Analytical Biochemistry 377 (2008) 162–169

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Investigating human immunodeficiency virus-1 proteinase specificity at positions P_4 to P_2 using a bacterial screening system

Petra Schlick [1], Tim Skern *

Max F. Perutz Laboratories, Medical University of Vienna, Dr. Bohr-Gasse 9/3, A-1030 Vienna, Austria

ARTICLE INFO

Article history:
Received 5 December 2007
Available online 14 March 2008

Keywords:
Random screening
Substrate specificity
Combinatorial library
Proteolytic enzyme
Antiviral compound

ABSTRACT

Inhibitors of human immunodeficiency virus-1(HIV-1) proteinase have been used for several years to treat acquired immunodeficiency syndrome patients. Despite intensive research, however, the substrate specificity of this enzyme is not completely elucidated. Here, we assessed the HIV-1 proteinase P_4 to P_2 substrate specificity using a bacterial screening system. In this system, the bacterial enzyme β-galactosidase has been transformed into an HIV-1 proteinase substrate by insertion of the p6/PR cleavage site. Consequently, HIV-1 processing can be determined by measuring the β-galactosidase activity on X-gal plates and by examination of the extent of cleavage of the β-galactosidase protein itself. We screened a library containing randomized sequences at the P_4 to P_2 positions and found strong preferences for Thr, Ser, and Pro at P_4, for Leu, Met, and Phe at P_3, and for Ser, Met, and Leu at P_2. The frequent observations of Thr at P_4 and Ser at P_2 extend previous findings and offer the possibility of producing inhibitors with different properties. These new data on HIV proteinase specificity illustrate the usefulness of random libraries in the genetic screening system. This approach can be applied to examine any proteinase that has a recognition site extending across several amino acids.

Human immunodeficiency virus (HIV)[2] was identified as the causative agent of acquired immunodeficiency syndrome (AIDS) almost 25 years ago [1,2]. In the meantime, AIDS has become well established as one of the top five leading causes of death worldwide. Despite intensive research efforts, a vaccine against HIV is still not available. However, the introduction of highly active antiretroviral therapy has led to significant declines in mortality rates. Currently approved antiretrovirals target vital enzymes of the HIV replication cycle such as reverse transcriptase and HIV-1 proteinase (HIV-1^{pro}) and the viral–cell fusion event (reviewed in [3]).

HIV-1^{pro} is responsible for the processing of the translated polyprotein; its inhibition leads to the release of noninfectious particles [4]. In contrast to cellular aspartic proteinases, HIV-1^{pro} is capable of cleaving at Tyr-Pro or Phe-Pro bonds. This unique feature of the enzyme provided an attractive target for drug design [5]. Initially, HIV-1^{pro} inhibitors were successfully designed as peptidomimetics of the viral peptide substrates with noncleavable structures in the scissile bond [6]. The crystal structure of HIV-1^{pro} [7–9] provided a basis for the design of new inhibitors to optimize the peptide-based substrates. Eight of the nine currently available inhibitors are based on this concept [10].

Initial use of one proteinase or one reverse transcriptase inhibitor (monotherapy) resulted in rapid development of drug resistance [11]. Use of a combination of inhibitors (triple therapy) lowered resistance development. However, even with triple therapy, a complete suppression of the appearance of drug-resistant viruses was not possible [12]. Therefore, the demand for novel developments is still very high. This need has led for example to the synthesis of the novel class of lysine sulfonamides active against HIV-1^{pro} [13]. To continue to improve currently available peptidomimetics, deeper insight into the essential determinants of substrate recognition is clearly required.

Cleavage by HIV-1^{pro} is performed at unique amino acid sequences on the viral polyprotein [14]. Comparison of cleavage site sequences suggests broad substrate specificity and a lack of consensus sequence; however, certain general features emerge. For instance, hydrophobic amino acids are preferred at P_1 and $P_{1'}$ (nomenclature of [15]), aliphatic and Glu/Gln residues are often found at P_2, aromatic residues are almost never found at $P_{3'}$, and small residues are preferred at P_2. Several sequences contain an aromatic residue at P_1 followed by Pro at $P_{1'}$. Accordingly, HIV-1^{pro} cleavage sites have been divided into two types, type 1 having Tyr(Phe)*Pro and type 2 having hydrophobic residues except Pro at the site of cleavage [16]. These types of cleavage sites were proposed to have different preferences for the P_2 and $P_{2'}$ positions

* Corresponding author. Fax: +43 1 4277 9616.
E-mail address: timothy.skern@meduniwien.ac.at (T. Skern).

