Pergamon

Bioorganic & Medicinal Chemistry Letters 12 (2002) 3619–3623

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Scanning the Prime-Site Substrate Specificity of Proteolytic Enzymes: A Novel Assay Based on Ligand-Enhanced Lanthanide Ion Fluorescence

Amy M. Barrios and Charles S. Craik*

Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143, USA

Received 17 June 2002; accepted 3 September 2002

Abstract—A novel method for assaying the substrate specificity of proteolytic enzymes has been developed utilizing ligand-enhanced lanthanide ion fluorescence. This approach was used to develop peptide libraries to probe substrate specificity in the prime sites of proteolytic enzymes. A positional scanning synthetic combinatorial library of fluorogenic peptides was synthesized and used to determine the extended prime site specificity of bovine α-chymotrypsin. The enzyme showed a preference for Lys and Arg in the P1′ position, rather broad specificity in the P2′ position, and a slight Arg specificity in the P3′ position. The specificity profile of bovine α-chymotrypsin agrees well with previously reported data, and the substrate library reported herein should provide valuable information about the prime site substrate specificities of other proteolytic enzymes as well. Furthermore, the continuous fluorogenic assay described may prove useful in analyzing the activity of other hydrolytic enzymes.

Proteolytic enzymes are ubiquitous in biology, comprising an estimated 2% of the human genome.[1] While well-regulated proteolysis is crucial to the development and survival of many organisms, unregulated proteolysis results in a number of pathogenic states including thrombotic disorders, hypertension, osteoarthritis, chronic degenerative disorders, and cancer.[1,2] One of the challenges facing those who study proteolysis is the task of elucidating the natural substrates of a protease and determining how this function is altered in the disease state.[3]

Knowledge of the primary sequence specificity of a protease provides a first approximation in determining its function in vivo and a number of researchers have developed substrate libraries for this purpose. Of these, the most widely applicable has been the use of coumarin-based fluorogenic substrate libraries to scan the substrate binding pockets on the N-terminal side of the scissile bond (the non-prime[4] subsites).[5–8] Synthetic combinatorial peptide libraries that systematically scan through the P1, P2, P3, and P4 subsites have been developed in our laboratory and used to profile the specificity of serine, cysteine and aspartyl proteases.[5,6,9–11] In addition to providing insight into the biochemistry of a protease, knowledge of substrate specificity can aid in the design of selective protease inhibitors. Simple chemical substrate libraries can yield a wealth of biochemically and therapeutically relevant data.

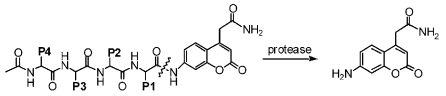
Although the existing libraries have provided valuable information about the specificity of numerous proteases, one of their main shortcomings is the fact that they only profile the non-prime subsites.[3] Phage display of peptide substrates,[12] FRET-based polypeptide substrates,[13,14] and acyl transfer from protease to substrate[15] have been used to map the prime site specificity of a limited number of proteolytic enzymes. Unfortunately, each of these methods has significant drawbacks such as poor solubility, restricted diversity and limited applicability. A simple chemical library that can be used to profile the prime site specificity of a protease is desirable. Ideally, a system could be designed in which (1) an easily detectable, highly sensitive signal (e.g., fluorescence) is activated upon substrate hydrolysis, (2) cleavage of many substrates can be observed simultaneously, and (3) the results can be rapidly deconvoluted to produce a specificity profile. In this manuscript we report the development of a novel method for determining protease substrate specificity C-terminal

*Corresponding author. Tel.: +1-415-476-8146; fax: +1-415-502-8298; e-mail: craik@cgl.ucsf.edu

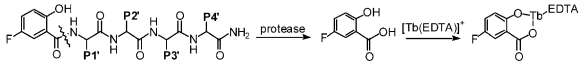
0960-894X/02/$ - see front matter
PII: S0960-894X(02)00786-2

to the scissile bond. The resulting libraries are readily synthesized from commercially available starting materials and should prove useful in profiling the specificity of a variety of proteases.

