1-1-2004

# Probing Proteinase Active Sites Using Oriented Peptide Mixture Libraries – ADAM-10

John L. Krstenansky
*Marshall University*, krstenansky@marshall.edu

Jihong Wang

Gregg R. Chenail

Matthew R. Tiffany

Geoffrey M. Kuesters

***See next page for additional authors***



## Recommended Citation

Krstenansky JL, Wang J, Chenail GR, Tiffany MR, Kuesters GM, Natke BC, Nestor Jr JL. Probing proteinase active sites using oriented peptide mixture libraries – ADAM-10. Lett Drug Design Disc. 2004;1:6 –13.

**Authors**
John L. Krstenansky, Jihong Wang, Gregg R. Chenail, Matthew R. Tiffany, Geoffrey M. Kuesters, Barbara C. Natke, and John L. Nestor Jr.

# Probing Proteinase Active Sites Using Oriented Peptide Mixture Libraries – ADAM-10

John L. Krstenansky, Jihong Wang, Gregg R. Chenail, Matthew R. Tiffany, Geoffrey M. Kuesters, Barbara C. Natke and John J. Nestor, Jr.*

*Consensus Pharmaceuticals, Inc., 200 Boston Ave., Medford, MA 02155, USA*

***Received June 20, 2003: Accepted September 4, 2003***

**Abstract:** Oriented Peptide Mixture Libraries can provide a full matrix of preferred and disfavored amino acids at each subsite of an optimal substrate for a new proteinase. This approach is rapid and convenient, requiring only two mixture libraries to complete the analysis. In this paper we demonstrate an extension of this type of analysis, using a focused library employing unnatural amino acids to probe the depth of the S1 position in the catalytic site of the alpha secretase ADAM-10. This analysis indicates that ADAM-10 will accept amino acids with substantial length and hydrophobicity (e.g. 2-naphthylalanine), but suggests that the S1 site has limitations in the apparent "width" of substituents being presented (e.g. 1-naphthylalanine; gamma branching). A highly selective and efficient substrate for ADAM-10, with a selectivity factor of 380,000 $M^{-1}$ $s^{-1}$, was derived from the predicted consensus substrate. This detailed analysis provides a starting point for the design of inhibitors of this interesting proteinase.

**Keywords:** proteinase, proteinase inhibitor, proteinase substrate, ADAM-10, consensus substrate motif

## INTRODUCTION

Speed and cost effectiveness are crucial considerations during the drug discovery process. Proteases are attractive therapeutic targets in a number of disease areas and several important drugs are marketed for the inhibition of protease function [1]. For research involving enzyme targets, the ability to quickly obtain effective substrates for HTS assays and to create inhibitors for x-ray crystallography / structure-based design is essential.

A number of techniques have been used to deduce the substrate requirements for novel proteases, with positional scanning [2,3], substrate phage display [4, 5], and substrate mixture digestion being frequently cited [6,7]. Recently an approach that makes use of Oriented Peptide Mixture Libraries (OPML) has been described [8] and several advantages of this approach are apparent. For OPML, the substrate mixture is present in solution, so no artifacts from linkage to a solid phase or presentation in a protein are seen. Preferences for and against amino acid residues are determined both for the primed and unprimed side of the substrate, in contrast to positional scanning. For the OPML approach, a full matrix of preferences for and against each proteinogenic amino acid at each subsite is obtained rapidly, through the use of only 2 libraries.

Typically the OPMLs are constructed to have positions that are fully degenerate (all proteinogenic amino acids except Cys) surrounding a sequence of orienting residues that force binding to the active site in a common fashion for the full chain. Such a library may contain tens of billions of potential substrates, but upon exposure to the enzyme for a short period of time (typically ca. 1-5% conversion) only the most favored substrates are processed. Separation of the unprocessed portion of the library and sequencing from the N-terminus give the relative amounts of each residue found at each position in the processed pool of peptides. This matrix of data provides the relative favoring or disfavoring of each amino acid residue at that position in an optimal substrate.

The experiment is carried out as a two-step process [8]. First a fully degenerate, N-terminally blocked mixture library (Ac-X-X-X-X-X-X-X-X-$NH_2$; where X represents 19 proteinogenic amino acids, lacking Cys) is incubated with the enzyme. These peptide libraries are obtained by using isokinetic mixtures of amino acids being explored at each position of the substrate peptide [9]. Cleavage of the peptide at any position then opens up a site for N-terminal Edman sequencing. The first cycle provides the P1' site preference information, the second cycle provides the P2' site population, etc. In the second step, an OPML is generated using the primed side information, while full degeneracy is maintained on the unprimed side [i.e. $NH_2$-X-X-X-X-O-O-O-O-Lys(Biotin)-$NH_2$; where O represents a fixed, orienting residue]. This OPML has an unblocked amino terminus, so that Edman sequencing can be carried out, but has a biotin residue at the C-terminus so that uncleaved library members and the C-terminus of cleaved substrates can be removed by streptavidin chromatography. Thus following the enzyme incubation with the OPML, the mixture can be run through a streptavidin column and the flow through will contain only the amino terminal residues of substrates cleaved by the enzyme. Edman sequencing of this mixture for the library shown above yields P4 preference information in the first

*Address correspondence to this author at the Sequenom Pharmaceuticals, Inc., 3595 John Hopkins Court, San Diego, CA 92121, USA; E-mail: jnestor@sequenom.com