[1] Present address: Intercell AG, Campus Vienna Biocenter 6, A-1030 Vienna, Austria.

[2] *Abbreviations used:* HIV, human immunodeficiency virus; AIDS, acquired immunodeficiency syndrome; eIF, eukaryotic initiation factor.

0003-2697/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2008.03.011

[17,18]; however, most of the specificity of HIV-1^{pro} is unexplored. It has been suggested that the preference for a residue at a particular position in the substrate depends strongly on the neighboring residues [19]. Furthermore, the analysis of crystal structures of six substrate complexes has shown that substrate shape rather than a particular amino acid sequence determines specificity [20].

Random screening systems provide a valuable tool to select for efficiently cleaved substrates and take into account their dependency on the surrounding residues. Thus, efficiently cleaved substrates have been identified by taking advantage of phage displays for stromelysin and matrelysin [21], human collagenase 3 [22], membrane type-1 matrix metalloproteinase [23], human kallikrein 2 [24], and HIV-1^{pro} [25]. Furthermore, chimeric Sindbis virus libraries have been used to identify cleavage sites of hepatitis C virus [26]. In this study, we investigated HIV-1^{pro} substrate specificity at positions P_4, P_3, and P_2 using a random cassette library introduced into β-galactosidase [27] as a reporter protein.

Materials and methods

Cloning procedures and library construction

The screening vector pET11c Δp6HIV^{proD} lacZ (Fig. 1A), containing genes for both the enzyme (HIV-1^{pro}) and the substrate (modified ß-galactosidase), was constructed as follows. The plasmid pCITE Δp6HIV^{proD} containing the cDNA for two covalently linked HIV-1^{pro} monomers preceded by 20 amino acids from the C terminus of the transframe region [28] was described in [29]. The Δp6HIV^{proD} expression cassette was amplified by standard PCR techniques and subcloned into pET11c (Novagen) using *Nde*I and *Bam*HI restriction sites. To support Δp6HIV^{proD} expression in *Escherichia coli* strain MC1061, the T7 promoter within pET11c Δp6HIV^{proD} was replaced with the *lac* promoter of pBluescript I KS (Stratagene), taking advantage of unique *Bgl*II and *Xba*I sites. The ß-galactosidase expression cassette was derived from plasmid pLex/LacZ (Invitrogen) and cloned blunt ended into the *Ava*I restriction site of pET11c Δp6HIV^{proD} to generate pET11c Δp6HIV-^{proD} lacZ (Fig. 1A).

To introduce specific cleavage sites for HIV-1^{pro} into the β-galactosidase protein, we made use of the *Bsu*36I restriction enzyme site (Fig. 1A), as described by Baum et al. [27]. This site lies at nucleotide 239 of the lacZ gene, corresponding to amino acid Glu80 of the β-galactosidase protein. It is unique in pET11c Δp6HIV^{proD} lacZ (Fig. 1A). Two sets of oligonucleotides with overhanging *Bsu*36I ends were designed corresponding to the HIV-1^{pro} cleavage sites at the p6/PR junction (5′-TGAAGTAAGCTTTAACTTCCCTCAG ATCACTCTGGG-3′ and 5′-TCACCCAGAGTGATCTGAGGGAAGTTAAA GCTTACT-3′) and within eIF4GI (5′-TGAAAATCGCTTCTCAGCCCTT CAACAAGCGGTAGG-3′ and 5′-TCACCTACCGCTTGTTGAAGGGCTGA GAAGCGATTT-3). The respective oligonucleotides were annealed to each other by heating in a PCR thermocycler to 94 °C for 5 min, followed by slow cooling to room temperature.

