Volume 328, number 1,2, 107–110 FEBS 12812 August 1993

The specificity of the S_1' subsite of cysteine proteases*

Robert Ménard, Euridice Carmona**, Céline Plouffe, Dieter Brömme, Yasuo Konishi, Jean Lefebvre and Andrew C. Storer

Biotechnology Research Institute, National Research Council of Canada, 6100 Avenue Royalmount, Montréal, Québec, H4P 2R2 Canada

Received 8 June 1993

The specificity of the S_1' subsite of the cysteine proteases cathepsin B, L, S and papain has been investigated using a series of intramolecularly quenched fluorogenic substrates (Dansyl-Phe-Arg-AA-Trp-Ala) where the P_1' amino acid (AA) has been varied. Taken individually, each enzyme displays a relatively broad S_1' subsite specificity and this subsite cannot be considered as a primary site of specificity. Notable differences do exist however between the various proteases. Cathepsin B prefers large hydrophobic residues in the P_1' position of a substrate while cathepsin L has an opposite trend, favoring amino acids with small (Ala, Ser) or long but non-branched (Asn, Gln, Lys) side chains. Cathepsin S and papain display a somewhat broader S_1' subsite specificity.

Cysteine protease; Papain; Cathepsin; Subsite specificity

1. INTRODUCTION

Knowledge of the specificity of cysteine proteases provides valuable information that can lead to a better understanding of the physiological roles of these enzymes or to the design of selective inhibitors. Early studies on the substrate specificity of cysteine proteases centered mainly on the plant enzyme papain. Schechter and Berger demonstrated that substrates could interact with seven subsites in papain and that the primary determinant of specificity was the nature of the P_2 residue [1,2]. This has been shown to be the case also with other cysteine proteases of the papain family, in particular the mammalian enzymes cathepsin B and L [3]. The specificity of the S subsites has been studied in greater detail due to the availability of small synthetic ester and amide substrates containing a chromophoric group attached to the C-terminus of the peptide and released upon hydrolysis. Very little is known about the S′ subsite specificities of cysteine proteases. Studies addressing this aspect have been carried out using partitioning experiments where an acyl enzyme intermediate reacts with different amino acid-derived nucleophiles [4]. However, the mechanism of nucleophilic attack on acyl enzymes can be quite complex and the interpretation of such experiments is not straightforward [5]. Another approach consists of using intramolecularly quenched fluorogenic substrates [6]. The main feature of such a substrate is the presence, in addition to a fluorescent donor, of an acceptor group that acts as an internal quencher when the peptidyl substrate is intact. Compounds of this type have been used recently as substrates for papain [7,8]. A few amino acid substitutions at positions P_1' and P_2' demonstrated the potential of the approach for the characterization of the S′ subsites specificities of cysteine proteases as well as other classes of proteolytic enzymes. The cathepsins have drawn interest due to their possible involvement in various physiological and pathophysiological processes [9]. Designing selective inhibitors to a given enzyme is a major task that could be aided by specific features of the S′ subsites of these enzymes. For example, the selectivity of a dipeptidyl epoxysuccinyl analog of E-64 for cathepsin B is a direct consequence of inhibitor interaction in the S′ subsites of cathepsin B [10,11]. In order to look for differences in specificities that could be exploited in a similar manner, the S_1' specificity of cathepsin B, S and L, as well as that of the model plant protease papain, has been determined against a series of intramolecularly quenched fluorogenic substrates.

2. MATERIALS AND METHODS

2.1. *Materials*

Papain was obtained from Sigma Chemical Co and was purified and activated as described previously [12]. Rat cathepsin B, kindly provided by Dr. John S. Mort (Shriners Hospital for Crippled Children, Montreal, Canada) and human cathepsin S were recombinant proteins obtained from yeast and purified as described previously

Correspondence address: R. Ménard, Biotechnology Research Institute, National Research Council of Canada, 6100 Avenue Royalmount, Montréal, Québec, Canada H4P 2R2. Fax: (1) (514) 496 5143.

**NRCC Publication No 36150*

***Dr. Euridice Carmona is a visiting scientist from Instituto Butantan, C.P. 65, 05504, São Paulo, Brazil.*

[13–15]. Rat cathepsin L was a generous gift from Dr. H. Kirschke (Martin-Luther-University, Halle (Saale), Germany). All peptidyl substrates were synthesized by the solid-phase method on an Applied Biosystem 430A peptide synthesizer, as described previously [16]. The dansyl group was incorporated by reacting dansyl chloride with the peptide in presence of *N*-methyl morpholine for 30 min, after deprotection of the N-terminal amino acid.

