein molecule containing a catalytic apparatus that will enable us to systematically tune the electronic properties of the protonated or ionized oxygen atoms that are directly involved in promoting the attack of the water nucleophile on the scissile peptide bond within the enzyme–substrate complex, and to measure effects on the rate of peptide bond hydrolysis. These data would provide a linear free energy relationship<sup>[5]</sup> revealing how $pK_a$ properties of the catalytic side chain carboxyl groups in the HIV-1 PR protein molecule affect catalysis in this prototypic aspartic acid proteinase enzyme. Note that this approach is not simply the incorporation into the protein of a ‘non-coded’ amino acid replacement for a catalytic amino acid; the essence of our approach was to design and build a non-peptide replacement moiety that can assume the precise geometric and electronic roles provided by the array of residues at the active site of an aspartyl proteinase enzyme.

In order to introduce an –O(H) catalytic moiety with tunable electronic properties, we use an ortho-substituted phenol(ate) in...
place of the Asp$_{25}$ side chain carboxyl (Fig. 3). The phenol O atom can align closely with the catalytically active tele-O atom of the Asp side chain β-carboxyl moiety.

Examination of the structure of the HIV-1 protease with a bound substrate-derived inhibitor suggested that there is sufficient room in the complex to be able to introduce ortho substituents on a suitably oriented phenol moiety in the enzyme’s active site. Thus encouraged, we then addressed how to introduce such a correctly oriented phenol moiety into the HIV-1 PR protein molecule.

The unique conformation of the peptide backbone of the conserved, catalytically essential –Leu$_{24}$–Asp$_{25}$–Thr$_{26}$–Gly$_{27}$– sequence in the HIV-1 protease (Fig. 4a) is important for interaction of the two monomers in the homodimeric protein molecule, and is critical for bringing the catalytic Asp side chains into the correct relative orientation for effective catalysis.[7] Energy-minimized modeling, using the interactive program Sculpt,[8] led us to consider fused aromatic and reduced rings as a mimic of the –Leu$_{24}$–Asp$_{25}$–Thr$_{26}$–Gly$_{27}$– backbone conformation, while at the same time providing a platform for decoration with desired catalytic –OH and other functional groups needed to introduce the artificial catalytic apparatus into the synthetic enzyme protein molecule. A suitably functionalized dihydroisoquinolinone fused ring entity is shown in Fig. 4b. In Fig. 4c, the functionalized dihydroisoquinolinone is shown superimposed on the –Leu$_{24}$–Asp$_{25}$–Thr$_{26}$–Gly$_{27}$– sequence as found in the HIV-1 PR protein molecule.

In order to validate this approach to investigating the chemistry of catalysis in HIV-1 PR, a simplified version of the artificial catalytic apparatus lacking a side chain substituent at carbon 3 was chosen as the initial entity to be introduced into the enzyme protein molecule by total chemical synthesis.

Here, we report the chemical synthesis of an HIV-1 PR protein molecule containing a 7-OH isoquinolinone catalytic moiety in place of the amino acid sequence Asp$_{25}$–Thr$_{26}$–Gly$_{27}$, and measurement of its catalytic properties. The implications of this work for understanding enzyme catalysis are discussed, and future research directions to refine and extend the approach are suggested.

**Experimental and Results**

The amino acid sequence of the polypeptide chain of the 7-OH-isoquinolinone moiety-containing HIV-1 PR is shown in Scheme 1a. The target polypeptide chain was prepared in a convergent fashion, by native chemical ligation (NCL) of two unprotected peptide segments, and the full-length synthetic

---

**Fig. 1.** The HIV-1 PR protein molecule. (Top) A ribbon representation of the homodimeric enzyme complexed to reduced isostere inhibitor MVT-101, based on Protein Data Bank (PDB) code 4HVP. Atoms of the two catalytic Asp$_{25}$ and Asp$_{25}$ side chains and inhibitor MVT-101 are shown as CPK space-filling representation. (Bottom) View from above the enzyme–inhibitor complex.

**Fig. 2.** Formal chemical mechanism of peptide bond hydrolysis by HIV-1 PR. General acid–general base catalysis by side chain carboxyls of Asp$_{25}$/Asp$_{25}$ in the enzyme–substrate complex.