The *Bsu*36I site was also used to introduce randomized oligonucleotides to construct the library of mutated sequences at the P_4, P_3, and P_2 positions of the HIV-1^{pro} cleavage site. The method of Reidhaar-Olson and Sauer [30] was used to synthesize the oligonucleotides. This strategy completely randomizes the first two nucleotides of each of the three codons; at the third nucleotide of each codon, only two substitutions are allowed. This method allows the generation of 31 codons that still encode all 20 amino acids [30]. The theoretical number of clones required for a complete coverage of all possible sequences for 3 mutated amino acids is 6^4 multiplied by 2^3, that is approximately 33,000. The sense oligonucleotide to randomize the positions P_4, P_3, and P_2 in the p6/PR cleavage site was (5′-TGAAGTA(AGCT)(AGCT)(GC)(AGCT)(AGCT)(GC)(AGCT)(AGCT)(GC)TTCCCTCAGATCACTCTGGG-3′). In the antisense oligonucleotide, inosine was provided at the positions corresponding to the randomized sites to enable annealing (5′-TCACCCAGAGTGATCTGAAGGGAIIIIIIIIITACT-3′ (I = inosine)). The oligonucleotides were annealed as described above for the specific oligonucleotides.

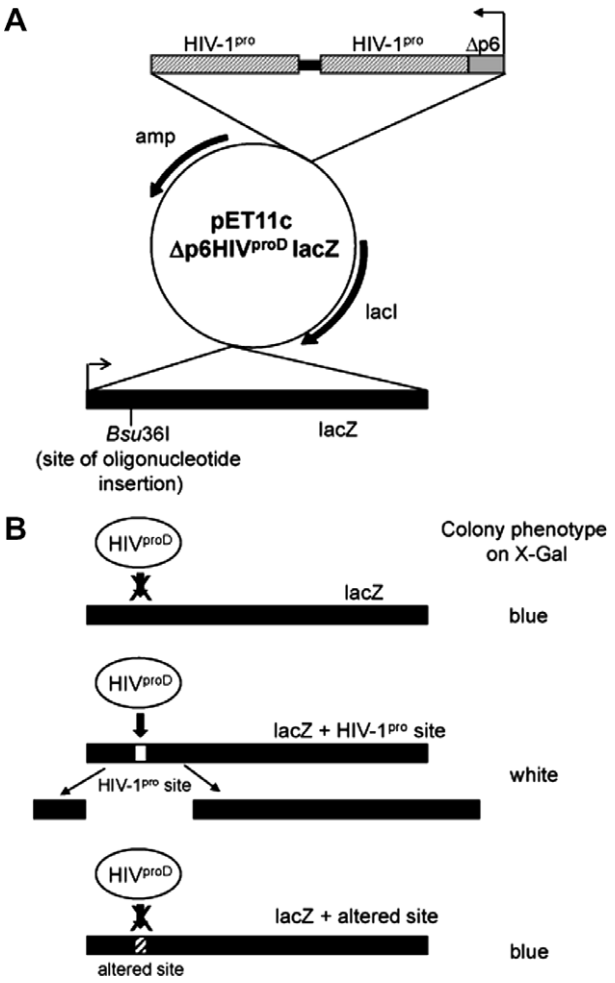

Fig. 1. Screening vector and system. (A) Plasmid pET11c Δp6HIV^{proD} lacZ used to screen HIV-1^{pro} activity on modified β-galactosidase proteins. The tethered HIV-1^{pro} dimer and the β-galactosidase expression cassette were inserted into the pET11c backbone. The HIV-1^{pro} coding sequence is expressed under control of the lac promoter (closed arrow), whereas the β-galactosidase gene is expressed from the P_L promoter (open arrow). On the β-galactosidase expression cassette, the site of oligonucleotide insertion (*Bsu*36I) is indicated. HIV-1^{pro} monomers are shaded, the transframe peptide p6 is gray and the β-galactosidase is black. (B) Schematic illustrating the blue/white color selection during screening. The β-galactosidase protein is black, the HIV-1^{pro} site is white, and the noncleavable site is shaded. White colonies on X-Gal are obtained only when the HIV-1^{proD} recognizes its cleavage site and as a consequence inactivates the β-galactosidase enzyme, preventing the cleavage of X-gal.

Bacterial transformation and screening

The specifically modified pET11c Δp6HIV^{proD} lacZ plasmids and the ligations for the random library were transformed into the β-galactosidase negative *E. coli* strain MC1061 using standard procedures. Blue-white selection to monitor β-galactosidase cleavage (Fig. 1B) was performed on LB agar plates supplemented with 100 μg/ml ampicillin and 40 μg/ml 5-bromo-4-chloro-3-indolyl β-galactoside (X-Gal). Plasmid DNA of selected clones was amplified to provide sufficient material for subsequent sequence analysis.