Volume 268, number 1, 35–38 FEBS 08631 July 1990

Sub-site preferences of the aspartic proteinase from the human immunodeficiency virus, HIV-1

Jan Konvalinka[1], Petr Strop[1], Jiri Velek[1], Vera Cerna[1], Vladimir Kostka[1], Lowri H. Phylip[2], Anthony D. Richards[2], Ben M. Dunn[3] and John Kay[2]

[1]*Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Science, Flemingovo namesti 2, 166 10 Praha 6, Czechoslovakia,* [2]*Department of Biochemistry, University of Wales College of Cardiff, PO Box 903, Cardiff CF1 1ST, Wales, UK* and [3]*Department of Biochemistry and Molecular Biology, J. Hillis Miller Health Centre, University of Florida, Gainesville, FL 32610, USA*

Received 10 May 1990

A series of synthetic, chromogenic substrates for HIV-1 proteinase with the general structure Ala-Thr-His-Xaa-Yaa-Zaa*Nph-Val-Arg-Lys-Ala was synthesised with a variety of residues introduced into the Xaa, Yaa and Zaa positions. Kinetics parameters for hydrolysis of each peptide by HIV-1 proteinase at pH 4.7, 37°C and $u = 1.0$ M were measured spectrophotometrically and/or by reverse phase FPLC. A variety of residues was found to be acceptable in the P_3 position whilst hydrophobic/aromatic residues were preferable in P_1. The nature of the residue occupying the P_2 position had a strong influence on k_{cat} (with little effect on K_m); β-branched residues Val or Ile in this position resulted in considerably faster peptide hydrolysis than when e.g. the Leu-containing analogue was present in P_2.

HIV-1 proteinase; Chromogenic substrate; Subsite; Specificity

1. INTRODUCTION

The aspartic proteinase encoded within the human immunodeficiency virus (HIV-1) genome is of vital importance in the maturation of infectious virions. It has thus become a strategic target for the development of specific inhibitors with considerable potential as therapeutic agents for the treatment of AIDS [1]. In order to facilitate the generation of such compounds, much attention has been focussed on structure/activity relationships in this important enzyme. Many reports (summarised in [2]) have indicated how synthetic peptides can be used as substrates to mimic the cleavage junctions that are required to be processed naturally in the viral polyprotein. Indeed, recently it has been shown that, by substitution of a *p*-nitrophenylalanine residue for the naturally-occurring amino acid in the P$_1'$ position, sensitive, soluble chromogenic substrates can be produced for the viral proteinase [3,4]. Assays using such substrates not only facilitate the convenient monitoring of enzymic activities but also permit detailed investigations in vitro into the active site preferences that are required to be met for effective interaction. In the present report, we describe further investigations into the preferences of the S_4, S_3, S_2 and S_1 sub-sites of the HIV-1 proteinase using a new series of such substrates.

Correspondence address: J. Kay, Department of Biochemistry, University of Wales College of Cardiff, PO Box 903, Cardiff CF1 1ST, Wales, UK

Abbreviations: Nle, norleucine; Nph, 4-NO_2-phenylalanine (The nomenclature system of Schechter and Berger [12], i.e. P_5-P_4-P_3-P_2-P_1*P_1'-P_2'-P_3', is used to depict amino acids adjacent to the residues in the P_1 and P_1' positions which contribute the scissile peptide bond (indicated by an asterisk *))

2. MATERIALS AND METHODS

The peptides examined as substrates were based initially upon consensus sequences of the cleavage junction pol P63/P32 of Rous sarcoma virus [5]. These contained the sequence Ala-Thr-His-Xaa-Yaa-Zaa*Nph-Val-Arg-Lys-Ala where a number of substitutions were introduced in the Xaa, Yaa and Zaa positions (Table I). The peptides were synthesised and purified by reverse phase HPLC as described previously [3]. They were all readily soluble in water; aqueous stock solutions of approximately 3 mM were thus prepared.

Homogenous preparations of HIV-1 proteinase were obtained as described previously [3]. The active site concentration of each sample used in the assays was determined by titration with a compound (Quinoline-2-carbonyl-Asn-Phe ψ[CHOH-CH_2N]-decahydro-isoquinoline-carbonyl-NH-*t*-butyl) which has been shown to have a subnanomolar K_i for its interaction with the HIV-1 proteinase [1]. Values for k_{cat} were derived from $V_{max} = k_{cat} \cdot [E]$.