**Fig. 3.** A phenol moiety designed to mimic the catalytic Asp side chain in the active site of the HIV-1 PR. Line drawing of the Asp side chains involved in catalyzing the attack of the water nucleophile on the scissile peptide bond in the substrate, with an ortho-substituted phenol moiety superimposed in magenta lines. Although shown for only one of the two catalytically active Asp side chains for the sake of clarity, various forms of the enzyme can be made with one or both catalytically active Asp side chains replaced by the phenol moiety.
polypeptide chain was then folded to form the homodimeric protein molecule (Scheme 1c). Synthesis of a control HIV-1 PR was carried out contemporaneously in identical fashion, using the same target polypeptide chain amino acid sequence except that the 7-OH-isooquinolinone moiety was replaced by residues Asp$^{25}$–Thr$^{26}$–Gly$^{27}$.

**Synthesis**

Protected building block 1 (Scheme 1b) was generously provided by Paul Alewood and Aline Dantas de Araujo [10] and was used in stepwise solid phase peptide synthesis (SPPS) to prepare the (1–40)$^{\text{thioester}}$ peptide segment, as described for the contemporaneous synthesis of the control HIV-1 PR by the same route. [7]

Building block 1 (160 mg, 0.321 mmol) was used (~2-fold molar excess, coupling with 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU)) in a 0.15 mmol scale synthesis of (1–40)$^{\text{thioester}}$ segment, to produce synthetic peptide of quality essentially identical to that of the control (1–40)$^{\text{thioester}}$ peptide segment.

NCL reaction with the (Cys$^{41}$-99) segment, prepared as previously described [7], was followed by alkylation with 2-bromoacetamide to give the desired product [(7-OH)dihydroisoquinolinone, des(25–27)]HIV-1 PR (1–99) polypeptide in crude form (see Supplementary Material).

**Purification and Characterization**

The crude polypeptide product was purified by reverse-phase HPLC. Appropriate precautions (including the use of disposable labware, meticulously cleaned equipment, and virgin HPLC columns) were taken to avoid contamination by traces of control HIV-1 PR polypeptide chain. After HPLC purification, the isolated yield of the [(7-OH)dihydroisoquinolinone, des(25–27)]HIV-1 PR (1–99) polypeptide chain was 8 mg. The purified polypeptide chain had the expected mass within experimental uncertainty, and was of high purity as determined by electrospray ionization mass spectrometry (MS) taken across the entirety of the single UV-absorbing HPLC peak.

**Folding**

Attempted folding of the polypeptide chain by dialysis using conditions previously established for folding of HIV-1 protease [11] led to near-quantitative precipitation of aggregated polypeptide. To avoid precipitation, the concentration of glyc erol was increased to 30–50 % (v/v) and the inhibitor MVT-101 was added in 100-fold molar excess. The enzyme was found to be more stable in solution at lower pH, such that no visible precipitation occurred at pH 5.2, whereas at pH 6.2, significant amounts of white precipitate were observed. The amount of folded [(7-OH)dihydroisoquinolinone, des(25–27)]HIV-1 PR

---

**Fig. 4.** Design of a non-peptide catalytic apparatus for HIV-1 PR. (a) Backbone and side chain conformations of the conserved sequence –Leu$^{24}$–Asp$^{25}$–Thr$^{26}$–Gly$^{27}$– from each monomer (yellow; green) in the active site region of HIV-1 PR. (b) A bicyclic phenol(ate) chiral mimic of the 3D stereochemical configuration of the catalytic apparatus of the HIV-1 PR. The stereochemical centre at C3 is intended to have the same relative configuration as the Cα of L-Thr (i.e. as if C4 were the Nα of the L-Thr residue). The stereochemistry at * should have the same configuration as the side chain of L-threonine. (c) Artificial bicyclic chiral ‘phenolate’ catalytic apparatus (magenta), superimposed on the 3D configuration of the catalytic sequence –Leu$^{24}$–Asp$^{25}$–Thr$^{26}$–Gly$^{27}$– (yellow) as found in the HIV-1 PR protein molecule. (d) CPK space-filling representation of the artificial catalytic apparatus modelled into HIV-1 PR structure based on coordinates of PDB code 4HVP. Orange sphere is the tunable X moiety. Modeling shows that there is sufficient space available for incorporation of various substituents at position X.
protein remaining in solution at pH 6.2 was quantified by HPLC using an OD 214 nm calibration curve based on control HIV-1 PR, the concentration of which was established by elemental analysis (total nitrogen content).