1570-1808/04 $45.00+.00 

| | A | D | E | F | G | H | I | K | L | M | N | P | Q | R | S | T | V | W | Y |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1' | 0.31 | 0.31 | 0.51 | 1.40 | 0.28 | 0.62 | 1.43 | 0.45 | ***3.08*** | 1.55 | 0.34 | 0.32 | 2.04 | 1.10 | 0.31 | 0.00 | 1.11 | 1.12 | 1.71 |
| 2' | 0.68 | 0.43 | 0.53 | 0.82 | 0.49 | 1.73 | 0.98 | ***2.29*** | 1.09 | 1.69 | 0.62 | 0.54 | 1.04 | 1.18 | 0.69 | 0.00 | 1.30 | 0.82 | 1.08 |
| 3' | 1.06 | 0.66 | 0.71 | 0.90 | 1.19 | 0.92 | 1.00 | 0.91 | 0.92 | 1.53 | 1.42 | 0.70 | 1.02 | 1.05 | 0.91 | 0.00 | 0.79 | 1.36 | 0.96 |
| 4' | 1.01 | 0.76 | 0.96 | 1.12 | 1.07 | 1.00 | 0.95 | 1.13 | 1.07 | 1.14 | 0.83 | 1.06 | 0.96 | 1.09 | 0.78 | 0.00 | 0.93 | 1.11 | 1.04 |
| 5' | 0.95 | 0.83 | 1.06 | 1.02 | 0.96 | 1.02 | 0.98 | 1.13 | 1.04 | 1.04 | 0.96 | 0.99 | 1.17 | 1.11 | 0.83 | 0.00 | 0.97 | 1.00 | 0.96 |
| | 0.80 | 0.60 | 0.76 | 1.05 | 0.80 | 1.06 | 1.07 | 1.18 | 1.44 | 1.39 | 0.83 | 0.72 | 1.24 | 1.10 | 0.70 | 0.00 | 1.02 | 1.08 | 1.15 |

**Fig. (1).** Matrix of ADAM 10 cleavage of the Ac-X8 library. An acetylated X8 library was incubated with 60 nM human ADAM 10 for 9 hrs at 37°C as described. The normalized data is presented for each amino acid (top row) at each prime-side position (left column). The preferred amino acids are printed in Italic.

**ADAM10**

**NXXXXLKXXK(biotin)** **Ac-X8**

| P4 | | P3 | | P2 | | P1 | | P1' | | P2' | | P3' | | P4' | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***R*** | ***1.76*** | Y | 1.52 | ***Q*** | ***1.55*** | ***Y*** | ***2.11*** | ***L*** | ***3.08*** | ***K*** | ***2.29*** | M | 1.53 | M | 1.14 |
| E | 1.52 | A | 1.44 | Y | 1.39 | A | 1.68 | Q | 2.04 | H | 1.73 | N | 1.42 | K | 1.13 |
| D | 1.28 | ***P*** | ***1.31*** | E | 1.38 | R | 1.41 | Y | 1.71 | M | 1.69 | W | 1.36 | F | 1.12 |
| Q | 1.24 | H | 1.28 | M | 1.32 | S | 1.39 | M | 1.55 | V | 1.30 | G | 1.19 | W | 1.11 |
| N | 1.09 | E | 1.26 | A | 1.26 | Q | 1.34 | I | 1.43 | R | 1.18 | A | 1.06 | R | 1.09 |
| Y | 1.08 | M | 1.14 | S | 1.24 | H | 1.23 | F | 1.40 | L | 1.09 | R | 1.05 | G | 1.07 |
| H | 1.06 | R | 1.14 | R | 1.18 | L | 1.09 | W | 1.12 | Y | 1.08 | Q | 1.02 | L | 1.07 |
| A | 1.04 | S | 1.12 | H | 1.17 | M | 1.08 | V | 1.11 | Q | 1.04 | I | 1.00 | P | 1.06 |
| M | 0.99 | N | 1.10 | N | 1.02 | F | 1.01 | R | 1.10 | I | 0.98 | Y | 0.96 | Y | 1.04 |
| P | 0.99 | Q | 1.03 | L | 0.97 | P | 1.01 | H | 0.62 | F | 0.82 | H | 0.92 | A | 1.01 |
| V | 0.94 | F | 1.01 | G | 0.87 | N | 0.96 | E | 0.51 | W | 0.82 | L | 0.92 | H | 1.00 |
| K | 0.94 | V | 0.95 | D | 0.84 | K | 0.82 | K | 0.45 | S | 0.69 | K | 0.91 | E | 0.96 |
| S | 0.93 | D | 0.83 | P | 0.78 | E | 0.77 | N | 0.34 | A | 0.68 | S | 0.91 | Q | 0.96 |
| I | 0.85 | I | 0.79 | F | 0.78 | V | 0.63 | P | 0.32 | N | 0.62 | F | 0.90 | I | 0.95 |
| L | 0.79 | G | 0.71 | V | 0.76 | D | 0.50 | A | 0.31 | P | 0.54 | V | 0.79 | V | 0.93 |
| F | 0.76 | L | 0.69 | I | 0.74 | G | 0.44 | D | 0.31 | E | 0.53 | E | 0.71 | N | 0.83 |
| G | 0.69 | K | 0.61 | K | 0.65 | I | 0.43 | S | 0.31 | G | 0.49 | P | 0.70 | S | 0.78 |
| W | 0.07 | W | 0.06 | W | 0.10 | W | 0.11 | G | 0.28 | D | 0.43 | D | 0.66 | D | 0.76 |
| T | 0.00 | T | 0.00 | T | 0.00 | T | 0.00 | T | 0.00 | T | 0.00 | T | 0.00 | T | 0.00 |

**Fig. (2).** Complete matrix for human ADAM-10. Human ADAM-10 was incubated with the focused library (NXXXXLKXXK-biotin) for 24 hours at 37°C as described. The uncleaved and prime side of the substrate was removed with streptavidin beads and the remaining peptides were subjected to n-terminal sequencing. In the left four columns, indicating the non-primed side of the peptide, the normalized data is presented next to each amino acid for each substrate position (top row). Amino acids chosen for the optimal substrate are shown in Italic. The data from Figure 1 (Primed side) has been reconfigured and placed next to the non-primed side data so that the complete Consensus Matrix™ can be visualized. The nomenclature of Schechter and Berger [29] is used to denote positions in substrates at the amino terminal (unprimed) and carboxy terminal (primed side) of the cleavage site.

cycle, P3 preference information in the second cycle, etc. The data so obtained (Figure 1) is then assembled into the Consensus Matrix ™ of amino acid preference information for each subsite in the proteinase (Figure 2).