Where possible, hydrolysis of the peptide substrates was monitored spectrophotometrically at 300 nm and 37°C as described previously [3]. The buffer used was 0.1 M sodium acetate, pH 4.7, containing 4 mM EDTA and sufficient NaCl to give a final ionic strength of 1 M. Mercaptoethanol (5 mM) was added to some assays but had no influence on the kinetic parameters determined. A final volume of 800 μl was used and the operational proteinase concentration was in the region of 5–10 nM.

When slow rates of hydrolysis were observed, kinetic parameters were obtained by following the time course of the proteolytic reactions by product appearance on reverse phase FPLC (as described previously [6]). Reaction conditions were the same as for the spectrophotometric measurements.

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

Table I

Substrates used for active site investigations with HIV-1 proteinase

Number	Sequence		
1	Ala-Thr-His-Gln-Val-Arg	*	Nph-Val-Arg-Lys-Ala
2	Ala-Thr-His-Gln-Val-Glu	*	Nph-Val-Arg-Lys-Ala
3	Ala-Thr-His-Gln-Val-Met	*	Nph-Val-Arg-Lys-Ala
4	Ala-Thr-His-Gln-Val-Phe	*	Nph-Val-Arg-Lys-Ala
5	Ala-Thr-His-Gln-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
6	Ala-Thr-Pro-Gln-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
7	Ala-Thr-Pro-Cys-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
8	Ala-Pro-Pro-Pro-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
9	Ala-Thr-His-Tyr-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
10	Ala-Thr-His-Arg-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
11	Ala-Thr-His-Val-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
12	Ala-Thr-His-Asn-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
13	Ala-Thr-His-Asp-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
14	Ala-Thr-His-Glu-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
15	Ala-Thr-His-Pro-Val-Tyr	*	Nph-Val-Arg-Lys-Ala
16	Ala-Thr-His-Gln-Ile-Tyr	*	Nph-Val-Arg-Lys-Ala
17	Ala-Thr-His-Gln-Phe-Tyr	*	Nph-Val-Arg-Lys-Ala
18	Ala-Thr-His-Gln-Ala-Tyr	*	Nph-Val-Arg-Lys-Ala
19	Ala-Thr-His-Gln-Ala-Tyr	*	Nph-Val-Arg-Lys-Ala
20	Ala-Thr-His-Gln-Gly-Tyr	*	Nph-Val-Arg-Lys-Ala

Initial velocities were measured for at least 7 substrate concentrations and the kinetic constants (K_m, V_{max}) were derived from a computer fit of the data using the Enzfitter program. In all cases, the values given for K_m and V_{max} are the means of at least two separate determinations. The precision of each individual estimation was in the range ± 5-20%. The identity of the (unique) peptide bond undergoing hydrolysis in each substrate was established by collecting the product peaks from reverse phase FPLC followed by amino acid analysis [3].

3. RESULTS AND DISCUSSION

The nature of the amino acid residue in the P_1 position (i.e. contributing to the -P_1*P_1'-scissile peptide bond) was first investigated in the series Ala-Thr-His-Gln-Val-Zaa*Nph-Val-Arg-Lys-Ala.

When Zaa was Arg or Glu (peptides 1 and 2; Table I), no hydrolysis was observed (Table II). However, insertion of Met, Phe or Tyr in this position resulted in substrates with excellent K_m values. The k_{cat} values obtained (Table II) were also comparable to those reported previously for the hydrolysis of a series of chromogenic substrates that were based directly on the sequence found at one of the CA/NC cleavage junctions in the HIV-1 polyprotein [3], i.e. Lys-Ala-Arg-Val-Nle*Nph-Glu-Ala-Nle-NH_2. Thus, the general nature of the proteinase is indicated by this ability to attack effectively substrates with considerably different sequences.

Since tyrosine provides a readily detectable signal (for monitoring by UV absorption in FPLC traces when poorer substrates are being examined), further modifications were carried out on peptides containing -Tyr*Nph- as the -P_1*P_1'- residues contributing the scissile peptide bond. Systematic replacement of the residue in the P_3 position was thus carried out in the series Ala-Thr-His-Xaa-Val-Tyr*Nph-Val-Arg-Lys-Ala.

Table II

Kinetic parameters for hydrolysis of chromogenic peptide substrates of general structure Ala-Thr-His-Gln-Val-Zaa*Nph-Val-Arg-Lys-Ala by HIV-1 proteinase

Peptide number	Zaa =	K_m (μM)	k_{cat} (s^{-1})
1	Arg		0
2	Glu		0
3	Met	15	15
4	Phe	15	13
5	Tyr	15	14

All reactions were carried out at 37°C in 0.1 M sodium acetate buffer, pH 4.7, containing 4 mM EDTA and sufficient NaCl to give a final ionic strength of 1 M. The peptide numbers relate to Table I.

