Pergamon

Bioorganic & Medicinal Chemistry Letters 10 (2000) 2291–2294

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Synthesis and Enzymatic Evaluation of a P_1 Arginine Aminocoumarin Substrate Library for Trypsin-Like Serine Proteases

Philip D. Edwards,* Russell C. Mauger, Kevin M. Cottrell, Frank X. Morris, Kara K. Pine, Mark A. Sylvester, Clay W. Scott and Stephen T. Furlong

Departments of Chemistry and Neuroscience, AstraZeneca Pharmaceuticals, PO Box 15437, Wilmington, DE 19850-5437, USA

Received 9 May 2000; accepted 2 August 2000

Abstract—A method for the solid-phase synthesis of P_1 arginine containing peptides via attachment of the arginine side-chain guanidine group is described. This procedure is applied to the preparation of a tetrapeptide, P_1 arginine aminocoumarin PS-SCL. This library was validated by using it to determine the P_4–P_2 specificity for thrombin and comparing the results to the known thrombin subsite specificity. This is the first reported example of a PS-SCL library containing a P_1 arginine.

Historically, the amino acid sequence flanking the cleavage site of a protein substrate has been used to develop synthetic substrates and peptidic inhibitors of proteolytic enzymes. However, the natural cleavage sequence may not be the optimal sequence for binding to the enzyme's active site. This was dramatically demonstrated for interleukin-converting enzyme (ICE) by Rano et al.[1] who used a positional scanning synthetic combinatorial library (PS-SCL) to prepare and analyze all possible tetrapeptide aminocoumarin substrates containing an aspartic acid residue at the P_1 position. When this library was screened against ICE, it was discovered that the optimal P_4–P_1 substrate sequence was WEHD not YVAD as previously believed. The successful construction of this PS-SCL was facilitated by the ability to attach the invariant P_1 aspartate residue to a solid-phase resin via the carboxylate of the P_1 aspartate side chain.

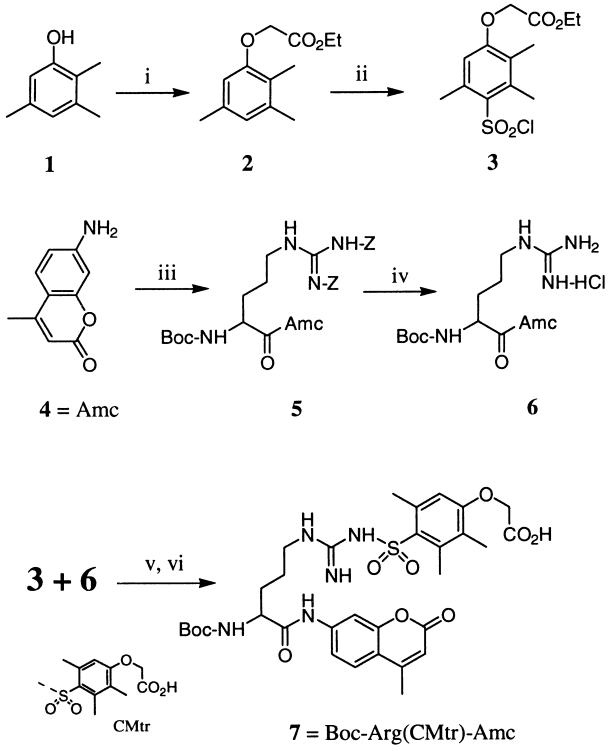
Since trypsin-like serine proteases prefer a basic arginine residue at the P_1 position, we sought to prepare a PS-SCL in which the invariant P_1 arginine residue was attached to the resin via its guanidine side chain. At the time we conceived this project, there were no reported methods for conducting SPPS using an arginine residue linked to a polymer via its side chain. In general, the versatility of the PS-SCL approach to substrate libraries is limited by the inability to attach most amino acids to a solid support via their side chain. Recently, Backes et al.[2] reported a strategy to overcome this generic problem in which the carboxyl group of the P_1 residue is attached to the resin via a safety catch linker. Displacement of the peptide with 7-amino-4-methylaminocoumarin followed by side-chain deprotection affords the peptide aminocoumarin substrate. This method avoids the necessity of attaching the P_1 residue to the resin via its side chain, and should, in principle, enable production of PS-SCL substrate libraries with any P_1 residue. However, in the report by Backes et al.,[2] only a P_1 lysine library for thrombin and plasmin was described, not a P_1 arginine library. Since arginine, not lysine, is the preferred P_1 residue for these trypsin-like enzymes, the question is raised as to whether synthetic difficulties prevented the preparation of a P_1 arginine substrate library. Due to the therapeutic importance of trypsin-like serine proteases, the need for P_1 arginine PS-SCL substrate libraries is great. In this paper, we describe the first reported process for constructing a P_1 arginine aminocoumarin (AMC) tetrapeptide substrate PS-SCL via solid-phase techniques, and validate the fidelity of the substrate library using the trypsin-like serine protease thrombin.

