Volume 260, number 2, 195-197 FEB 08041 January 1990

Selective cleavage of glycyl bonds by papaya proteinase IV

David J. Buttle, Anka Ritonja[+], Laurence H. Pearl[°], Vito Turk[+] and Alan J. Barrett

Biochemistry Department, Strangeways Research Laboratory, Worts Causeway, Cambridge CB1 4RN, England, [+]Department of Biochemistry, J. Stefan Institute, Jamova 39, 61000 Ljubljana, Yugoslavia and *[°]Department of Biochemistry, University College, London WC1E 6BT, England*

Received 28 November 1989

The specificity of papaya proteinase IV (PPIV) has been examined with small substrates and a protein. With both classes of substrate, the enzyme shows a marked selectivity for cleaving glycyl bonds. Boc-Ala-Ala-Gly-NHPhNO_2 is a convenient substrate for routine assays that discriminate well against chymopapain, the most common contaminant of PPIV. Sixteen cleavage points in β-trypsin were identified, of which 13 are glycyl bonds. Tentative suggestions are made as to the reasons for lack of cleavage of some other glycyl bonds. The structure of PPIV has been modelled on that of papain, and we suggest that the replacement of the highly conserved residues Gly-65 and Gly-23 by arginine and glutamic acid, respectively, can account for the specificity of PPIV.

Specificity; Papain

1. INTRODUCTION

PPIV is a cysteine proteinase that is abundant in latex from the unripe fruit of the papaya (*Carica papaya*) plant [1]. It displays a high degree of sequence identity with other proteinases from the same source, and is clearly a member of the papain superfamily of cysteine proteinases [2]. Despite the similarity in primary structure, there have been strong indications that the substrate specificity of PPIV is very different from that of other members of the superfamily. PPIV failed to cleave several amide substrates, and caused only limited hydrolysis of casein [1]. The purpose of the present study was to elucidate the specificity of PPIV with small substrates and a protein.

2. EXPERIMENTAL

2.1. *Materials*

β-Trypsin was prepared according to [3]. The papaya proteinases were purified from papaya latex as previously described: papain [1,2], chymopapain [4] (with modification to complete the removal of PPIV), PPIII [4] and PPIV [1,2]. The cysteine proteinases were standardized by active-site titration as described [5], and for active-site titration of PPIV, this method was adapted for the use of Boc-Ala-Ala-Gly-NHPhNO_2 as substrate (from Bachem Feinchemikalien AG, Bubendorf, Switzerland). Boc-Ala-Ala-Gly-NHMec and Boc-Ala-Ala-Ala-NHMec were synthesized by Dr C.G. Knight of the Strangeways Laboratory. The chemicals used for Edman degradation were of sequenal grade from applied Biosystems.

2.2. *Determination of kinetic parameters*

Continuous rate assays in 100 mM sodium phosphate, 1 mM EDTA, 0.005% Brij 35, 2 mM cysteine, pH 6.8, were used to determine K_m and k_{cat} values. For the analysis of the hydrolysis of Boc-Ala-Ala-Gly-NHPhNO_2, experiments were done at 20°C in a recording spectrophotometer. Substrate concentrations ranged from 0.10 to 4.35 mM. Hydrolysis of Boc-Ala-Ala-Gly-NHMec and Boc-Ala-Ala-Ala-NHMec was followed in a Perkin-Elmer fluorimeter (excitation 360 nm, emission 460 nm) at 30°C with the aid of a personal computer running the FLUSYS software [6]. Substrate concentrations ranged from 5 to 80 μM for the glycine substrate, and 16 to 130 μM for the alanine analogue, the upper limit being determined in both cases by solubility. In all cases, K_m and k_{cat} values were found by non-linear regression analysis (Enzfitter, Elsevier-Biosoft, Cambridge, England).

2.3. *Hydrolysis of bovine β-trypsin with papaya proteinase IV*

Trypsin was reduced with 2-mercaptoethanol in 6 M guanidine HCl at 40°C for 10 h and alkylated with iodoacetic acid. The salts were removed by dialysis in 0.1 M potassium phosphate buffer, pH 6.5, containing 1 mM EDTA and 2 mM dithiothreitol. The carboxymethylated trypsin (3.0 mg in 1.0 ml) was not completely soluble at the start of the digestion, when 2% (w/w) PPIV was added. The mixture was allowed to stand at room temperature (about 22°C) and samples were taken after 10, 30, 60, 120 and 240 min for analysis. The enzymatic reaction was stopped by adding formic acid to a final concentration of 20% (v/v). Peptides were purified by HPLC on a Bio-Rad C_{18} column eluted with aqueous acetonitrile containing trifluoroacetic acid. Individual peptides were sequenced with an Applied Biosystems gas phase sequenator model 470 A [7]. Phenylthiohydantoin derivatives were identified on line with the 120A HPLC [8].