2.2 *Methods*

Cleavage sites of the substrates were identified by reverse-phase HPLC on a Vydac C-18 column (25 × 0.46 cm) using a linear gradient (10% CH_3CN to 50% CH_3CN in 40 min, all in 0.1% TFA) at a flow rate of 1 ml/min The retention time of the intact substrates were between 24 and 33 min. Substrate hydrolysis resulted in the appearance of two new peaks The fragment Dns-Phe-Arg was easily detectable since it had a typical retention time of ca. 16 min and was present in all enzyme-substrate mixtures. Retention time of the second hydrolysis product varied according to the nature of the peptide For cathepsin B, a second cleavage site was evidenced by the presence of additional peaks in the chromatogram.

The reactions were followed by monitoring the decrease in fluorescence of the dansyl group upon hydrolysis of the P_1-P_1' bond. This was done by setting the excitation monochromator at the wavelength for tryptophan absorption (290 nm) and the emission monochromator at the wavelength for dansyl emission (550 nm). The reaction conditions consisted of: 50 mM phosphate buffer, 0.2 M NaCl, 5 mM EDTA, 5% CH_3CN, and 1–2 mM DTT. The assays were done at pH 6.5 for papain and cathepsin S, 6 0 for cathepsin B and 5.8 for cathepsin L. All kinetic experiments were performed at 25°C. The experiments were carried out at low substrate concentrations where the reactions followed first-order kinetics with a rate constant k_{obs}, determined by non-linear regression of the progress curve to a single exponential equation. The value of k_{obs} was found to be independent of the substrate concentrations used indicating that $[S] \ll K_m$ and k_{cat}/K_M can be obtained simply by dividing k_{obs} by the enzyme concentration. For reactions where hydrolysis at a second site was significant, it can be shown that the measured k_{cat}/K_M corresponds to the sum of the individual values of k_{cat}/K_M for each cleavage site and that the ratio of the resulting products from each site corresponds to the ratio of the individual k_{cat}/K_M values [4,17]. This ratio was determined from HPLC experiments and the measured k_{cat}/K_M was corrected to reflect solely the specificity constant for hydrolysis at the P_1-P_1' position