*Corresponding author. Tel.: +1-302-886-8371; fax: +1-302-886-5382; e-mail: philip.edwards@astrazeneca.com

Abbreviations: DMAP, 4-(dimethylamino)pyridine; DIPEA, diisopropylethylamine; DIPCI, diisopropyl carbodiimide; HATU, *N*-[(dimethylamino)-1*H*-1,2,3-triazol[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethan aminium hexafluorophosphate *N*-oxide; HOBt, 1-hydroxybenxtriazole hydrate; PS-SCL, positional scanning synthetic combinatorial library; SPPS, solid-phase peptide synthesis.

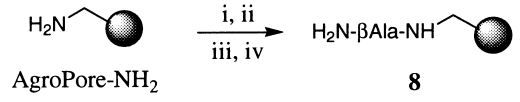
0960-894X/00/$ - see front matter
PII: S0960-894X(00)00460-1

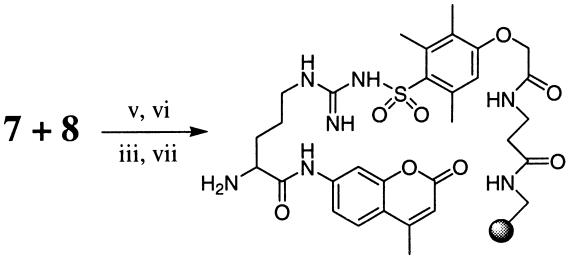
While there have been several reports of the solid-phase synthesis of amidines and guanidines,[3] little has been published on the solid-phase synthesis of peptides containing arginine linked to the resin through the side-chain guanidine group. As we began considering a strategy for attaching an arginine to a solid-phase resin via its side chain guanidine group, two reports appeared in the literature describing such a procedure.[4,5] Zhong et al.[4] employed an aromatic sulfonyl linker which is conceptually similar to the procedure we ultimately employed. However, attachment of the arginine guanidine to the resin-bound sulfonyl group required 4 N KOH at 75 °C for 2 days, conditions incompatible with the presence of an aminocoumarin amide required in our synthesis. Most relevant to our needs was a poster presentation by Urban and Dattilo describing the solid-phase preparation of a P_1 arginine tripeptide *p*-nitroanilide substrate using a novel CMtr linker. The CMtr is a modified Mtr containing a carboxyl group for attachment to the resin.[5] We used a modified version of this procedure for preparing the intermediate Boc-Arg (CMtr)-Amc **7** (Scheme 1). The chemistry is straightforward with a few notable exceptions. The sulfonyl chloride **3** is relatively unstable, and the reaction time for its formation should be limited to 1 h. Furthermore, **3** should be used in the subsequent sulfonamide forming reaction immediately upon solvent removal. We have also found that the best overall yields are obtained when sulfonamide formation and ester hydrolysis are carried out in two independent, sequential reactions.



Scheme 1. (i) Ethyl bromoacetate, K_2CO_3, DMF, rt, 18 h, 75%; (ii) HSO_3Cl, CH_2Cl_2, 4 °C, 1 h, 85%; (iii) Boc-Arg(Z)$_2$-OH, Boc_2O, Py, THF, rt, 18 h, 76%; (iv) $Pd(OH)_2$, EtOH, 1 N HCl, H_2/48 psi, 4 h, 93%; (v) 1 N NaOH (2 equiv), acetone, rt, 0.5 h, 32%; (vi) 1 N NaOH (1 equiv), THF/EtOH (1:1), rt, 3 h, 82%.

