JOURNAL OF
MEDICINAL
CHEMISTRY

© Copyright 1995 by the American Chemical Society

Volume 38, Number 2 January 20, 1995

Communications to the Editor

Validation of Screening Immobilized Peptide Libraries for Discovery of Protease Substrates

Jasbir Singh,*,† Martin P. Allen,† Mark A. Ator,‡ James A. Gainor,† David A. Whipple,† James E. Solowiej,‡ Adi M. Treasurywala,§ Barry A. Morgan,† Thomas D. Gordon,† and Donald A. Upson†

Departments of Medicinal Chemistry, Enzyme and Receptor Biochemistry, and Biophysics and Computational Chemistry, Sterling Winthrop Pharmaceuticals Research Division, 1250 South Collegeville Road, P.O. Box 5000, Collegeville, Pennsylvania 19426

Received October 21, 1994

With the explosive progress in molecular biology, new proteases have been identified as attractive therapeutic targets. For most of these new proteases we lack detailed specificity and selectivity information, which is necessary for the design of efficient screens for small molecule inhibitors of the enzyme. This information also contributes to the synthetic design of potent and selective mechanism-based inhibitors. The combinatorial approach is an attractive strategy for rapidly assessing this comprehensive information. Implementation of automation is critical for timely success with this method, which demands rapid synthesis and high throughput screening of large numbers of samples. Automated solid-phase-based synthesis is an efficient and practical approach to obtain these required samples. We reasoned that if one could carry out the screening of these solid support-bound ("immobilized") peptides,[1] then the goal of protease specificity assessment[2] could be accomplished with minimal sample manipulation. However, the critical question is whether the hits obtained from the screening of immobilized samples have any relevance to the hits obtained from screening of the corresponding soluble samples.[3] In this paper, we report the strategy for screening an unbiased peptide library and the correlation of the rates of hydrolysis of immobilized and soluble peptide substrates for the matrix metalloprotease, collagenase.

Known, potent inhibitors for proteases (e.g., matrix metalloproteases, human immunodeficiency virus protease, human leukocyte elastase, etc.) span generally four residues in length. Therefore, the key design features of our combinatorial peptide library were (a) focused acquisition of substrate specificity and selectivity information on a four amino acid window and (b) construction of a generic library which neither requires nor presumes any prior knowledge of the substrate specificity of the protease under investigation. The combinatorial peptide arrays AAAXXBA, AAXAXBA and AAXXABA, satisfy these requirements. In these arrays,[4] X represents 20 amino acids present one at a time, B represents a degenerate position in which all 20 amino acids are present in equimolar amounts, and A is an alanyl residue (Scheme 1). These arrays represent 24 000 sequences.

Scheme 1

COP-AAAXXBA-$(Acp)_5$-βAla-AMP-CPG

COP-AAXAXBA-$(Acp)_5$-βAla-AMP-CPG

COP-AAXXABA-$(Acp)_5$-βAla-AMP-CPG

where: COP = 7-Hydroxycoumarin-4-propanoyl
AMP = 3-Aminopropyl
Acp = 6-Aminocaproyl
X and B = See text for the definitions
CPG = Controlled-pore glass

monomers used are A (Ala), D (Asp), E (Glu), F (Phe), G (Gly), H (His), I (Ile), K (Lys), L (Leu), M (Met), N (Asn), P (Pro), Q (Gln), R (Arg), S (Ser), T (Thr), U (Smc), V (Val), W (Trp), Y (Tyr), where Smc = *S*-methylcysteine

A cornerstone of our strategy has been the assembly and subsequent protease assay of the combinatorial

† Department of Medicinal Chemistry.
‡ Department of Enzyme and Receptor Biochemistry.
§ Department of Biophysics and Computational Chemistry.