Catalytic Activity

We used synthetic peptide substrates based on our previous work. The peptides mimicked the viral Gag-Pol p24/p15 (GHKARVL-AEAMSVQTVNATIM-MQGRGNFRQKR) and p17/p24 (RRSNQVSQNY-PIVQNIQGRR) cleavage sites. Assays were performed using high pressure liquid chromatography-electrospray ionization mass spectrometry (LC-MS) to identify and quantify the cleavage products. The TIM-MQR junction was cleaved most rapidly; others were slower, which is in the same order as for wild-type HIV-1 PR. Prior to assay, MVT-101 was removed by dialysis. Concentrations were 0.5 mM for substrates and 7 μM for the [(7-OH)dihydroisoquinolinone, des(25–27)]HIV-1 PR analogue protein, which is 100–1000 times higher than that used to assay the control HIV-1 PR. The derived kinetic constants for the [(7-OH)dihydroisoquinolinone, des(25–27)]HIV-1 PR analogue were turnover number \( k_{cat} = 1 \times 10^{-3} \) s\(^{-1}\) and Michaelis constant \( K_m = 0.4 \) mM at pH 6.2 and are compared with the kinetic constants for the control HIV-1 PR in Table 1.

Discussion

In the work reported here, we designed an unnatural catalytic apparatus for the HIV-1 protease to allow systematic tuning of the electronic properties of the O atoms that are directly involved in the general acid–general base mechanism of peptide bond hydrolysis in the enzyme–substrate complex. A bicyclic aromatic and reduced fused ring system containing a phenolic –OH with a neighbouring \( \text{ortho} \) substituent \( X \) was designed to mimic the peptide backbone conformation in the active site, and to retain natural stereochemistries at key atoms. This bicyclic fused ring moiety, appropriately functionalized, was designed to replace the conserved –Asp25–Thr26–Gly27 tripeptide sequence in the enzyme’s active site. The substituent \( X \) can be a proton, an electron-donating group (e.g. methyl or methoxy), or an electron-withdrawing group (e.g. F or NO\(_2\)); these substitutions will enable control over the electronic properties of the catalytic phenol or phenolate O atom.

In order to provide proof of principle for this approach to the elucidation of key aspects of the chemistry of enzyme catalysis in HIV-1 PR, we used total chemical synthesis to prepare HIV PR ‘ART’, an enzyme analogue that contained a simplified version of the designed unnatural catalytic apparatus, one that

---

**Table 1. Enzymatic activities**

<table>
<thead>
<tr>
<th></th>
<th>( k_{cat} ) [s(^{-1})]</th>
<th>( K_m ) [μM]</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV PR ‘ART’(^\dagger)</td>
<td>1 \times 10(^{-3})</td>
<td>400</td>
</tr>
<tr>
<td>Control HIV-1 PR</td>
<td>23</td>
<td>25</td>
</tr>
</tbody>
</table>

\(^\dagger\)HIV PR ‘ART’ = [(7-OH)dihydroisoquinolinone, des(25–27)]HIV-1 PR analogue.
lacked the atoms corresponding to the side chain of Thr$^{26}$. The HIV PR ‘ART’ synthetic protein had detectable catalytic activity when measured using synthetic peptide substrates derived from natural Gag-Pol cleavage sites. The turnover number $k_{\text{cat}}$, which is a measure of the efficiency of chemical catalysis in the enzyme–substrate complex, was reduced by a factor of $\approx 10000$-fold compared with the control HIV-1 protease assayed under similar conditions. Nonetheless, this maximal velocity of catalysis by HIV PR ‘ART’ represents an $\approx 10^3$-fold acceleration of peptide bond hydrolysis over the non-catalyzed reaction. This catalytic activity of HIV PR ‘ART’ compares favourably with the rate accelerations observed for other designed protein enzymes, such as catalytic antibodies and the activities of de novo designed enzyme molecules.