A significant amount of effort was required to identify appropriate conditions for attaching **7** to a solid support (Scheme 2). We initially investigated two different PEG polystyrene resins. Since the amino group attachment point in these resins is distant from the cross-linked resin backbone, we felt that these resins would afford a more efficient amide coupling with the relatively bulky acid **7**. However, the physical properties of the coupled products were such that severe mechanical losses were encountered. These losses continued to plague us throughout the solid-phase synthetic sequence. To overcome this problem and still maintain an extended attachment point to the resin, we attached a β-alanine linker to an AgroPore^{TM}-NH_2 resin.[6] When coupled with **7**, the resin bound intermediate **9** had excellent handling properties and no significant mechanical losses were observed. The couplings of both the β-alanine to the resin and of **7** to **8** were monitored by following the growth of the amide carbonyl absorption peak of a bead using solid-phase IR since a negative Kaiser test could not be obtained. Maximum loading was considered the point where there was no further increase in the carbonyl absorption, and was achieved using a two-step sequence involving basic coupling conditions (HATU and DIPEA) followed by a more acidic system (DIPCI





viii, ix

repeat viii, ix twice

iii (twice), x

20 pots, known P_2 **Ac-O-O-X-Arg-Amc** = P_2 sublibrary

20 pots, known P_3 **Ac-O-X-O-Arg-Amc** = P_3 sublibrary

20 pots, known P_4 **Ac-X-O-O-Arg-Amc** = P_4 sublibrary

Scheme 2. (i) Boc-βAla-OH, HATU, DIPEA, DMF, 18 h, repeat for 3 h; (ii) HOBT, DIPCI, DMAP (cat), DMF, 18 h, repeat for 3 h; (iii) Ac_2O, DIPEA, DMF, 2 h; (iv) 40% TFA/CH_2Cl_2, 0.5 h; (v) HATU, DIPEA, DMAP (cat), DMF, 18 h; (vi) HOBt, DIPCI, DMAP (cat), DMF, 18 h; (vii) 40% TFA/CH_2Cl_2, 1 h; (viii) isokinetic mixture or known amino acid, HATU, DIPEA, DMF, 2 h, repeat procedure once; (ix) 20% piperidine, DMF, 15 min, repeat procedure once; (x) thioanisole/TMSBr/triisopropyl silane/phenol/TFA (ref 8).

and HOBt). Unreacted amino groups on the resin were then capped with acetic anhydride.

The strategy behind the tetrapeptide PS-SCL is the generation of three sublibraries, one library for each sequence position evaluated (P_2, P_3, and P_4). Each sublibrary is composed of one pot for each amino acid evaluated. In the current study, 20 amino acids were investigated resulting in 20 pots per sublibrary. In each pot, the residue at one sequence position (e.g., P_2) is known. Using a P_2 sublibrary as an example, this is achieved by dividing **9** into 20 pots and coupling a different amino acid in each pot. In the next coupling sequence, an isokinetic mixture of the 20 amino acids is used to generate all possible 20 tripeptides in each pot in equal proportions. For the final coupling, the same isokinetic mixture is used to generate in each of the 20 pots all possible 400 tetrapeptide substrates with a different, known P_2 residue. Following cleavage from the resin, each pot in the sublibrary is assayed against the enzyme and the optical density (OD) measured. Since each pot contains all possible tetrapeptides with a known P_2 residue, those pots displaying the highest OD contain substrates that are most efficiently cleaved and therefore the preferred P_2 residues. Similar analysis of the P_3 and P_4 sublibraries identifies the preferred P_3 and P_4 residues.

Solid-phase synthesis of each sublibrary was accomplished using Fmoc protected single amino acids or an isokinetic mixture of 20 amino acids, HATU coupling, and piperidine deprotection. The loading was monitored using the Kaiser test, with a negative test being obtained before proceeding with the deprotection. The HATU coupling conditions and proportions of amino acids in the isokinetic mixture were those developed by Herman et al.[7] Following cleavage from the resin, the crude material was isolated by solvent removal and subsequent ether trituration. No further purification was employed so as to avoid selective removal of any peptides. Initially, cleavage of the peptides from the resin using a standard TFA:H_2O:phenol:TIS (87.5:5:5:2.5) cocktail and a PEG polystyrene resin afforded only a 5% yield of peptide mixture. Changing to the AgroPore^{TM}-NH_2 resin and employing a stronger, TMSBr-based cleavage cocktail[8] increased the yield to 50–70% (75–105 mg/pot).