0022-2623/95/1838-0217$09.00/0 © 1995 American Chemical Society

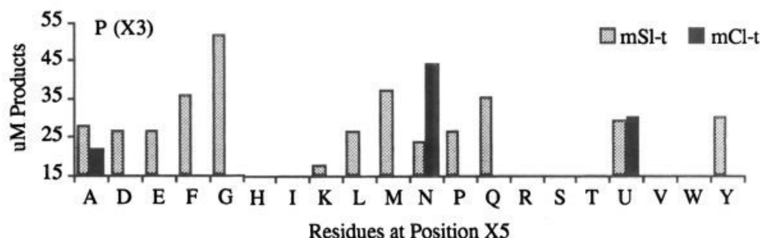

Figure 1. Screening results for truncated mature human fibroblast collagenase (mCl-t) and truncated mature human fibroblast stromelysin (mSl-t) against COP-AAPAX$_5$BA-(Acp)$_5$-βAla-AMP-CPG samples. Activity reported as micromolar product is based on the fluorescence of the soluble product released in solution following proteolyis. A fixed time point assay was utilized for screening of the whole immobilized peptide library.

peptide library immobilized on controlled-pore glass support,[5] a practical approach for screening a large number of samples. This approach allows analysis of active samples in two steps: quantitation of protease activity by fluorogenic detection of the soluble product followed by microsequencing of the remaining solid support-bound peptide to determine the site of processing.

The Ala-(Acp)$_5$-βAla-AMP-CPG portion, which is common to the entire library, was prepared by manual synthesis on a 250 g scale.[6-8] The degenerate position "B" was assembled by the "mix and split" method first described by Furka et al.[9] to provide BA-(Acp)$_5$-βAla-AMP-CPG (BA-glass). Assembly of the immobilized combinatorial peptide library was carried out by automated multiple peptide synthesis on BA-glass, as reported previously.[10] The automated synthesis provides preparation of every sample of this immobilized combinatorial peptide library in quantities sufficient for 10–15 screening assays. The use of controlled-pore glass has been a key for the success of this approach by permitting automation of synthesis, weighing and enzymology utilizing the ACT-MPS 350, a modified Hewlett-Packard ORCA and Packard Multi-PROBE robots, respectively. The results of screening the immobilized array,[11] $A_1A_2X_3A_4X_5B_6A_7$ vs truncated[12] mature human fibroblast collagenase and stromelysin, highlight the utility of this approach for the determination of substrate selectivity among closely related proteases. Analysis of the screening results (mixture of 20 sequences per tube) reveals that proline[13] at position X_3 is essential for the peptides to be substrates for human fibroblast stromelysin and not as critical for human fibroblast collagenase (data not shown). Within the sequence COP-AAPAX$_5$BA-(Acp)$_5$-βAla-AMP-CPG, stromelysin tolerates the presence of a variety of residues at X_5, while collagenase demonstrates a higher degree of specificity (Figure 1). Microsequencing on the glass-bound peptides following the proteolysis reveals the contents at position B leading to the observed substrate activity for a given sample. Sequencing results (Figure 2) of the active samples AAPAABA and AAPANBA (from reaction with human fibroblast collagenase) revealed the site of processing and the composition AAPANUA (Ala-Ala-Pro-Ala-Asn-Smc-Ala) to be a good substrate for human fibroblast collagenase. In addition, the P_3 = Pro (P), P_1 = Asn (N), and P_1' = Smc (U) information is useful for the design of human fibroblast collagenase inhibitors. The analysis of these two active samples [AAPA(A,N)BA] serves to illustrate our strategy for a complete resolution of every active component of the peptide library under study from one round of screening without requiring synthesis of "iterative" peptide libraries.

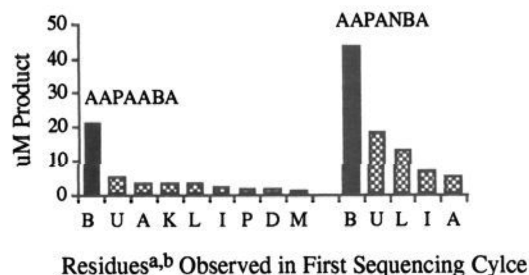

Figure 2. Results of the Edman microsequencing of active tubes vs truncated mature human fibroblast collagenase. Following identification of active tubes, the solid portion of the reaction mixture for these tubes following proteolysis was thoroughly washed and subjected to Edman microsequencing: (a) see monomer list in Scheme 1 for the corresponding three letter codes, (b) columns marked with "B" represent total amount of solid-bound product present, which is identical to the amount of soluble product as shown in figure 1 for X_5 = A and X_5 = N samples).