The OPML approach has been used successfully to determine substrate subsite preferences for a wide variety of, kinases and protein binding domains [10-12], and recently has been extended to proteases [8]. This approach provides a matrix of information describing the preferred and disfavored

residues at each position of the substrate. Hence, one can rapidly and simply, using a minimum number of libraries, obtain a large amount of detailed structural information regarding the enzyme target. This is valuable not only for designing substrates and inhibitors for the enzyme, but also, when used in combination with matrix information of related enzymes, for providing selectivity information. This powerful tool focuses on the positions and modifications necessary to give the enzyme selectivity for a particular molecule being designed, be it as a drug or tool. To date, this technology has primarily utilized the 20 proteinogenic amino acids. In this paper we report an important extension of this technology wherein a set of structurally diverse, unnatural amino acid probes is used to provide more detailed information regarding a particular binding pocket of interest.

As an example, we report a study of the active site requirements of the metalloproteinase ADAM-10. The ADAM family (A Disintegrin And Metalloproteinase) shares structural homology to certain snake venoms and certain of its members appear to play major roles in modifying the surface exposure of critical signaling molecules [13]. Human ADAM-10 is a 59 kDa, single chain metalloproteinase [14,15] initially isolated as a putative Tumor Necrosis Factor-α Convertase (TACE). Although the physiological TACE now is felt to be ADAM-17 [16], it has been suggested that ADAM 10 may play a major role as a "sheddase", causing release of proteins from the cell surface [17]. Although some potential natural substrates have been suggested [18,19], no optimal *in vitro* substrate has been reported.

In this paper we have analyzed the active site preferences for and against each proteinogenic amino acid and identified a suitable peptide substrate for ADAM-10. We then extended these studies to further probe the size of the P1 binding pocket through the use of OPML containing unnatural amino acids.

## EXPERIMENTAL PROCEDURES

### Reagents

Fmoc-protected amino acids, peptide synthesis resins, and coupling reagents were purchased from Novabiochem (San Diego, CA) or Chem-Impex International (Woodale, IL). PalPeg resign was purchased from Applied Biosystems. Other peptide synthesis reagents were purchased from the Sigma-Aldrich Corporation (St. Louis, MO). Solvents for synthesis and HPLC were JT Baker Brand. Activated human ADAM-10 was purchased from R&D Systems (Minneapolis, MN). Streptavidin beads are from Upstate Biotechnology (Lake Placid, NY). All other Chemicals for making buffers are obtained from Sigma-Aldrich Corporation (St. Louis, MO). All Edman (Procise) sequencing and 433 synthesis reagents are from Applied Biosystems, Inc. (Foster City, CA).

### Library Synthesis

All fully degenerate libraries, focused libraries and unnatural amino acid libraries were synthesized using an ABI 433A Peptide Synthesizer™ (Perkin Elmer-Applied Biosystems) on a 0.25 mmole scale using Rink Amide MBHA resin (Nova Biochem). The degenerate positions, 'X', were attached using experimentally determined isokinetic mixtures of 19 natural amino acids (Cys deleted) containing the following ratios of Fmoc amino acids: 1.0 (Ala, Gly, Leu, Met); 1.3 (Phe, Tyr(*t*-Bu)); 1.7 (Asn(Trt), Asp(*t*-Bu), Glu(*t*-Bu), Gln(Trt), Lys(*t*-Boc), Pro); 2.0 (Arg(Pmc), His(Trt), Ile, Val, Ser(*t*-Bu), Thr(*t*-Bu)); 2.3 (Trp(*t*-Boc)).

For each coupling step 4 equivalents (1.0 mmole) of isokinetic mixture or a single amino acid was weighed into an Applied Biosystems Amino Acid Cartridge™ (ca. 440 mg/cartridge). Each library was made using the standard *FastMoc 0.25 mmol* chemistry protocol with *Single Couple* protocols provided in the Applied Biosystems software. Acetylation of the degenerate libraries was conducted using 0.3 M N-acetylimidazole for 3 hours at room temperature. Complete acetylation was determined through ninhydrin testing. The focused libraries were synthesized in the same manner with the exception of the addition of the first amino acid. Fmoc-Lys(biotinyl)-OH (ChemImpex CAT # 4988) was coupled manually to the resin using 1.0 mmole amino acid and HBTU with 2.0 mmole DIEA in DMF. The reaction was done in a 10 mL syringe with a polystyrene frit and allowed to shake overnight at room temperature. The resin was then loaded onto the 433 Synthesizer and the remaining amino acids were added automatically.

The unnatural amino acid focused library was synthesized using the same procedures as the focused libraries. The ratios of the unnatural Fmoc-protected amino acids in the isokinetic mixture were: 1.0 (Phe, hPhe); 1.1 (Cha); 1.2 (Nal(2)); 1.3 (Bip, Phg); 1.4 (1-Nal, Chg).

Once the resin synthesis was complete, the libraries were cleaved from the solid support and deprotected using a TFA/anisole/water/triisopropyl silane (95:2:2:1) solution and allowed to stand for four hours with occasional stirring [20]. The mixture was filtered and the resin washed with TFA (2 X 2 mL). The combined eluant was cooled and the libraries were then precipitated with $Et_2O$. The mixture was centrifuged and the $Et_2O$ decanted from the precipitate. The precipitate was triturated with $Et_2O$ two more times. After the final $Et_2O$ wash, the libraries were dried under $N_2$, dissolved in 2 mL HOAc and lyophylized. The lyophilized solid was used directly in the enzyme incubations.