The data obtained (Table III) indicate that there is a considerable tolerance in this position by the enzyme. With the sole exception of a proline residue (peptide 15), all of the other variants investigated (whether acidic, basic, aromatic or hydrophobic) showed k_{cat} values in the range of 2-15 s^{-1} and K_m values of less than 30 μM (Table III). Thus, irrespective of the slight individual variations in kinetic parameters, most of these peptides (peptides 5, 9, 10, 13, 14) are essentially as good substrates as those described by Richards et al. [3] and are better than other peptide substrates described hitherto by other investigators (e.g. [4]).

However, the importance of the ability of the substrate peptide to adopt a productive conformation in the active site cylinder of the enzyme [7] is indicated by the data obtained with the proline-containing peptide (peptide 15; Table III). In a parallel manner, when proline was introduced in place of the His residue in the more peripheral P_4 position (peptide 6), once again the k_{cat} measured for the hydrolysis of this peptide was substantially lowered (albeit to a still measurable value; Table III). However, a further replacement of the P_3 Gln residue with cysteine (peptide 7) restored the k_{cat} value back closer to the original values (Table III). Thus, the flexibility of the overall sequence that can be tolerated by the enzyme is rather remarkable. Even a sequence containing 3 proline residues in the P_5-P_4-P_3 positions (peptide 8) was able to find effectively to the HIV-1 proteinase but the unproductive orientation of the scissile peptide bond engendered by this interaction is reflected in the poor k_{cat} value (Table III). A similar influence of a proline residue, positioned appropriately in a peptide substrate, in lowering k_{cat} has been described previously for hydrolysis by a mammalian aspartic proteinase [8].

Finally, the nature of the residue occupying the P_2 position was investigated and found to be particularly critical (Table IV). To document this, the series Ala-Thr-His-Gln-Yaa-Tyr*Nph-Val-Arg-Lys-Ala was examined. The presence of a β-branched residue (Val, Ile, - peptides 5, 16) was found to be beneficial (Table IV).

Table III

Kinetic parameters for hydrolysis by HIV-1 proteinase of synthetic chromogenic peptides varying in the residues occupying the P_4 and P_3 positions and with the general structure Ala-Thr-P_4-P_3-Val-Tyr*Nph-Val-Arg-Lys-Ala

Peptide number	P_6-P_5-P_4-P_3-sequence	K_m (μM)	k_{cat} (s^{-1})
9	Ala-Thr-His-Tyr-	6	7
10	Ala-Thr-His-Arg-	15	10[b]
11	Ala-Thr-His-Val-	<30[a]	< 3[a]
12	Ala-Thr-His-Asn-	11	2[a]
13	Ala-Thr-His-Asp-	16	10[b]
14	Ala-Thr-His-Glu-	30	15[b]
5	Ala-Thr-His-Gln-	15	14
15	Ala-Thr-His-Pro-		0
6	Ala-Thr-Pro-Gln-	<40[a]	0.3[a]
7	Ala-Thr-Pro-Cys-	10	7
8	Ala-Pro-Pro-Pro-	<50[a]	0.2[a]

Mercaptoethanol (5 mM) was added to some assays but had no influence on the kinetic parameters determined (particularly with peptide 7)

[a] Slow reactions were analysed by FPLC for determination of V_{max} values. The K_m values in these cases are upper estimates since it is difficult to obtain accurate values from the low peak sizes of the products generated at μM concentrations of substrate

[b] Values determined both by spectrophotometric and FPLC analysis. Other details as in the legend to Table II

By contrast, replacement of the Ile residue with an isomeric Leu (peptide 17) had a relatively minor effect on K_m but resulted in a 5-fold lowering of k_{cat}. Insertion of a Phe residue into the P_2 position (peptide 18) diminished k_{cat} (Table IV) to a value sufficiently low that it had to be determined by FPLC (and, consequently an overestimate was obtained for K_m; see legend to Table III). A slightly poorer K_m value was measured for peptide 19, with Ala in the P_2 position, but the k_{cat} value derived was comparable to that observed for leucine (Table IV). Thus, it would appear that the increased size resulting from the addition of a benzene ring in the Phe side chain in P_2 of peptide 18 leads to distortion upon binding that is unfavourable to catalysis. Removal of the side-chain altogether to leave a Gly residue (peptide 20) resulted in a peptide that was essentially resistant to hydrolysis (Table IV).