A justifiable concern about the results reported here is that, although stringent precautions were taken to prevent it, the observed activity of HIV PR ‘ART’ may have arisen from contamination by trace amounts of the control HIV-1 PR. We do not believe that to be the case, a belief that is bolstered by the large increase in the $K_m$ value observed for the HIV PR ‘ART’ enzymatic activity: 400 μM (versus a $K_m$ of 25 μM for control HIV-1 PR). Contamination by 1 part in 10000 of control HIV-1 PR would have left $K_m$ unchanged. Furthermore, the observed catalytic activity of HIV PR ‘ART’ was pH-dependent. At pH 6.2 after 30 h, ~30% of the TIM-MQR... site was cleaved, whereas at pH 5.2, only 3% cleavage was observed under otherwise identical conditions. This is consistent with catalysis arising from a phenolic –OH with $pK_a$ greater than 9, rather than what is observed with the control HIV-1 PR where the rate increases at pH 5.2 compared with pH 6.2. Both the elevated $K_m$ and the reduced activity at lower pH are consistent with the observed enzymatic activity being an intrinsic property of the HIV PR ‘ART’ protein molecule itself.

Why is HIV PR ‘ART’ less catalytically active than HIV-1 PR? The extensive precipitation observed when attempting to fold the HIV PR ‘ART’ polypeptide chain dramatically illustrated that the homodimeric protein containing the simplified version of the unnatural catalytic apparatus is much less stable than the native homodimer. The reason for this is readily evident: the native sequence –Asp$^{25}$-Thr$^{26}$-Gly$^{27}$– adopts a backbone conformation that provides multiple H-bonds between the two monomers, stabilising the homodimeric enzyme molecule (Fig. 5). In this so-called ‘fireman’s grip’, there are strong inter-monomer H-bonds that involve the side chain OH of each Thr$^{26}$ that are absent in HIV PR ‘ART’, because the atoms corresponding to the side chain of atoms of Thr$^{26}$ were omitted in order to simplify the synthesis of the unnatural catalytic apparatus.

How can the design of an HIV-1 PR-derived enzyme containing an artificial catalytic apparatus with tuneable electronic properties be improved, and how can the rigour with which the catalytic properties of such an enzyme are studied be increased? First, the Thr side chain atoms must be included in the functionalized 7-OH dihydroisooquinolinone used to replace Asp$^{25}$/Asp$^{25}$ (see Fig. 4b, c). This will restore some of the H-bonds found in the fireman’s grip network, and thus increase the stability of the HIV PR ‘ART.v2’ homodimeric protein molecule and allow a more thorough investigation of its catalytic properties over a wider (and higher) pH range. Determination of the structure of the resulting HIV PR ‘ART.v2’ protein molecule, with bound substrate-derived peptide inhibitors, should be carried out at high resolution by X-ray crystallography. That will allow determination of the disposition of the catalytic phenolic O atoms with respect to the scissile peptide bond in the substrate, and establish whether their postulated role in catalysis is reasonable. Use of a ketomethylene, hydrateable inhibitor will be particularly revealing.

With an HIV PR ‘ART.v2’ that contains a fully functionalized artificial catalytic apparatus in hand, linear free energy relationship studies should be carried out, with $X = H$, CH$_2$, F, NO$_2$, etc. Molecular modeling studies have shown that there is room in the active site of the HIV-1 PR protein molecule for $X = F$, or CH$_3$, without steric clashes (Fig. 4d); for $X = NO_2$, there is some apparent crowding with O-to-O contact distances ~0.2 Å closer than the van der Waal’s contact radii. Whether or not the enzyme will tolerate these clashes can only be determined by experiment. Further extensions of these studies would include the chemical synthesis of a covalent dimer form of HIV PR ‘ART’; that would enable the use of different catalytic moieties in each monomer of the enzyme, e.g. a covalent dimer with one monomer containing ‘ART.v2’ catalytic moiety and the other monomer containing the native Asp$^{25}$, or different ‘X’ groups in each monomer of an ‘ART.v2’ covalent dimer.

Conclusion
The use of total chemical synthesis to create an enzyme with an artificial catalytic apparatus as described here can give ‘at-will’ control of an enzyme’s catalytic properties. It provides an enzyme protein molecule in which catalysis is subject to the direct and systematic control of the investigator, and will give fundamental insights into electronic and other aspects of the mechanism of enzyme catalysis in the aspartic acid proteinases.

Supplementary Material
Technical details of the synthesis and characterization of peptide segments and the target HIV-1 PR analogue protein are available on the Journal’s website.

Conflicts of Interest
The authors declare no conflicts of interest.

Acknowledgements
This research did not receive any specific funding.
References