### Substrate Synthesis

PalPeg resin (4.5 g) was put into a 50 ml syringe with a frit on the bottom, washed 3X with DMF and 3X with $CH_2Cl_2$. The Fmoc group was removed with three washes with 20% PIP/DMF and tested with ninhydrin. Fmoc-Arg-(Pbf)-OH was added to the resin (2.6 g) using HBTU and DIEA and shaken for 1 hour at room temperature. Deprotection was performed as above. Fmoc-Lys(Dabcyl)-OH (2.5 g) was added using HBTU and DIEA and the reaction was shaken overnight at room temperature. A ninhydrin test confirmed the coupling. The resin was then placed on a Millipore 9050 peptide synthesizer to couple the amino acids up to the Glu-EDANS automatically, using standard protocols. The Glu-EDANS and final R residue were added manually as above. The peptide was cleaved from the resin with 20% PIP/DMF and processed as above

(library synthesis). Each peptide was purified by rp HPLC using a $CH_3CN$/0.05% TFA gradient system to > 90%.

### Enzyme Incubation for Matrix Determination

The reaction mixtures containing 3 mg/ml of peptide library solutions (1-2 mM) and 60 nM of ADAM 10 in 50 mM HEPES, pH 7.5, 5 μM $ZnCl_2$, 0.01% brij-35 were incubated at 37ºC. 20 μL of aliquots were taken out and quenched by 3 μL of 5% TFA at a series of time points of 0 min, 10 min, 30 min, 1 hr, 2hr, 4 hr, 9hr, and 24hr.

For the N-terminal acetylated, totally degenerate peptide libraries, 10 μL aliquots from certain time points (1-5 % cleavage) were sequenced.

For the biotin-containing, focused libraries, 1 μL of the aliquot at different time points was sequenced for 1-2 cycles to determine the percentage of cleavage. The appropriate samples containing 3-5% cleavage were incubated with streptavidin beads to remove the C-terminal cleaved product and uncleaved peptide. Briefly, 20 ul of the stopped enzyme reaction was incubated with 400 ul washed (3X PBS and 3X $NH_4HCO_3$) streptavidin beads with 400 μL $NH_4HCO_3$ for one hour at 4°C. Following incubation, the reaction was transferred to BioRad disposable columns (Cat #731-1550) and the flowthrough was collected. Each column was washed 4 X with 150 ul $NH_4HCO_3$. The flowthrough and washes for each sample were combined and dried on a SpeedVac, washed 2X with 100 ul 10% HOAc, and dried between each wash. Following the final wash, the samples were resuspended in 10 μL of sequencing buffer (20% $CH_3CN$, 0.1% TFA) and the entire sample was loaded onto the sequencer.

### Sequencing

Samples were sequenced using Edman chemistry on an Applied Biosystems Procise 494 Protein Sequencer. Peptide libraries (typically 200 to 500 pmol) were loaded on the sequencer according to standard protocols for the instrument. HPLC purification of each cycle was carried out from 0-48% buffer B ($CH_3CN$/2-Propanol; Applied Biosystems #401570) at 1% per minute. For the peptide libraries containing unnatural amino acids, a longer gradient was developed. This gradient was extended to 51% buffer B at 1% per min. The flask cycles were also modified using standard Procise protocols to ensure enough equilibration time for HPLC separation.

### Data Analysis

The relative preference for each amino acid in a given cycle (A(ij)) was calculated by dividing the amount of that particular residue (in pmol) after digestion (D(ij)) by the relative amount of that particular amino acid, at that particular position, in the starting mixture library (R(ij)). This can be expressed by: (A(ij)=D(ij)/ R(ij)).The data was then normalized to the total number of amino acids used in each X position, (An(ij)), by the equation: (An(ij)=(A(ij)/sum(Aij)) x total number of AA). Sum A(ij) equals the sum of all amino acids A(ij)s at that position. The values grater than one indicate a preference above the average [21].

For the unprimed side an additional value was added to the equation (B(ij)) to account for the remaining undigested library following streptavidin purification. A(ij)=[D(ij)-B(ij)]/R(ij). The data was then normalized as above.

For the non-natural amino acid containing peptide libraries, HPLC peak area was used.

### Enzyme Kinetics

The ADAM-10 enzyme solution (20 nM in 25 mM TRIS pH 9.0, 2.5 uM $ZnCl_2$, 0.005% Brij-35) and substrate dilutions (10-160 uM in above buffer) were equilibrated to 37°C for 10 minutes prior to mixing. ADAM-10 (10 nM) was incubated with the consensus substrate or the ADAM-17 substrate at various concentrations (5-80 uM) in a Molecular Devices fMax Gemini 96-well plate reader. The reader was set to Ex 350 nm/ Em 490 nm, kinetic read every 9 seconds for 20 minutes, mixing between reads. Data was analyzed using the SoftMax Pro software and curves were generated using the Michaelis Menten equation with KaleidaGraph software. Substrate concentrations were corrected for peptide content, determined by amino acid analysis.

## RESULTS

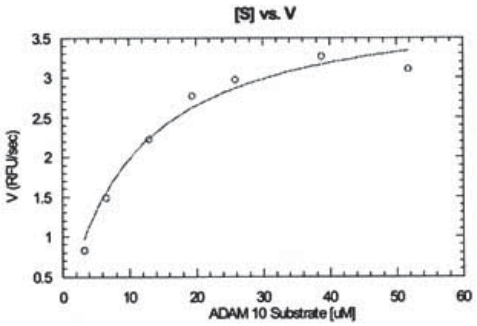
### Matrix Determination

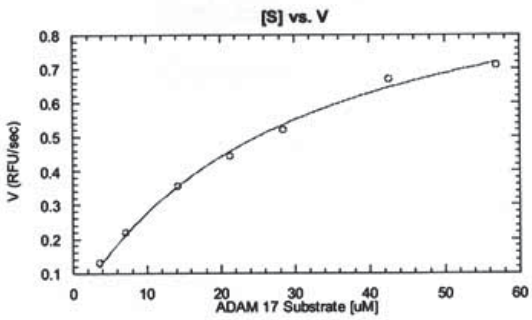
Initial screening of an Ac-XXXXXXXX-$NH_2$ library (where X = 19 natural L-amino acids excluding Cys) with ADAM-10 gave a matrix shown in Figure 1. The Ac on the N-terminal blocks Edman sequencing of the uncleaved peptides in the mixture and of the P-side of the cleaved substrates. This allows only the P' side of the most optimal substrates (first 1-5% cleaved) to be available for sequencing. A strong preference for Leu at $P_1$' and a weaker preference for Lys at $P_2$' can be seen in the figure. There was limited amino acid selectivity at positions $P_3$' and $P_4$' as indicated by each amino acid being represented by a number close to 1.0 (normalized data), indicating that such amino acid residues are neither favored nor disfavored at this position in a good substrate. Thr was not taken into consideration for these experiments due to a large contaminating buffer peak co-eluting at the Thr position.