In order to develop a fluorogenic peptide library that can be used to probe the prime site binding pockets of a protease, an appropriate fluorophore must be chosen. In the case of the non-prime side coumarin-based fluorogenic libraries, the fluorescence of the coumarin moiety is quenched by attachment through an amide bond to the amino acid chain. Upon cleavage of the coumarin-peptide amide bond, the amino group of the coumarin is released and fluorescence is greatly enhanced, as shown in eq 1. A complementary approach is needed in the development of a fluorophore for use in prime site libraries, wherein cleavage of the peptide-fluorophore bond releases a fluorophore with a free carboxylic acid moiety. For this study, we have devised an assay in which a 5-fluorosalicylic acid (fsa) peptide conjugate is hydrolyzed and the resulting free fsa ligand coordinates to a terbium ion in the reaction solution to produce a fluorescent signal (eq 2). The intrinsic fluorescence of the lanthanide ion terbium(III) is very weak, but can be enhanced by the use of laser excitation or a sensitizing ligand such as fsa.[16] Because non-radiative decay of the lanthanide ion excited state can occur through vibronic coupling to water molecules bound to the metal ion, an EDTA ligand or another similar shielding ligand is required to exclude water from the metal coordination sphere.[17]



(1)



(2)

Fsa-enhanced terbium ion fluorescence has been reported as a discontinuous assay for alkaline phosphatase activity[18,19] wherein the reaction mixture was made basic with NaOH before fluorescence was measured. Our initial investigations into the fluorescence of the [Tb(EDTA)(fsa)] conjugate provided confirmation that this technique could also be valuable in the detection of peptide substrate hydrolysis. The sensitivity of the [Tb(EDTA)(fsa)] assay was measured at both elevated pH and pH 8.0. The fsa detection limit was approximately 1 μM in both cases. The slight increase in fluorescence intensity at elevated pH was not enough to justify the use of a discontinuous assay in which the reaction is quenched by addition of NaOH before measuring fluorescence. Instead we chose to use a continuous assay in which the increasing fluorescence is monitored in real time as substrate is hydrolyzed and fsa is released into the $[Tb(EDTA)]^+$ containing reaction solution. At pH 8.0, maximum sensitivity was obtained with a $[Tb(EDTA)]^+$ concentration of 10 μM.

The fsa ligand was incorporated into a tetrapeptide library utilizing the positional scanning approach.[20] In this library, each substrate is a tetrapeptide with an amidated C-terminus and an fsa ligand attached to the N-terminus. Three sublibraries scanning the P1′, P2′, and P3′ positions were synthesized in 96-well plate format, as shown in Figure 1. The P4′ position consists of an equimolar mixture[21,22] of 19 amino acids—the 20 naturally occurring amino acids except methionine and cysteine but including the unnatural amino acid norleucine, a methionine isostere, were used in all cases. For example, in the P2′ sublibrary, the P2′ position is spatially addressed, with each of its 19 members contained in a separate well of the library plate.[23] The P1′, P3′, and P4′ positions are occupied by an equimolar mixture[21,22] of the 19 possible amino acids. By incorporating all 19 of the relevant amino acids at each of these positions, every possible tetrapeptide sequence is represented, facilitating a sampling of all sequence space (Fig. 1).

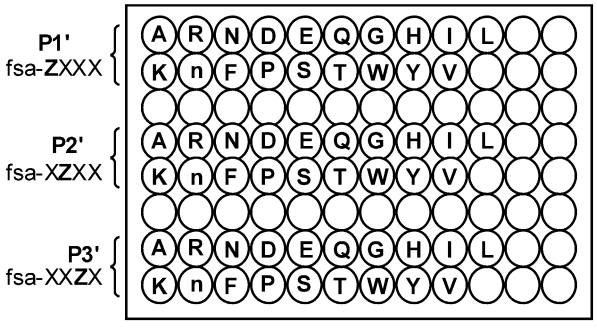
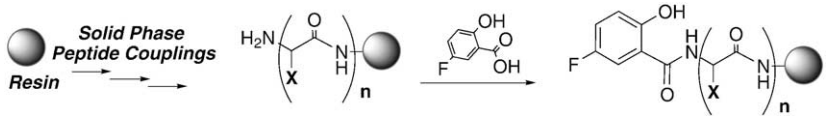


Figure 1. A representation of the substrate library setup. In each sublibrary, X represents an equimolar mixture of all 19 amino acids and Z represents the spatially addressed position.