We conducted a limited chemical characterization of the library. The P_3 sublibrary Ac-O-Val-O-Arg-Amc and Ac-O-Asp-O-Arg-Amc pots were analyzed by means of MALDI-TOF mass spectrometry. The mass spectral data was assessed for the presence of three features: (1) that the smallest (-G-V-G-) and largest (W-V-W-) molecular weight peptides were present; (2) that the center of the distribution for each of the two pots was shifted from each other by 16, the difference in molecular weight between valine and aspartic acid; and (3) that the molecular weight distribution was centered around the average molecular weight. Figure 1 is a comparison of the mass spectra for the -Val- and -Asp- pots. As anticipated, the high and low molecular weight compounds were present and the center of distribution between each library was shifted by 16. The average molecular weights for the Ac-O-Val-O-Arg-Amc and Ac-O-Asp-O-Arg-Amc pots are 705 and 721, respectively. The centers of distribution in Figure 1 are in good agreement with these calculated values.

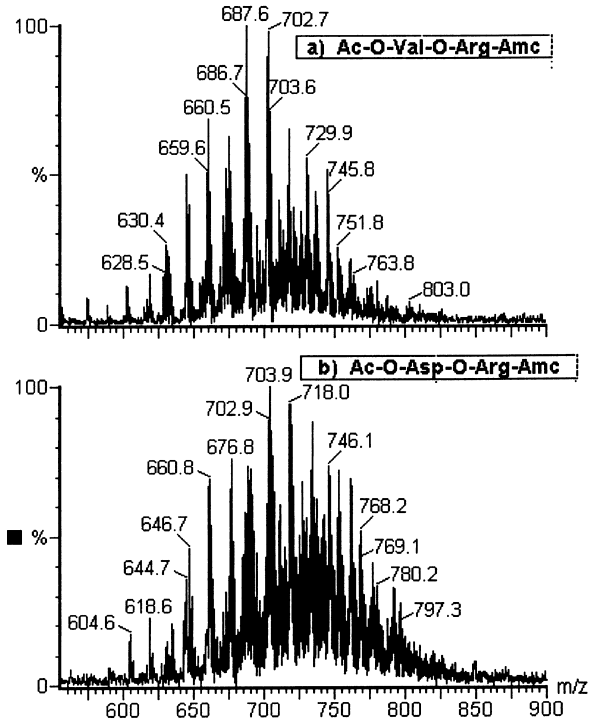


Figure 1. MALDI-TOF mass spectra (reflectron mode, alpha-cyano hydroxycinnamic acid matrix) of P_3 sublibrary pots (a) Ac-O-Val-O-Arg-Amc and (b) Ac-O-Asp-O-Arg-Amc.

While the chemical characterization of the library suggested that equimolar mixtures of peptides were present in the pots analyzed, the most important and practical means for validating the fidelity of the library was to assay each sublibrary against an enzyme with known substrate specificity. Thrombin was chosen because it is a trypsin-like serine protease of current medicinal interest. In addition, thrombin has a varied subsite specificity that we felt would be ideal for validating the library: an extremely strong preference for proline at P_2; a relatively broad specificity at P_3; and a preference for hydrophobic residues at P_4.[9] The libraries were assayed by incubating each sample with enzyme and measuring the fluorescence.

The results against each sublibrary are depicted in Figure 2. In the P_2 library, only five amino acids demonstrated any significant hydrolysis by thrombin, all small hydrophobic residues. Proline (P) was clearly preferred. Consistent with the literature, a broad variety of residues were tolerated at P_3, including basic, polar, and both large and small hydrophobic residues. Proline, which is a unique cyclic amino acid that imparts a turn into the peptide backbone, was completely inactive at the P_3 position. Significantly, valine (V), leucine (L) and isoleucine (I) were preferred over phenylalanine (F) as recently reported for peptidic chloromethyl ketone inhibitors.[10] Particularly interesting is the comparison of aspartic acid (D) and its amide asparagine (N) with