Based on this result, a focused library $NH_2$XXXX-LKXXK(biotin)-$NH_2$ (where X = 19 natural L-amino acids excluding Cys) was synthesized and incubated with ADAM-10 as described. Following removal of the C-terminal fragments and uncleaved peptides using streptavidin beads, the resulting matrix for positions P4-P1 can be seen in Figure 2. At the P1 site a preference can be seen for Tyr. At the P2 position, Gln is the most represented amino acid. The P3 position shows a weak preference for Tyr, Ala and Pro. Since Pro is generally not tolerated in most positions, when it is tolerated, as identified in position P3 for ADAM 10, we feel it is a good idea to use it in the peptide to help orient the peptide in the binding site. Finally, at the P4 position, Arg is preferred.

### ADAM 10 Consensus Substrate Analysis

To generate the consensus peptide, we brought together the information from both the P- side and the P'. Because




| Substrate | Km (uM) | Vmax uM/sec | kcat/Km $M^{-1} s^{-1}$ |
|---|---|---|---|
| ADAM-10 | 10.24 | 0.034 | 380,000 |
| ADAM-17 | 29.07 | 0.009 | 36,000 |

**Fig. (3).** Determination of Km and Vmax for ADAM 10 consensus sequence. A peptide representing the determined consensus sequence for ADAM-10 ($NH_2$-D-Arg-Arg-Glu(Edans)-Arg-Pro-Gln-Tyr-Leu-Lys-Ser-Ser-Lys(Dabcyl)-Arg-$NH_2$) was compared to a known ADAM-17 substrate ($NH_2$-Arg-Glu(Edans)-Leu-Ala-Gln-Ala-Val-Arg-Ser-Ser-Lys(Dabcyl)-Arg-$NH_2$). Human ADAM 10 (20 nM) was incubated with each substrate at varying concentrations as described. The Km and Vmax were determined by plotting the Michaelis-Menten equation using KelidaGraph software.

there was no particular selectivity determined in the matrix for positions P3' and P4', and considering there are similarities between ADAM-10 and ADAM-17 [15,22], we chose to use the relatively innocuous amino acid Ser for those positions in our substrate peptide. Polar residues would be preferred since we interpret the lack of selectivity at these positions as their being present in the bulk aqueous phase when bound to the enzyme. The known substrate for ADAM-17 contains Ser at positions P3' and P4' [23].

## A Peptide Containing the Putative Consensus Sequence for ADAM-10

($NH_2$-D-Arg-Arg-Glu(Edans)-Arg-Pro-Gln-Tyr-Leu-Lys-Ser-Ser-Lys(Dabcyl)-Arg-$NH_2$) was synthesized and subjected to a series of kinetic analyses (Figure 3). For comparison, a known substrate for ADAM-17 ($NH_2$-Arg-Glu(Edans)-Leu-Ala-Gln-Ala-Val-Arg-Ser-Ser-Lys(Dabcyl)-Arg-$NH_2$) (Bachem #M-2155) was characterized. The Km for the ADAM-10 substrate was determined to be 10 uM as compared to almost 30 uM for the ADAM-17 substrate. The Vmax for the ADAM 10 substrate was determined to be 0.034 uM/sec yielding a "selectivity factor" (kcat/Km) of 380,000 $M^{-1}s^{-1}$. This can be compared to the Vmax for the ADAM-17 substrate of 0.009 uM/sec which yields a selectivity factor of 36,000 $M^{-1}s^{-1}$. By this analysis, ADAM 10 has a 10 fold greater selectivity for the ADAM-10 substrate identified using the OPLM method over the ADAM 17 substrate.

## Unnatural Amino Acid Library Design

Having determined that the substrate generated from the ADAM 10 matrix was an efficient substrate (kcat/Km = 380,000 $M^{-1}s^{-1}$), we were interested in probing further the character of favored P1 substitutions in ADAM-10. The identification of Tyr as the favored residue at P1 suggested that the S1 pocket (site in the enzyme that interacts with the P1 position of the substrate peptide) might be a pocket with significant depth. Since such pockets may have great significance for inhibitor design, we decided to explore the nature of the S1 pocket using an OPML with a diverse set of unnatural lipophilic and aromatic amino acids at the P1 position.

Since the technology being used to probe the P1 binding site relies on Edman sequencing of the digested library mixture, the unnatural amino acids chosen need to be identifiable in standard sequencing protocols. Hence, the choice of amino acids was based on a balance of structural diversity and their ability to be separated in the HPLC step of the sequencing protocol. The chosen set of unnatural amino acids is shown in Figure 4. As can be seen, they serve as probes of depth [Phe, hPhe, Bip, Nal(2), Phg], width [Nal(1), Cha] and the need for aromaticity (either specific interaction or "flatness") [Cha, Chg]. From our experience with isokinetic mixtures of natural amino acids, we approximated the amount for each unnatural amino acid in the isokinetic mixture. Having an exact equimolar ratio was not essential since the unprocessed library is used as a control and we report the ratio of a particular amino acid in the cleaved library sample to its corresponding level in the unprocessed library. Hence, preferred and disfavored amino acids are reflected properly as the ratio to their abundance in the unprocessed library.

The oriented library made using these unnatural amino acid probes was:

Asp-Arg-Pro-Glu-X-Leu-Arg-Ser-Ser-Lys(Biotin)-$NH_2$, where X represents the isokinetic mixture of amino acids shown in Figure 4.