The positionally scanned library was synthesized using standard solid phase amino acid coupling procedures as illustrated in Scheme 1. Rink amide resin was used as the solid support. The tetrapeptide chain was built upon the resin and the fsa was incorporated at the N-terminus using standard coupling protocols.[5] The library was then cleaved from the resin under acidic conditions. Following the removal of solvent from the library, the residue was dissolved in DMSO to a final concentration of approximately 20 mM total substrate, estimated based on the approximately 40% yields obtained from similar single substrate syntheses. As each sublibrary has 3 randomized positions, there are 19^3, or 6859 different substrates in each library well and the concentration of each individual substrate approximately 2.9 μM. It is important to keep the concentration of substrate below the K_m so that the rate of hydrolysis remains directly proportional to the specificity constant, k_{cat}/K_m.

The combinatorial tetrapeptide library thus obtained was used to assay the substrate specificity of bovine α-chymotrypsin as follows. A 1 μL aliquot of each library member in DMSO was diluted to 100 μL in one well of a black polystyrene 96-well plate with a solution containing 50 mM Tris·HCl pH 8.0, 100 mM NaCl, 20 mM $CaCl_2$, 10 μM $[Tb(EDTA)]^+$ and 10 μM enzyme. The resulting mixture was allowed to react at 25 °C for



Scheme 1.

30 min. Hydrolysis was followed by observing the increase in fluorescence at 546 nm upon excitation at 250 nm. As shown in Figure 2, the enzyme cleaves substrates with Arg and Lys in the P1′ position most efficiently. The P2′ position is more broad in its specificity and the P3′ position shows a slight preference for Arg.

The tetrapeptide library provides an averaged and independent view of the amino acid preferences at each subsite and does not address the possibility of cooperative effects between the subsites. To elucidate any cooperativity between subsites, a series of single substrates was synthesized and assayed with chymotrypsin. The individual substrates were synthesized in a similar fashion as the library and purified by reverse-phase HPLC. Peptide sequences were designed to compare the most reactive amino acid in each sublibrary (Arg) with an amino acid with moderate reactivity and neutral functionality (Ala). Alanine, a small, neutral amino acid, was included in the P4′ position in all peptides, providing extended backbone interactions to enhance turnover. The kinetic data obtained from single substrate hydrolysis were analyzed using the standard Michaelis–Menten equation. As outlined in Table 1, an arginine residue in the P1′ position is necessary for efficient turnover of the substrate. Of the single substrates listed in Table 1, the only substrate without a P1′ Arg residue for which satisfactory kinetic data could be obtained was fsa-ARRA, which had a k_{cat}/K_m an order of magnitude lower than the substrates with P1′ Arg. A detailed kinetic analysis of the hydrolysis of fsa-RRAA provided a k_{cat} value of 2.2 ± 0.2 s^{-1} and a K_m value of 0.5 ± 0.1 mM. Substrates with Arg at both P1′ and P3′ were hydrolyzed somewhat less efficiently than those with Arg at only P1′, supporting the proposal by Schellenberger et al.[24] that both of these positively charged residues may interact with the same negatively charged residue(s) in the enzyme.

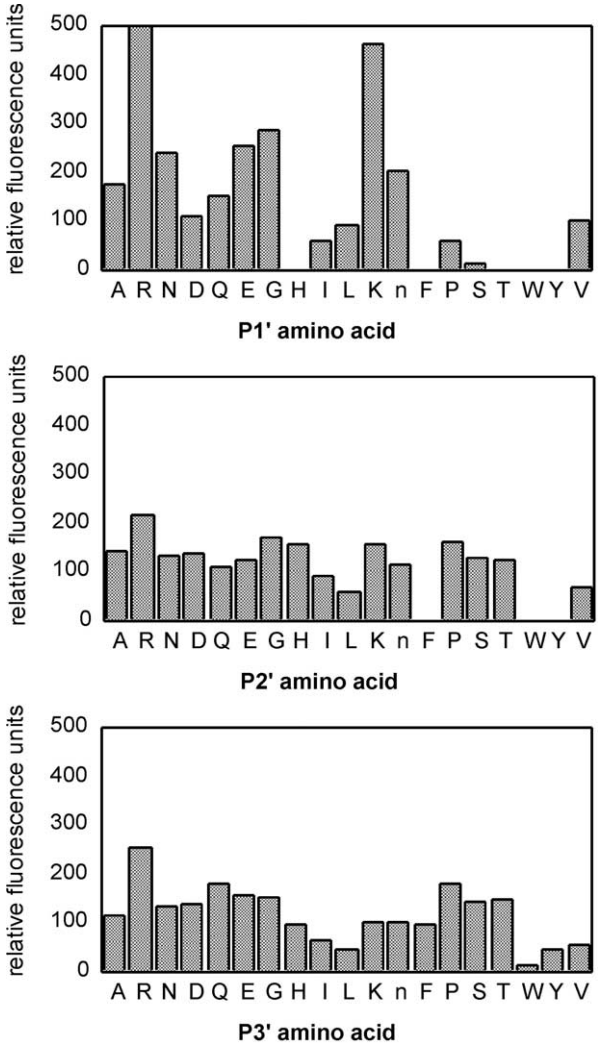


Figure 2. The prime side substrate specificity of bovine α-chymotrypsin as determined by the combinatorial fsa tetrapeptide substrate library.