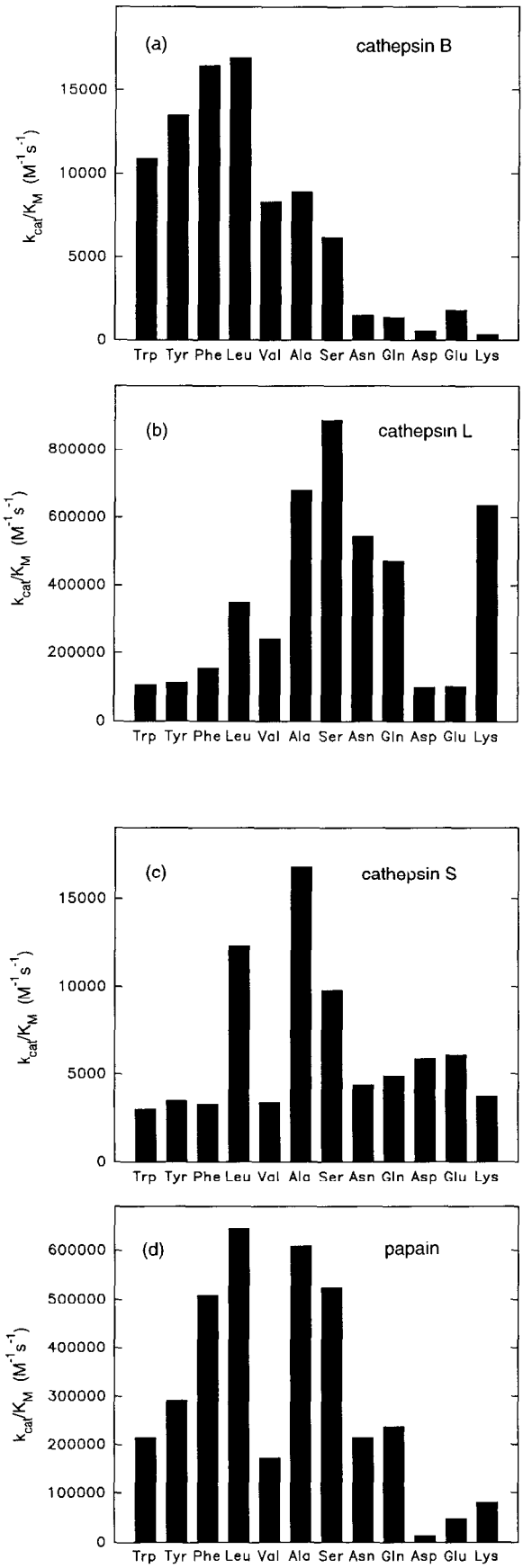
3. RESULTS AND DISCUSSION

The substrates used to investigate the S_1' specificity are Dns-Phe-Arg-AA-Trp-Ala, where AA = Trp, Tyr, Phe, Leu, Val, Ala, Ser, Asn, Gln, Asp, Glu or Lys. These substrates were designed for a number of reasons. They contain a dansyl group on the N-terminal portion of the peptide and a Trp residue on the C-terminal side of the scissile bond which is expected to be between residues Arg (P_1) and AA (P_1'). The presence of the dansyl group, which displays an absorption band in the 330 nm region, will quench the fluorescence of the Trp residue (λ emission ca. 340 nm) in the intact peptide. This transfer of excitation energy also causes the dansyl group to emit fluorescence at 550 nm. Upon hydrolysis of the P_1-P_1' bond, the quenching of Trp fluorescence is relieved resulting in an increase of the fluorescence at 340 nm as well as a decrease of the dansyl fluorescence at 550 nm. In this study, the fluorescence of the dansyl group resulting from the energy transfer from the Trp residue in the intact peptide was used to follow the reactions in order to avoid background fluorescence from the Trp residues present in the cysteine proteases.

Dns-Phe-Arg-**AA**-Trp-Ala
P_2 P_1 P_1' P_2' P_3'

Other features of these substrates ensure that the hydrolysis will occur between residues Arg and AA. Phe was placed at position P_2 since it is a residue either highly favored or accepted in the S_2 subsites of the cysteine proteases under investigation. Arg was chosen as the P_1 residue since it is well accepted in the S_1 subsites of the cysteine proteases but is a highly unfavorable residue for the S_2 subsite of most enzymes. In addition, a free Ala was placed at position P_3' of the peptide. Since cathepsin S and L as well as papain do not possess carboxypeptidase activity, the presence of the negative charge avoids placing the Ala residue in subsites S_1' or S_2'. Reactions with cathepsin B deserve special attention for two reasons. Contrary to the other enzymes used in this study, cathepsin B has a dipeptidyl carboxypeptidase activity [18] and is also able to accept Arg in its S_2 subsite [3]. Depending on the nature of residue AA, hydrolysis could therefore occur at the AA- Trp bond with cathepsin B. For this reason, and to verify that the reaction was occurring at the predicted site with all cysteine proteases, the hydrolysis fragments were analysed by HPLC. Single site cleavage at the predicted location was observed for all enzyme–substrate systems except for cathepsin B where cleavage at a second site (probably the AA-Trp bond) was observed with AA = Trp, Tyr, Phe, Ala, Ser, Asn or Gln. However, the major site of hydrolysis was still the Arg-AA bond. With AA = Lys, hydrolysis occurred mainly at the AA-Trp bond. The kinetic parameter k_{cat}/K_M was corrected to take into account the reaction at a second site and reflects only the rate constant for hydrolysis at the Arg-AA bond.

The specificity constants k_{cat}/K_M for hydrolysis of the substrates Dns-Phe-Arg-AA-Trp-Ala by the four enzymes under investigation are illustrated in Fig. 1. The standard deviation on k_{cat}/K_M values was usually 5–10% and never greater than 25% except for reactions of cathepsin B with the substrates where AA = Asp and Glu (standard deviations of 30% and 40%, respectively), where the k_{cat}/K_M values are very low. By considering each enzyme individually, it can be seen that the S_1' specificity of these enzymes is relatively broad. This is particularly true of cathepsin S where there is only 6-fold difference in k_{cat}/K_M between the best and worst substrates. For cathepsin B and papain, an acid residue in P_1' is particularly unfavorable and the substrate with Asp in P_1' has a k_{cat}/K_M value that is 52-fold and 48-fold lower than that of the best substrate for cathepsin B and papain, respectively. These variations in the specificity constant are much less than what is observed with the P_2 residue so that the S_1' subsite cannot be considered



as a primary determinant of specificity for these cysteine proteases.

When comparing the specificity profiles obtained for the various proteases (Fig. 1a–d), clear differences in the general specificity of each enzyme can be noted particularly between cathepsin B and L. Cathepsin B displays a preference for large hydrophobic residues in the P_1' position of a substrate (Trp, Tyr, Phe, Leu) over residues such as Asn and Gln. Cathepsin L has an opposite trend favoring substrates where the P_1' amino acid side chain is small (Ala, Ser) or long but non-branched (Asn, Gln, Lys). It must be noted also that cathepsin L is the only enzyme in our study that displays a high level of activity against the substrate with Lys in P_1' (k_{cat}/K_M = 638,000 $M^{-1}\cdot s^{-1}$). The specificity is somewhat broader with papain and cathepsin S and a distinctive trend in the general specificity profile cannot easily be established. Based on the nucleophile partitioning experiments of Alecio et al. [19], the S_1' subsite of papain has been considered to show a preference for hydrophobic residues, especially Leu. This is only partly supported by our results since the substrate with AA = Leu is only slightly better than those with AA = Ala or Ser. Our work also is not in agreement with the finding that Asn is a much poorer residue than Leu for the S_1' subsite of papain. In the present study, the S_1' specificity of papain has been investigated in a direct manner by measuring the specificity constant k_{cat}/K_M for hydrolysis of substrates covering a wide range of residues in the P_1 position. Clearly, the most significant comment that can be made from the data presented in Fig. 1d is that the specificity of the S_1' subsite of papain is relatively broad.