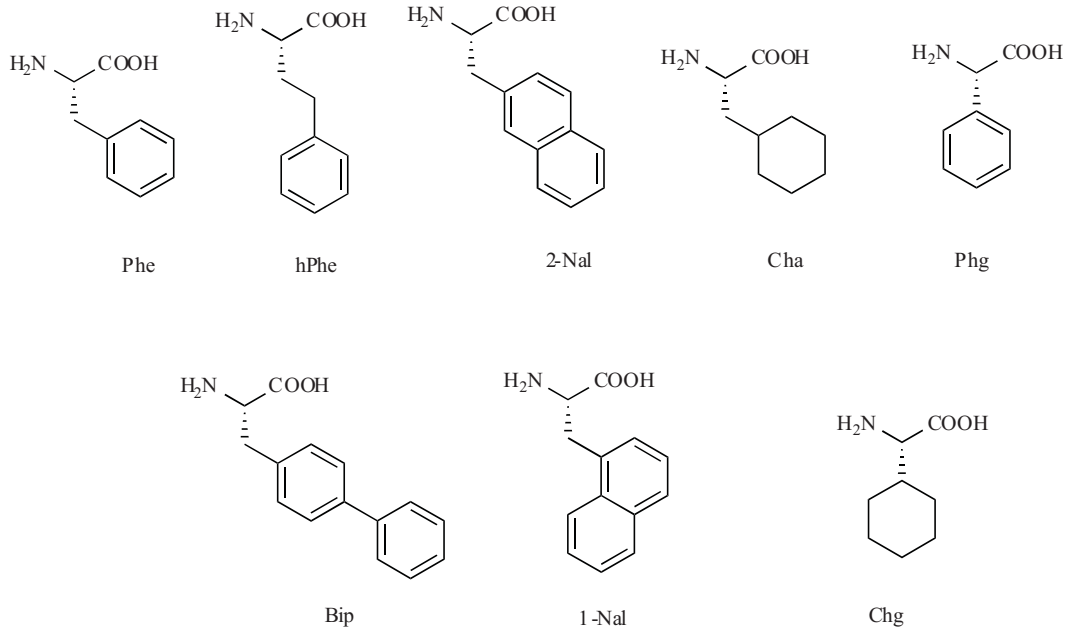



**Fig. (4).** Set of unnatural amino acid structural probes.

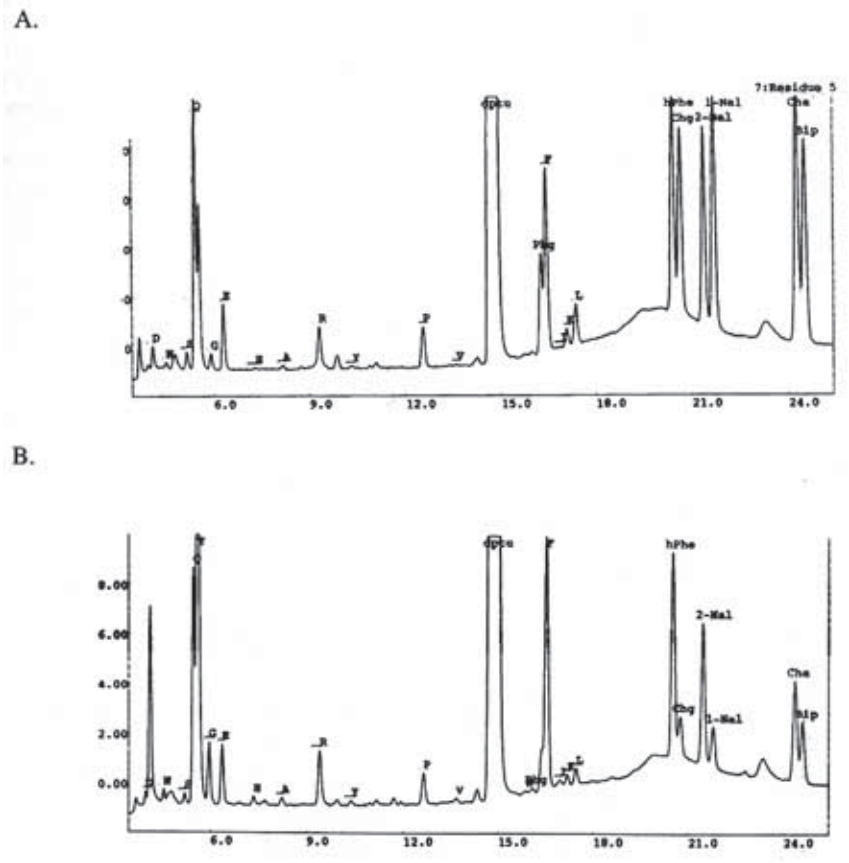



**Fig. (5).** Sequencing at position $P_1$ of library mixture before (A) and after 10 min. incubation with enzyme (B). The focused library Asp-Arg-Pro-Glu-X-Leu-Arg-Ser-Ser-Lys(Biotin)-$NH_2$ was synthesized where X represents an isokinetic mixture of the amino acids shown in Figure 4. This library was incubated with ADAM-10 for 30 minutes at 37 °C followed by purification and sequencing as described. The cleaved library is represented in the B profile, while the starting library is represented in the A profile.

This library was subjected to the standard enzyme cleavage protocol and processed in the same way as the previously described samples.

### $P_1$ Preference Determination Using the Unnatural Amino Acid Library

ADAM-10 was incubated with the above focused, oriented library for 1 min. Following removal of the uncleaved peptides and the C-terminal fragments using a streptavidin column the sample was sequenced as described. The HPLC profile is shown in Figure 5B for the cleaved library and in Figure 5A for the uncleaved, reference library. The peak area was determined for each unnatural amino acids in both profiles (data not shown). The ratio of peak area B/ peak area A was determined and is shown in Table 1. Amino acids that were strongly favored in the P1 position, as indicated by a value of greater than 1, include Phe and hPhe. Na1(2) was also favored, although to a lesser extent. The amino acids that were tolerated include Cha, Phg and Bip. The amino acids that were strongly disfavored were the 1-Nal and Chg.