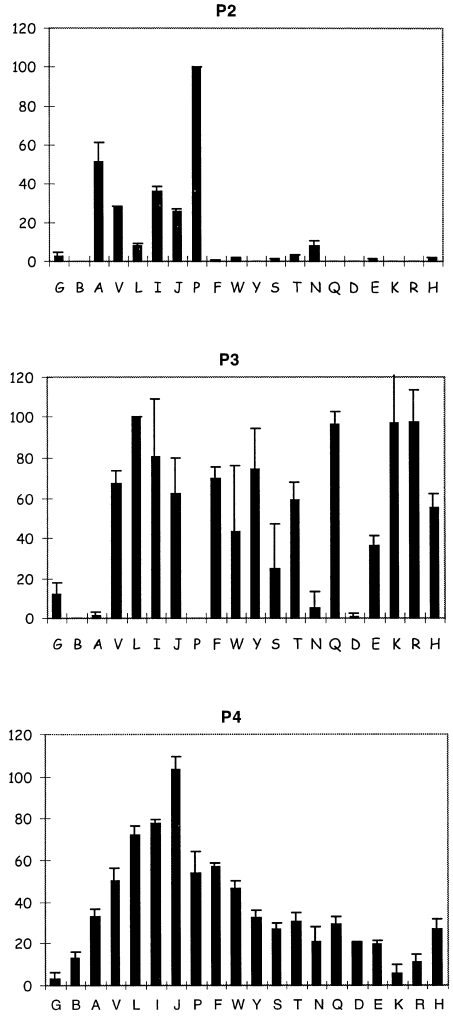


Figure 2. Substrate specificities for human thrombin. Each bar is an *n* of 3. The *x* axis represents substituted amino acids (B = D-alanine; J = norleucine). The *y* axis is AMC production expressed as a percentage of the maximum rate observed in each experiment.

glutamic acid (E) and its amide glutamine (Q). Both D and N were inactive, while E and Q, which have an additional methylene in their side chain, were good substrates, suggesting a critical side chain length requirement. Finally, hydrophobic residues are preferred at P_4. These results agree remarkably well with those obtained by Backes et al.[2] using their P_1 lysine substrate library. The subtle differences observed between the two libraries, such as detectable hydrolysis of nonproline P_2 substrates in the current P_1 arginine library, may be due either to different detection threshold levels in the enzyme assays, or due to different subsite specificities between P_1 arginine and P_1 lysine substrates.

In summary, a procedure has been developed for efficiently attaching an arginine to a solid support via its guanidine side chain, and cleaving it from the resin following subsequent elaboration. This methodology was used to construct a P_1 arginine aminocoumarin substrate PS-SCL, and validate the library against thrombin, a trypsin-like serine protease. This extremely important class of enzymes prefers an arginine at P_1, and the methodology described herein is the first process reported to allow construction of PS-SCL substrate libraries containing a P_1 arginine.

References and Notes

1. Rano, T. A.; Timkey, T.; Peterson, E. P.; Rotonda, J.; Nicholson, D. W.; Becker, J. W.; Chapman, K. T.; Thronberry, N. A. *Chem. Biol.* **1997**, *4*, 149.
2. Backes, B. J.; Harris, J. L.; Leonetti, F.; Craik, C. S.; Ellman, J. A. *Nat. Biotechnol.* **2000**, *18*, 187.
3. (a) Josey, J. A.; Tarlton, C. A.; Payne, C. E. *Tetrahedron Lett.* **1998**, *39*, 5899. (b) Roussell, P.; Bradley, M.; Matthews, I.; Kane, P. *Tetrahedron Lett.* **1997**, *38*, 4861. (c) Bonnat, M.; Bradley, M.; Kilburn, J. *Tetrahedron Lett.* **1996**, *37*, 5409.
4. Zhong, H. M.; Greco, M. N.; Maryanoff, B. E. *J. Org. Chem.* **1997**, *62*, 9326.
5. Urban, J.; Dattilo, J. W. *Abstracts of Papers*, 15th American Peptide Symposium, Nashville, TN, 14–19 June, 1997.
6. AgroPore^{TM}-NH2 resin was purchased from Argonaut Technologies.
7. Herman, L. W.; Tarr, G.; Kates, S. A. *Mol. Divers.* **1996**, *2*, 147.
8. (a) Hughes, J. L.; Leopold, E. J. *Tetrahedron Lett.* **1993**, *48*, 7713. (b) Guo, L.; Funakoshi, S.; Fujii, N.; Yajima, H. *Chem. Pharm. Bull.* **1988**, *36*, 4989.
9. Bode, W.; Turk, D.; Karshikov, A. *Protein Sci.* **1992**, *1*, 426.
10. Vindigni, A.; Dang, Q. D.; DiCera, E. *Nat. Biotechnol.* **1997**, *15*, 891.