Utilization of the results from screening of an immobilized combinatorial peptide library for the design of either an inhibitor or a continuous substrate for assays in the absence of a solution vs immobilized correlation requires a leap of faith. Therefore, we have undertaken to show the usefulness and validation of this strategy. Octapeptide substrate sequences for which the k_{cat}/K_m values vary over 2 orders of magnitude for human fibroblast collagenase[14] were chosen. The k_{cat}/K_m data for the 38 octapeptide substrates were analyzed in SYBYL as histograms. A set of six sequences[15] representing three and four standard deviations separation were selected, and the corresponding immobilized sequences were prepared using the immobilized combinatorial peptide library construct, [COP-sequence-(Acp)$_5$-βAla-AMP-CPG]. These immobilized peptides were assayed vs human fibroblast collagenase. A plot (Figure 3) of the soluble (rel ratio of k_{cat}/K_m):immobilized (rel ratio of products formed) peptide substrates assay data shows an excellent correlation of $r^2 = 0.994$. This result demonstrates that the kinetics of hydrolysis of immobilized peptides are predictive of the reaction of their soluble counterparts, validating the use of immobilized peptides for high throughput screening. We believe that this report will provide a basis for extending such a correlation to receptor-binding ligands. Furthermore, screening of immobilized samples of non-peptides should be evaluated as an approach for rapid lead identification and development.

In conclusion, we have defined a method for the rapid synthesis of an unbiased peptide library and demon-

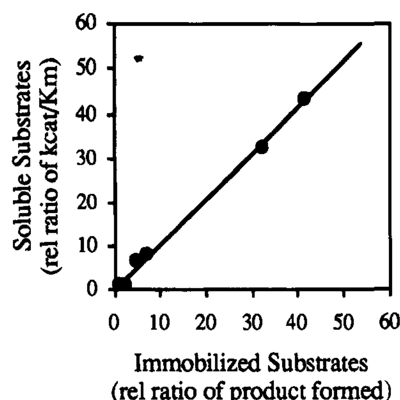

Figure 3. Correlation of kinetics of hydrolysis of soluble and immobilized octapeptide substrates for collagenase. See ref 15 for sequence of six octapeptides for which the solution and the immobilized assay data is shown in the graph above.

strated a correlation between the relative k_{cat}/K_m ratio (a measure of substrate specificity) for soluble peptides and the relative substrate activity of corresponding immobilized peptides. Therefore, screening of an immobilized peptide combinatorial library offers a valid approach for the determination of protease substrate specificity and selectivity.

Acknowledgment. We acknowledge the support and encouragement of Dr. Lawrence I. Kruse for developing the immobilized peptide libraries.

References

(1) Immobilized peptide libraries have been used extensively for screening against antibodies and receptors. "Pin" technology developed by Geysen et al. (*Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 3998–4002) has been the earliest development in this rapidly growing field. Another major advance in the area of synthesis and screening of solid support-bound peptides has been made by Fodor et al. (*Science* **1991**, *251*, 767) with the development of Light-Directed, Spacially Addressable Parallel Chemical Synthesis technology. For an excellent review, see: Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P. A.; Gordon, E. M. *J. Med. Chem.* **1994**, *37*, 1233–1251 and references cited therein.

(2) (a) Since we carried out our initial studies, a series of papers have appeared which involve screening of peptides immobilized on polymer membrane as protease substrates: Wondrak, E. M.; Copeland, T. D.; Louis, J. M.; Oroszlan, S. A Solid Phase Assay for the Protease of Human Immunodeficiency Virus. *Anal. Biochem.* **1990**, *188*, 82–85. Ito, Y.; Lin-Shu, L.; Imanishi, Y. Interaction of Thrombin with Synthetic Fluorescent Substrate Immobilized on Polymer Membrane. *Biomaterials* **1992**, *13*, 375–381. Laursen, R. A.; Duan, Y. Protease Substrate Specificity Mapping Using Membrane-Bound Peptides. *Anal. Biochem.* **1994**, *216*, 431–438. (b) Recently, an example utilizing a biopolymer (proteins) as a solid support has also been described for the selection of protease substrates: Matthews, D. J.; Wells, J. A. Substrate Phage: Selection of Protease Substrates by Monovalent Phage Display. *Science* **1993**, *260*, 1113–1117. (c) The following papers have described protease specificity determination using screenings of soluble peptide libraries/mixtures: Petithory, J. R.; Masiarz, F. R.; Kirsch, J. F.; Santi, D. V.; Malcolm, B. A. A Rapid Method of Determination of Endoproteinase Substrate Specificity: Specificity of the 3C Proteinase from Hepatitis A Virus. *Proc. Natl. Acad. Sci. U.S.A* **1991**, *88*, 11510–11514. Berman, J.; Green, M.; Sugg, E.; Anderegg, R.; Millington, D. S.; Norwood, D. L.; McGeehan, J.; Wiseman, J. Rapid Optimization of Enzyme Substrates Using Defined Substrates Mixtures. *J. Biol. Chem.* **1992**, *267*, 1434–1437.