**Table 1. Preference Ratios of the Amino Acids at the $P_1$ Position of the Oriented Mixture Substrate. Area Under the Curve for each Unnatural Amino Acid Shown in Figure 5 was Determined. The Difference was Calculated by the Ratio Area Peak B/Area Peak A.**

| | | |
|---|---|---|
| **FAVORED** | | |
| Phe | hPhe | 2-Nal |
| 2.95 | 1.72 | 1.29 |
| **TOLERATED** | | |
| Cha | Phg | Bip |
| 0.67 | 0.48 | 0.47 |
| **DISFAVORED** | | |
| 1-Nal | Chg | |
| 0.25 | 0.17 | |

## DISCUSSION

It is clear that the OPML analysis of proteinase substrate requirements can be used to accelerate inhibitor drug discovery programs. As is apparent from the data in Figures 2 and 3, this approach can provide a comprehensive analysis of proteinase active site preferences very rapidly with the use of only two libraries. The consensus substrate motif is read off the tops of the columns (most abundant, i.e. most favored residues at each subsite) in the Consensus Matrix™. Several peptide substrates, incorporating residues in positions of high selectivity near the top of the columns, are generally made to investigate fine tuning of the behavior of the substrate for choice of a reagent for HTS assay development. The substrate chosen for synthesis in this study has a very high selectivity factor (380,000 $M^{-1}s^{-1}$), suggesting that it is likely to be an excellent substrate for assay development. The Consensus Matrix™ can also be used to design selective substrates, since it also indicates which residues are strongly disfavored at critical subsites. Judicious choice of residues allowed in a subsite for one proteinase, but disfavored for another proteinase can be used to derive high selectivity.

The OPML analysis is also used to design first generation proteinase inhibitors. Since there are well accepted methods for going directly from substrate identification to peptide inhibitor design for each of the four major classes of proteinases [1], the information in the Consensus Matrix™ can be used to rapidly generate first generation inhibitors for use in target validation and x-ray crystal structure studies.

The physiological role and substrates for ADAM-10 are a subject of much scientific interest. Although ADAM-10 was originally isolated as a putative TNF convertase, ADAM-17 now is generally accepted to play this role. Mice with an ADAM-10 knockout genotype die in day 9.5 of embryogenesis due to multiple defects in the CNS and cardiovascular system [24]. ADAM-10 and –17 are felt to play a key role in the activation of the Notch signaling pathway [25] and in the benign processing of the Amyloid Precursor Protein through “alpha secretase” cleavage at a site near the outer cell membrane [26].

Metalloproteinases typically exhibit strong selectivity dependence on the residue at position P1’ [27]. The presence of Leu or Tyr is generally taken as evidence of a shallow or deep S1’ site, respectively [28]. For ADAM-10 we find Leu preferred and Tyr accepted in the P1’ position. We were intrigued to find Tyr favored at P1 and took it as an indication that P1 might be a deep pocket. In order to investigate the characteristics of this subsite we studied a focused OPML that incorporated a mixture of unnatural, hydrophobic amino acids at the P1 position. The amino acid mixture incorporated in P1 was chosen to test the depth and width of S1 binding subsite (Figure 4; discussed above). Pth-amino acid elution conditions were developed which allowed quantitation of these individual amino acids in a single sequencing run (Figure 5). As outlined in Table 1, amino acids of considerable length could be accommodated in S1 (hPhe, Nal(2), Bip), but wide or “fat” amino acids (e.g. Nal(1) and Cha; Chg, respectively) were not accepted. Additional probing of this or other subsites with amino acid mixtures designed to ask other questions (e.g. electron rich or poor heterocyclic side chains) can be visualized readily. This type of information would be valuable in the design of inhibitors of ADAM-10.

We have provided another illustration of how the OPML approach [8, 30] to the analysis of the optimal substrate requirements of a new enzyme can be executed very rapidly with only two libraries, providing a very powerful matrix of information for substrate and inhibitor design. We have further shown that the approach can be extended to the use of unnatural amino acid mixtures for detailed probing of active site binding character. The rapid development of a suitable HTS assay for a novel protease, as well as inhibitors for target validation and x-ray studies, provides a comprehensive approach for the acceleration of drug discovery.

## ACKNOWLEDGEMENTS

The advice and guidance of the developers of the OPML approach to proteinase substrate determination (Prof. Lewis Cantley and Dr. Benjamin Turk) is gratefully acknowledged.

## ABBREVIATIONS

| | | |
|---|---|---|
| Bip | = | 3-(4-biphenyl)-L-alanine |
| Cha | = | 3-cyclohexyl-L-alanine |
| Chg | = | L-cyclohexylglycine |
| hPhe | = | L-homophenylalanine |
| HTS | = | high throughput screening |
| 1-Nal | = | 3-(1-naphthyl)-L-alanine |
| 2-Nal | = | 3-(2-naphthyl)-L-alanine |
| Phg | = | L-phenylglycine |
| Fmoc | = | Fluorenylmethyloxycarbonyl |
| DIEA | = | Di-iso-propylethyl amine |
| HBTU | = | 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate |
| DMF | = | *N*,*N*-dimethylformamide |
| FRET | = | Fluorescence resonance energy transfer |
| TFA | = | Trifluoroacetic acid |

## REFERENCES


[1] Leung, D., Abbenante, G., Fairlie, D.P. *J. Med. Chem.* **2000**, *43*, 305-41.

[2] 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*, 149-155.

[3] Harris, J.L., Backes, B.J., Leonetti, F., Mahrus, S., Ellman, J.A., Craik, C.S. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7754-7759.

[4] Matthews, J., Wells, J.A. *Science* **1993**, *260*, 1113-1117.