Table 1. Single substrate kinetics with bovine α-chymotrypsin

Substrate	k_{cat}/K_m (M^{-1} s^{-1})
fsa-RRAA	4400 ± 1000
fsa-RAAA	3500 ± 800
fsa-RRRA	2200 ± 500
fsa-RARA	1900 ± 500
fsa-ARRA	160 ± 60
fsa-ARAA	NA
fsa-AARA	NA

The prime site specificity for bovine α-chymotrypsin determined in these experiments matches quite well with that previously obtained using the acyl transfer method[15,24,25] and can be interpreted based on the crystal structures of the enzyme complexed with macromolecular inhibitors as summarized in Figure 3.[26] The residues Asp35 and Asp64 have both been predicted to interact with the P1′ and P3′ residues. Closer inspection of the crystal structure of bovine α-chymotrypsin complexed with turkey ovomucoid protein third domain reveals a possible cation–π interaction[27] (approximately 3.89 Å) between the P3′ Arg residue of the inhibitor and Phe41, providing another potentially favorable interaction with the positively charged residue. Our data provide further evidence for the cooperativity between the S1′ and S3′ subsites in chymotrypsin,[24] confirming that single substrates with Arg in both the P1′ and P3′

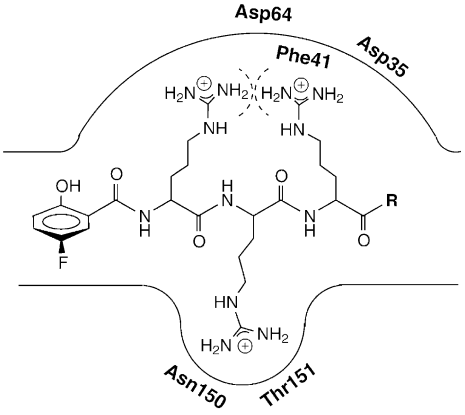


Figure 3. A representation of the prime side subsites in bovine α-chymotrypsin. The P1′ and P3′ subsites overlap, resulting in an interdependence between these two sites with Asp35 and Asp64 providing the major specificity determinants. Phe41 lies beneath the peptide, providing a potential cation–π interaction with the Arg residue in P3′ and its backbone oxygen atom providing a hydrogen bond acceptor for the amide N–H moiety of the P2′ residue. The P2′ pocket is less well defined, with potential hydrogen bond donors and acceptors such as Asn150 and Thr151 available.

positions are hydrolyzed less efficiently than those with only one Arg. In the S2′ subsite, a hydrogen bonding interaction between the amide NH group of the substrate and an oxygen from Phe41 of the enzyme has been the major determinant of specificity reported.[24] Our library results show a relatively broad specificity in this site, but the single substrate kinetics data indicate that Arg may be slightly more favorable than Ala at P2′. Chymotrypsin does not seem to display a well-defined pocket for the P2′ residue of the substrate, but there are numerous opportunities for hydrogen bonding interactions between the P2′ guanidinium group and backbone and side chains from the enzyme, with Asn150 and Thr151 as two likely candidates.

In this study, we have developed a novel fluorogenic assay for proteolytic cleavage of peptide substrates based on ligand enhancement of lanthanide ion fluorescence. The fsa fluorophore has been incorporated into a positional scanning synthetic combinatorial tetrapeptide library. The resulting library shows promise in allowing a facile determination of the extended prime site specificity of proteolytic enzymes. Work is currently underway in our laboratory to expand the utility of these prime site libraries and apply them in the determination of the prime site specificity of a variety proteolytic enzymes including both endo- and exopeptidases. It is anticipated that enzymes with a preference for aromatic residues in the P1 position will be most reactive with this library containing the fsa ligand in the P1 site. In addition, we expect that the lanthanide-based fluorogenic assay described herein will prove useful in assaying the activity of other hydrolytic enzymes including protein tyrosine phosphatases, alkaline phosphatases, and lipases.