Correspondence address: D.J. Buttle, Biochemistry Department, Strangeways Research Laboratory, Worts Causeway, Cambridge CB1 4RN, England

Abbreviations: Boc, *t*-butyloxycarbonyl; Bz, benzoyl; compound E-64, L-3-carboxy-2,3-*trans*-epoxypropionyl-leucylamido(4-guanidino)butane; NHPhNO_2, *p*-nitroanilide; NHMec, 7-(4-methyl)coumarylamide; OPhNO_2, *p*-nitrophenyl ester; PPIII, papaya proteinase III; PPIV, papaya proteinase IV; Z, benzyloxycarbonyl

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50 © 1990 Federation of European Biochemical Societies

3. RESULTS AND DISCUSSION

3.1. *Action on small substrates*

The earlier finding that PPIV rapidly cleaved Boc-Gly-OPhNO$_2$ [1], was followed up by the testing of analogous amide substrates which are potentially more suitable for routine use. K_m and k_{cat} values were determined for three substrates with all four known papaya cysteine proteinases. The results (table 1) show that Boc-Ala-Ala-Gly-NHPhNO$_2$ would discriminate in favor of the detection of PPIV in mixtures of these proteinases, largely through differences in k_{cat}. The substrate is significantly sensitive to papain, also, but its low sensitivity to chymopapain is of practical importance, because chymopapain can be a particularly troublesome contaminant in the isolation of PPIV [1]. Used in conjunction with Bz-Arg-NHPhNO$_2$, which is cleaved by the other papaya proteinases but not PPIV [1], Boc-Ala-Ala-Gly-NHPhNO$_2$ readily distinguishes PPIV from the other enzymes in fractions from chromatographic columns.

A fluorimetric analogue of the nitroanilide, Boc-Ala-Ala-Gly-NHMec, was also cleaved most effectively by PPIV, the major difference between the papaya proteinases again being in k_{cat}. PPIV does not cleave peptide-NHMec substrates that are used for the other papaya proteinases such as Z-Phe-Arg-NHMec [1,5], so that the use of two fluorimetric substrates can show the proportions of individual proteinases in mixtures. In addition, the sensitivity of Boc-Ala-Ala-Gly-NHMec allows the detection of PPIV at the picomolar concentrations encountered in work with tight-binding inhibitors.

Substitution of Ala for Gly in the NHMec substrate at position P1 (in the terminology of [9]) greatly decreased the value of k_{cat}, causing a 30-fold reduction in the specificity constant, k_{cat}/K_m.

3.2. *Action on β-trypsin*

Analysis of the samples from the digestion mixture by HPLC showed that virtually no uncleaved trypsin remained at 30 min. The enzyme hydrolyzed the insoluble fraction of the carboxymethylated trypsin, and initially

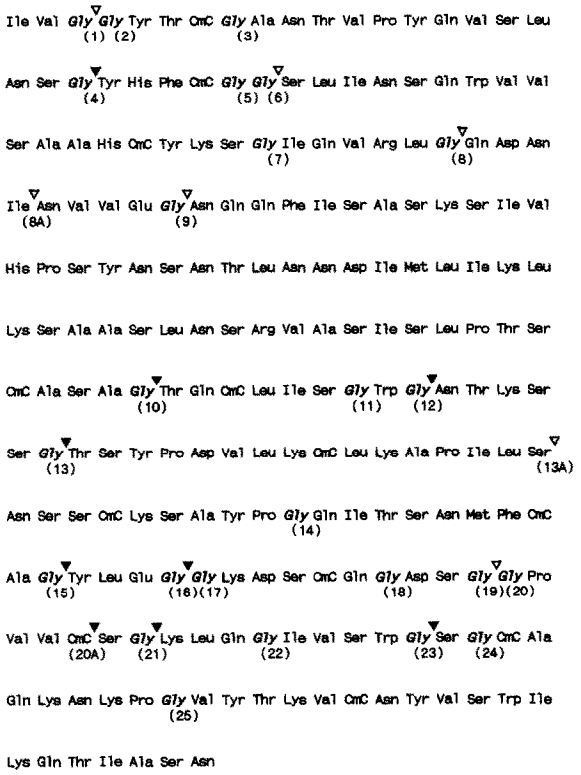

Fig.1. Points of cleavage of reduced, carboxymethylated β-trypsin by PPIV. Glycyl bonds, and others that were cleaved, are numbered for reference. CmC, carboxymethyl-cysteine; ▼, cleavage detected after 10 min; ▽, cleavage detected only after 30 min (see text for details).