(3) To the best of our knowledge, no study showing a correlation between the hits obtained by screening immobilized samples to the corresponding soluble counterparts has been reported to date.

(4) Positions which are not varied have been fixed with alanine: a chiral amino acid residue with the smallest side chain. The ψ and ϕ torsions for alanine are representative of all nautral amino acids, except proline. Glycine and proline were not chosen for the fixed positions, as these residues will provide a strong bias for the peptide conformation.

(5) Our goal was to select an appropriate solid support which was compatible with both the chemistry of solid-phase peptide synthesis and enzymology in aqueous media. Fmoc-ProLeuAlaLeuPhe, a known (soluble) substrate for the metalloprotease thermolysin, was attached to (i) polystyrene resins thru a hexa(6-aminocaproyl) linker, (ii) CH-Sepharose 4b thru a *N,N*-dimethylethylenediamine linker, and (iii) controlled-pore glass (CPG) thru a hexa(6-aminocaproyl) linker. Peptide attached to Sepharose and CPG were processed by thermolysin, while the polystyrene-bound peptide was not processed at all. However, Sepharose is not compatible with the organic media employed in solid-phase peptide synthesis. In contrast, controlled-pore glass (CPG) seemed to possess the required properties, being compatible with both the chemistry of solid-phase peptide synthesis and enzymology in aqeous media.

(6) Controlled-pore glass has not been widely used for solid phase peptide synthesis. See: Albericio, F.; Porta, A.; Pedroso, E.; Giralt, E. Solid phase coupling of protected peptide segments. Influence of the polymer support. *Peptides* **1986**, 167–170.

(7) Controlled-pore glass (CPG) can be purchased from CPG Inc. (Fairfield, NJ) with two functionalities which are suitable for initiation of solid-phase peptide synthesis: we chose to use CPG functionalized with an aminopropyl group (AMP) with a loading of 86 μmol/g and a pore size of 500 Å.

(8) Incorporation of a β-Ala in the linker allowed an experimental validation of the homogeneity of the linker. Amino acid analysis of a small sample using β-Ala as an internal standard provided determination of the extent of coupling reactions during the synthesis of the linker. Multiple couplings were performed at each step to achieve quantitative yields.

(9) Furka, A.; Sebestyén, F.; Asgedom, M.; Dibó, G. General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **1991**, *37*, 487–493.

(10) Ator, M.; Beigel, S.; Dankanich, T.; Echols, M.; Gainor, J.; Gilliam, C.; Gordon, T.; Koch, D.; Kruse, L.; Morgan, B.; Olsen, R.; Siahaan, T.; Singh, J.; and Whipple, D. Immobilized Peptide Arrays: A New Technology for the Characterization of Protease Function. *Peptides: Chemistry Structure and Biology*; Proceedings of the 13th American Peptide Symposium; Hodges, R., Smith, J., Eds., 1994, pp 1012–1016 and references cited therein.

(11) Position numbering is shown for discussion purpose only.

(12) Brownell, J.; Earley, W.; Kunec, E.; Morgan, B. A.; Olyslager, B.; Wahl, R. C.; Houck, D. R. Comparison of Native Matrix Metalloproteases and Their Recombinant Catalytic Domains Using a Novel Radiometric Assay. *Arch. Biochem. Biophy.* **1994**, in press.

(13) For sake of clarity, the most dramatic and significant panel for mSl-t, X_3 = P has been shown here.

(14) Netzel-Arnett, S.; Fields, G.; Birkedal-Hansen, H.; Van Wart, H. Sequence Specificity of Human Fibroblast and Neutrophil Collagenases. *J. Biol. Chem.* **1991**, *266*, 6747–6755.

(15) The P_4–P_4' based substrate sequences, in the descending order of activity, are as follows: GlyProGlnGlyIleTrpGlyGln; GlyProGlnAlaIleAlaGlyGln; GlyProLeuGlyIleAlaGlyGln; GlyProGlnGlyIleAlaGlyGln; GlyProGlnGlyTyrAlaGlyGln; GlyProGlnGlyPheAlaGlyGln.

JM940697X