Acknowledgements

This work was supported by NIH grants AI35707 and CA72006 to CSC. AMB thanks the NIH for a postdoctoral fellowship and Ms. Hannah Haigler for assistance in synthesizing the tetrapeptide library.

References and Notes

1. Barrett, A. J.; Rawlings, N. D.; Woessner, J. F. *Handbook of Proteolytic Enzymes*; Academic Publishers: London, 1998.
2. Harris, J. L.; Craik, C. S. *Cell* **2000**, *101*, 136.
3. Dunn, B. *Nat. Biotechnol.* **2000**, *18*, 149.
4. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157.
5. Backes, B. J.; Harris, J. L.; Leonetti, F.; Craik, C. S.; Ellman, J. A. *Nat. Biotechnol.* **2000**, *18*, 187.
6. Harris, J. L.; Backes, B. J.; Leonetti, F.; Mahrus, S.; Ellman, J. A.; Craik, C. S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7754.
7. Rano, T. A.; Timkey, T.; Peterson, E. P.; Rotonda, J.; Nicholson, D. W.; Becker, J. W.; Chapman, K. T.; Thornberry, N. A. *Chemistry Biology* **1997**, *4*, 149.
8. Thornberry, N. A.; Rano, T. A.; Peterson, E. P.; Rasper, D. M.; Timkey, T.; Garcia-Calvo, M.; Houtzager, V. M.; Nordstrom, P. A.; Roy, S.; Vaillancourt, J. P.; Chapman, K. T.; Nicholson, D. W. *J. Biol. Chem.* **1997**, *272*, 17907.
9. Takeuchi, T.; Shuman, M. A.; Craik, C. S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11054.
10. Dauber, D. S.; Ziermann, R.; Parkin, N.; Maly, D. J.; Mahrus, S.; Harris, J. L.; Ellman, J. A.; Petropoulos, C.; Craik, C. S. *J. Virol.* **2002**, *76*, 1359.
11. Harris, J. L.; Peterson, E. P.; Hudig, D.; Thornberry, N. A.; Craik, C. S. *J. Biol. Chem.* **1998**, *273*, 27364.
12. Matthews, D. J.; Wells, J. A. *Science* **1993**, *260*, 1113.
13. Krafft, G. A.; Wang, G. T. *Methods Enzymol.* **1994**, *241*, 70.
14. Matayoshi, E. D.; Wang, G. T.; Krafft, G. A.; Erickson, J. *Science* **1990**, *247*, 954.
15. Schellenberger, V.; Turck, C. W.; Hedstrom, L.; Rutter, W. J. *Biochemistry* **1993**, *32*, 4349.
16. Coates, J.; Sammes, P. G.; West, R. M. *J. Chem. Soc., Perkin Trans. 2* **1996**, 1275.
17. Coates, J.; Sammes, P. G.; West, R. M. *J. Chem. Soc., Perkin Trans. 2* **1996**, 1283.
18. Christopoulos, T. K.; Diamandis, E. P. *Anal. Chem.* **1992**, *64*, 342.
19. Evangelista, R. A.; Pollak, A.; Templeton, E. F. G. *Anal. Biochem.* **1991**, *197*, 213.
20. Pinilla, C.; Appel, J. R.; Blanc, P.; Houghten, R. A. *Biotechniques* **1992**, *13*, 901.
21. Houghten, R. A.; Pinilla, C.; Appel, J. R.; Blondelle, S. E.; Dooley, C. T.; Eichler, J.; Nefzi, A.; Ostresh, J. M. *J. Med. Chem.* **1999**, *42*, 3743.
22. Ostresh, J. M.; Winkle, J. H.; Hamashin, V. T.; Houghten, R. A. *Biopolymers* **1994**, *34*, 1681.
23. All 20 naturally occurring amino acids except methionine and cysteine were included in the library. Norleucine, an unnatural amino acid isosteric with methionine was also included.

24. Schellenberger, V.; Turck, C. W.; Rutter, W. J. *Biochemistry* **1994**, *33*, 4251.
25. Schellenberger, V.; Schellenberger, U.; Mitin, Y. V.; Jazkube, H.-D. *Eur. J. Biochem.* **1990**, *187*, 163.
26. Fujinaga, M.; Sielecki, A. R.; Read, R. J.; Ardelt, W.; Laskowski, M., Jr.; James, M. N. G. *J. Mol. Biol.* **1987**, *195*, 397.
27. Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303.