Thus, a peptide (peptide 5) that is a very acceptable substrate for the HIV-1 proteinase is transformed into one that is refractory to hydrolysis by the presence/absence of three C-atoms, not in a central position contributing to the scissile peptide bond but in the adjacent P_2 flanking position. The importance of the nature of the residue occupying this position has been identified previously with archetypal aspartic proteinases [9] where, for example, the presence of a His residue in P_2 has been shown to be very unfavourable for substrate or inhibitor interaction with pepsin [10].

From these findings, it is clear that inhibitors of the aspartic proteinase from the AIDS virus can be designed which do not mimic exactly the sequence of residues

Table IV

Kinetic parameters for hydrolysis of synthetic chromogenic peptides of general structure Ala-Thr-His-Gln-Yaa-Tyr*Nph-Val-Arg-Lys-Ala by HIV-1 proteinase

Peptide number	Yaa =	K_m (μM)	k_{cat} (s^{-1})
5	Val	15	14
16	Ile	15	11
17	Leu	30	2[a]
18	Phe	<100[a]	0.7[a]
19	Ala	60	2[a]
20	Gly		~ 0[a]

[a] Slow reactions were analysed by FPLC for determination of V_{max} values (see legend to Table III for details). The identity of the sole cleavage site was confirmed by FPLC analysis in each case

surrounding cleavage junctions in the HIV polyprotein. It is acceptable to place two large aromatic/hydrophobic groups in -P_1*P_1'- provided that they are flanked by residues of an appropriate nature. Since it has been shown previously that HIV-proteinase needs 6/7 residues at most for adequate substrate interactions, it would seem that the P_2/P_2' positions are likely to be of paramount importance. While this manuscript was in preparation, a report appeared describing variations in the P_2' position in an HPLC-type substrate [11]. Just as was found for the P_2 position in the present study, it was reported that the presence of a β-branched Ile residue in P_2' resulted in higher k_{cat} values than were measured for either Ala or Leu, with K_m relatively unaffected. The nature/combination of residues occupying these 2 positions is thus of sufficient importance to warrant further investigations into this important enzyme.

Acknowledgements: The contributions of Dr Noel Roberts, Mary Graves and Herve Hirel in generously supplying inhibitors, proteinase, enthusiasm, ideas and encouragement are gratefully acknowledged. This work was supported by awards from the MRC-AIDS Directed Programme, from the NIH (AI28571) and from the Royal Society Joint Scientific Research Programme.

REFERENCES

[1] 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. and Machin, P.J. (1990) Science 248, 358-361.

[2] Kay, J. and Dunn, B.M. (1990) Biochim. Biophys. Acta 1048, 1-17.

[3] Richards, A.D., Phylip, L.H., Farmerie, W.G., Scarborough, P.E., Alvarez, A., Dunn, B.M., Hirel, Ph.-H., Konvalinka, J., Strop, P., Pavlickova, L., Kostka, V. and Kay, J. (1990) J. Biol. Chem. 265, 7733-7736.

[4] Tomaszek, T.A., Magaard, V.W., Bryan, H.G., Moore, M.L. and Meek, T.D. (1990) Biochem. Biophys. Res. Commun. 168, 274-280.

[5] Current Communications in Molecular Biology:Viral Proteinases as Targets for Chemotherapy (1989) (Krausslich, H.-G., Oroszlan, S. and Wimmer, E. eds) pp. 259-267, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[6] Richards, A.D., Roberts, R.F., Dunn, B.M., Graves, M.C. and Kay, J. (1989) FEBS Lett. 247, 113–117.
[7] Miller, M., Sathyanaryana, B.Y., Wlodawer, A., Toth, M.V., Marshall, G.R., Clawson, L., Selk, L., Schneider, J. and Kent, S.B.H. (1989) Science 246, 1149–1152.
[8] Pohl, J. (1990) in: Peptides: Chemistry, Structure and Biology (Rivier, J.E. and Marshall, G.R. eds) pp. 365–366, Escom, Leiden, Netherlands.
[9] Jupp, R.A., Dunn, B.M., Jacobs, J.W., Vlasuk, G., Arcuri, K.E., Veber, D.F., Perlow, D.S., Payne, L.S., Boger, J., De Laszlo, S., Chakravarty, P.K., TenBroeke, J., Hangauer, D.G., Ondeyka, D., Greenlee, W.J. and Kay, J. (1990) Biochem. J. 265, 871–878.
[10] Maibaum, J. and Rich, D.H. (1988) J. Med. Chem. 31, 625–629.
[11] Margolin, N., Heath, W., Osborne, E., Lai, M. and Vlahos, C. (1990) Biochem. Biophys. Res. Commun. 167, 554–560.
[12] Schechter, I. and Berger, A. (1967) Biochem. Biophys. Res. Commun. 27, 157–162.