One must keep in mind that other types of substrates could yield slightly different results regarding S_1' specificities due for example to synergistic effects of binding at different subsites. Our results reflect the S_1' specificity of cysteine proteases solely for substrates with a Trp residue in the P_2' position. Clearly, the nature of the amino acids in positions P_2, P_1 and P_3' could also have an influence on the S_1' specificity of the enzymes. This study allows however a qualitative evaluation of the nature of S_1' subsite interactions in cysteine proteases. From our results, we can state that the S_1' subsite of cysteine proteases is not a primary site of specificity when compared to the S_2 subsite. Notable differences in trend were observed between the various enzymes used in this study which can be interpreted to reflect differences in size and hydrophobicity of their S_1' subsites. It has been shown in the past that the backbone hydrogen

←

Fig. 1. Specificity constant (k_{cat}/K_M) for the hydrolysis of Dns-Phe-Arg-AA-Trp-Ala by cysteine proteases. (a) cathepsin B, (b) cathepsin L, (c) cathepsin S, (d) papain. The experiments were repeated four to ten times and the average value is reported. The k_{cat}/K_M for reactions of cathepsin B have been corrected to take into account the competing reaction at a second site, as described in section 2.

bond donors and acceptors of substrates (i.e. NH and C = O) interact with the S′ subsites of papain and that these interactions contribute significantly to binding and catalysis [17]. In that aspect, side-chain contacts in the S_1' subsite seem to be less important and cysteine proteases can accept a broad range of amino acids in the P_1' position of a substrate.

Acknowledgements: One of the authors (E.C.) is grateful to the Conselho Nacional de Desenvolvimento Cientifico e Tecnologico (Brazil) for financial support via a CNPq/postdoctoral fellowship

REFERENCES

[1] Schechter, I. and Berger, A. (1967) Biochem. Biophys. Res. Commun. 27, 157–162.
[2] Berger, A. and Schechter, I (1970) Phil Trans Roy. soc Lond B257, 249–264.
[3] Barrett, A J. and Kirschke, H (1981) Methods Enzymol. 80, 535–561.
[4] Fersht, A. (1985) Enzyme Structure and Mechanism, W.H Freeman and Co., New York.
[5] Fink, L and Bender, M.L. (1969) Biochemistry 8, 5109–5118
[6] Yaron, A., Carmel, A. and Katchalski-Katzir, E. (1979) Anal. Biochem. 95, 228–235.
[7] Garcia-Echeverria, C and Rich, D H. (1992) FEBS Lett. 297, 100–102.
[8] Garcia-Echeverria, C. and Rich, D.H (1992) Biochem. Biophys. Res. Commun. 187, 615–619
[9] Rich, D.H. (1986) in: Proteinase Inhibitors (A J. Barrett and G Salvesen, Eds.) pp 153–178, Elsevier, Amsterdam.
[10] Buttle, D.J , Murata, M , Knight, C G and Barrett, A.J. (1992) Arch. Biochem. Biophys. 299, 377–380.
[11] Gour-Salin, B.J , Lachance, P , Plouffe, C , Storer, A.C and Ménard, R. (1993) J. Med. Chem 36, 720–725
[12] Ménard, R., Khouri, H E., Plouffe, C., Dupras, R., Ripoll, D , Vernet, T., Tessier, D C , Laliberté, F., Thomas, D Y and Storer, A.C. (1990) Biochemistry 29, 6706–6713.
[13] Fox, T., de Miguel, E., Mort, J.S and Storer, A.C. (1992) Biochemistry 31, 12571–12576
[14] Rowan, A D., Mason, P., Mach, L. and Mort, J.S. (1992) J. Biol. Chem. 267, 15993–15999.
[15] Brömme, D., Bonneau, P R., Lachance, P., Wiederanders, B., Kirschke, H , Peters, C , Thomas, D Y , Storer, A.C. and Vernet, T. (1993) J. Biol. Chem. 268, 4832–4838.
[16] Szewczuk, Z., Gibbs, B.F., Yue, S Y., Purisima, E.O and Konishi, Y. (1992) Biochemistry 31, 9132–9140.
[17] Berti, P.J , Faerman, C H and Storer, A C (1991) Biochemistry 30, 1394–1402.
[18] Aronson, N.N and Barrett, A.J. (1978) Biochem. J. 171, 759–765.
[19] Alecio, M.R., Dann, M.L and Lowe, G. (1974) Biochem. J. 141, 495–501.