produced long peptides that were then further hydrolyzed as the time of incubation increased. Peptides were isolated from the 10-min and 30-min digests for identification (fig.1). Thirteen of the sixteen cleavages were at glycyl bonds, the remainder having isoleucine, serine and cysteic acid in P1. Four of the glycyl bonds not cleaved (bonds 2, 5, 17, 20) were adjacent to others that were, so that the glycine residues would have become terminal or sub-terminal in the digest, and thus resistant to endopeptidase action. Some of the uncleaved bonds contained proline in P2 or P1′ (bonds 14, 20, 25) or carboxymethyl-cysteine in these positions (bonds 3, 5, 24), which were not seen in any of the bonds cleaved,

Table 1

Kinetic constants for the hydrolysis of substrates of PPIV by the papaya proteinases

Enzyme	Substrate								
	Boc-Ala-Ala-Gly-NHPhNO$_2$			Boc-Ala-Ala-Gly-NHMec			Boc-Ala-Ala-Ala-NHMec		
	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)
Papain	1.4	1.6	1.14	0.14	2.0	14.3	ND	ND	ND
Chymopapain	4.2	0.5	0.12	0.23	0.4	1.7	ND	ND	ND
PPIII	4.2	0.5	0.12	0.17	0.1	0.6	ND	ND	ND
PPIV	5.2	22.0	4.23	0.16	5.0	31.2	0.08	0.08	1.00

Hydrolysis of the nitroanilide substrate was at 20°C, and of the coumarylamide substrates at 30°C, as described in the text. The SEM for all values given was less than 10% of the mean. ND, not determined

and by analogy with the specificity of some other endopeptidase, may be excluded by the specificity of PPIV.

In view of its narrow specificity, PPIV shows promise of being a useful reagent for limited proteolysis in the laboratory, in amino acid sequencing and dissection of proteins for other purposes.

3.3. *Molecular modelling of PPIV*

In an attempt to account for the distinctive specificity of PPIV, we have constructed a three-dimensional model on the basis of the amino acid sequence and the known structure of papain (Brookhaven Protein Structure Data Bank, Entry PIPPD). Since PPIV is 67% identical with papain in amino acid sequence [2], the polypeptide backbone structure can be expected to be very similar, as is that of actinidin with only 48% sequence identity to papain [10]. The model shows that the structure of the S1 substrate binding site of PPIV is greatly altered by the presence of glutamic acid and arginine, respectively, at positions 23 and 65, which are both glycine residues in all other cysteine proteinases for which sequences are available. These residues are located in one wall of the active site cleft, close to the catalytic cysteine 25, and their side chains form a barrier across the upper part of the cleft effectively excluding residues with long side chains from the S1 binding site. Although some flexibility may be present, the glutamate/arginine barrier would be substantially rigidified by hydrophobic packing of the methylene groups and strong charge-assisted hydrogen bonding between the guanidino and carboxyl head groups of their respective side chains. The presence of the barrier would greatly hinder the binding of substrates with side chains that extend past Cβ. Substrates with glycine at P1 would bind with no difficulty, and although small side chains such as those of alanine or serine may be accommodated, this would occur at the expense of a decrease in the stability of the transition state relative to a substrate binding with no strain. The barrier may also prevent binding in a conformation optimal for attack by the reactive thiol of Cys-25, so that such substrates would be hydrolyzed with a lower catalytic efficiency. The restricted specificity of PPIV may therefore result from both a sterically restricted binding site and poorer catalysis for substrates with P1 side chains.

REFERENCES

[1] Buttle, D.J., Kembhavi, A.A., Sharp, S.L., Schute, R.E., Rich, D.H. and Barrett, A.J. (1989) Biochem. J. 261, 469–476.
[2] Ritonja, A., Buttle, D.J., Rawlings, N.D., Turk, V. and Barrett, A.J., FEBS Lett., in press.
[3] Strop, P. and Cechova, D. (1981) J. Chromatogr. 207, 55–62.
[4] Buttle, D.J. and Barrett, A.J. (1984) Biochem. J. 223, 81–88.
[5] Zucker, S., Buttle, D.J., Nicklin, M.J.H. and Barrett, A.J. (1985) Biochim. Biophys. Acta 828, 196–204.
[6] Rawlings, N.D. and Barrett, A.J., Comput. Appl. Biosci., in press.
[7] Hewick, R.M., Hunkapiller, M.W., Hood, L.E. and Dreyer, W.J. (1981) J. Biol. Chem. 256, 7990–7997.
[8] Hunkapiller, M.W. and Hood, L.E. (1983) Methods Enzymol. 91, 486–493.
[9] Berger, A. and Schechter, I. (1970) Phil. Trans. Roy. Soc. Lond. Ser. B 257, 249–264.
[10] Kamphuis, I.G., Drenth, J. and Baker, E.N. (1985) J. Mol. Biol. 182, 317–329.