[5] Smith, M.M., Shi, L., Navre, M. *J. Biol. Chem.* **1995**, *270*, 6440-6449.

[6] Birkett, A.J., Soler, D.F., Wolz, R.L., Bond, J.S., Wiseman, J., Berman, J., Harris, R.B. *J. Analyt. Biochem.* **1991**, *196*, 137-148.

[7] Petithory, J.R., Masiarz, F.R., Kirsch, J.F., Santi, D.V., Malcolm, B.A. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 11510-11514.

[8] Turk B.E., Huang L.L., Piro E.T., Cantley L.C. *Nat. Biotechnol.* **2001**, *19*, 661-667.

[9] Ivanetich, K.M., Santi, D.V. *Methods Enzymol*. **1996**, *267*, 247-60.

[10] Sonyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., Neel, B.G., Birge, R.B., Fajardo, J.E., Chou, M.M., Hanafusa, H., Schaffhausen, B., Cantley, L.C. *Cell* **1993**, *72*, 767-778.

[11] Sonyang, Z., Blechner, S., Hoagland, N., Hoekstra, M.F., Piwnica-Worms, H., Cantley, L.C. *Current Biol.* **1994**, *4*, 973-982.

[12] O'Neill T., Dwyer A.J., Ziv Y., Chan D.W., Lees-Miller S.P., Abraham R.H., Lai J.H., Hill D., Shiloh Y., Cantley L.C. & Rathbun G.A. *J. Biol. Chem.* **2000**, *275* (30) 22719-22727.

[13] Primakoff, P., Myles, D.G. *Trends Genet.* **2000**, *16* (2), 83-87.

[14] Rosendahl, M.S., Ko, S.C., Long D.L., Brewer, M.T., Rosenzweig, B., Hedl, E., Anderson, L., Pyle, S.M., Moreland, J., Meyers, M.A., Kohno, T., Lyons, D., Lechenstein, J.S. *J. Biol. Chem*. **1997**, *272* (39), 24588-24593.

[15] Moss, M.L., Jin, S.L., Milla, M.E., Bickett, D.M., Burkhart, W., Cater, H.L., Chen, W.J., Clay, W.C., Didsbury, J.R., Hassler, D., Hoffman, C.R., Kost, T.A., Lambert, M.H., Leesnitzer, M.A., McCauley, P., McGeehan, G., Mitchell, J., Moyer, M., Pahel, G., Rocpue, W., Overton, L.K., Schoenen, F., Seaton, T., Su, J.L., Becherer, J.D., Warner, J., Willard, D. *Nature* **1997**, *385*, (6618), 733-736.

[16] Black, R.A., Rauch, C.T., Kozlosky, C.J., Peschon, J.J., Slack, J.L., Wolfson, M.F., Castner, B.J., Stocking, K.L., Reddy, P., Srinivasan, S., Nelson, N., Boiani, N., Schooley, K.A., Gerhart, M., Davis, R., Fitzner, J.N., Johnson, R.S., Paxton, R.J., March, C.J., Cerretti, D.P. *Nature* **1997**, *385*, (6618), 729-733.

[17] Moss, M. Lambert, M.H. *Essays Biochem.* **2002**, *38*, 141-153.

[18] Lammich, S. Kjro, E., Postina, R., Gilbert, S., Pfeiffer, R., Jasionowski, M., Haass, C., Fahrenholz, F. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3922-3927.

[19] Qi, H., Rand, M.D., Wu, X., Sestan, N., Wang, W., Rakic, P., Xu, T., Artavanis-Tsakonas, S. *Science* **1999**, *283*, 91-94.

[20] Solé, N.A., Barany, G. *J. Org. Chem.* **1992**, *57*, 5399-5403.

[21] Songyang, Z., Cantley, L.C. *Methods Enzymol.* **1995**, *254*, 523-535.

[22] Maskos, K., Fernandez-Catalan, C., Huber, R., Bourenkov, G.P., Bartunik, H., Ellestad, G.A., Reddy, P., Wolfson, M.F., Rauch, C.T., Castner, B.J., Davis, R., Clarke, H.R.G., Petersen, M., Fitzner, J.N., Cerretti, D.P., March, C.J., Paxton, R.J., Black, R.A., Bode, W. *Proc. Natl. Acad. Sci USA* **1998**, *95*, 3408-3412.

[23] Mohler, K.M., Sleath, P.R., Fitzner, J.N., Cerretti, D.P., Alderson, M., Kerwar, S.S., Torrance, D.S., Otten-Evans, C., Greenstreet, T., Weerawarna, K., Kronheim, M.P., Gerhart, M., Kozlosky, C.J., March, C.J., Black, R.A. *Nature* **1994**, *370*, (6486) 218-220.

[24] Hartmann, D., de Strooper, B., Serneels, L., Craessaerts, K., Herreman, A., Annaert W., Umans, L., Lubke T., Lena Illert, A., von Figura, K., Saftig, P. *Hum. Mol. Genet.* **2002**, *11*, 2615-2624.

[25] Hartmann, D., Tournoy, J., Saftig, P., Annaert, W., de Strooper, B. *J. Mo. Neurosci.* **2001**, *17*, 171-181.

[26] Marcinkiewicz, M., Seidah, N.G. *J. Neurochem*. **2000**, *75*, 2133-2143.

[27] Johnson, L.L., Dyer, R., Hupe, D.J. *Curr. Op. Chem. Biol.* **1998**, *2*, 466-471.

[28] Lovejoy, B., Welch, A.R., Carr, S., Luong, C., Broka, C., Hendricks, R.T., Campbell, J.A., Walker, K.A., Martin, R., Van Wart, H., Browner, M.F. *Nature Struct. Biol.* **1999**, *6,* 217-21.

[29] Schecter, I., Berger, A. *Biochem. Biophys. Res. Comm.* **1967**, *27*, 157-162.

[30] Turk, B.E., Cantley, L.C. *Curr. Opin. Chem. Biol.* **2003**, *7*, 84-90.